大學圖書館
14卷1期（民國99年3月） 頁37-57

臺灣學術機構典藏共享平台之建置

Constructing a Sharing Platform for Academic Institutional Repositories in Taiwan

陳光華 Kuang-hua Chen
國立臺灣大學圖書資訊學系副教授兼圖書館副館長
Associate Professor, Department of Library and Information Science, National Taiwan University
Deputy Director, National Taiwan University Library
E-mail: khchen@ntu.edu.tw

陳彥良 Yen-liang Chen*
臺灣積體電路製造股份有限公司12廠電腦整合製造工程師
Computer Integrated Manufacturing Engineer,
Taiwan Semiconductor Manufacturing Company Fab 12
E-mail: r93546001@ntu.edu.tw

【摘　要】

機構典藏系統保存學術機構的研究成果，並作為使用者取得學術研究成果的新管道，已是現代學術圖書館必須積極推動的資訊服務。目前臺灣各大專校院在教育部的推動下，紛紛建置各自的機構典藏系統，已充分展現各自的研究特色與學術影響。為了以臺灣作為一個主體，向國際學術社群展現臺灣整體的學術成果與研究能量，建置一機構典藏共享平台，被視為現階段臺灣機構典藏發展的重點。「臺灣學術機構典藏（Taiwan Academic Institutional Repository，簡稱TAIR）」的建置與發展，是遵循著「分散建置、整合應用」的精神，建構整合的跨機構的學術成果典藏與共享平台系統，呈現出臺灣學術機構典藏的豐富面貌，共同分享臺灣各大專校院的學術研究成果，最大化臺灣的學術能量。本文將介紹國際上相關平台系統的發展，說明臺灣學術機構典藏共享平台的發展與功能，討論系統平台的建置過程與實作經驗，分享機構典藏共享平台的建置技術、遭遇的困難與挑戰。

【Abstract】

The institutional repository (IR) preserves the research outputs of academic institutions and provides a new channel for accessing scholarly information. It has become a “must-have” information service of any university library or academic library,

收稿日期：民國98年8月11日；接受日期：民國99年1月28日
*於民國97年1月至98年7月曾擔任國立臺灣大學圖書館系統資訊組研究助理。

with many universities creating their own independent repositories. Funded by the Ministry of Education, Taiwanese Institutional repositories enable each university to demonstrate their specialties and academic achievements. To further unify research resources and demonstrate the academic prowess of Taiwan, an “umbrella” organisation was established to integrate the nation’s institutional repositories. The founding ethos of the Taiwan Academic Institutional Repository (TAIR) was“distributed construction, integrated application” with the objective of harvesting resources from all IR systems and then forming an academic portal in order to share research results. This article considers related systems from around the world, examines the features of TAIR and discusses the challenges faced during its development, implementation and technical evolution.

關 鍵 詞：機構典藏；系統開發；開放典藏首創協定
Keywords：Institutional repository; System development; Open Archive Initiative

壹、緒論

機構典藏（Institutional repository，簡稱IR），意指學術機構將內部的研究產出，包括學術論文、研究計畫、課程教材、甚至圖片影像，以數位化的方法保存，並建立網路化系統平台，提供使用者瀏覽與檢索服務。對學術機構而言，不但可以保存歷來的學術產出與研究歷程，也可以了解整體的研究成果；對使用者而言，機構典藏突破了傳統上學術期刊以學術領域為基礎的學術資訊傳播模式，提供了另一個新的學術資訊的傳播管道，並提供完善的全文搜尋功能，讓使用者更方便地取得線上學術研究資訊。線上學術研究資訊的重要性在於使用者可以快速地取得，可降低取得學術研究成果的時間成本，亦有許多研究顯示線上全文論文被引用的次數會顯著提高（Lawrence, 2001；Hajjem, Gingras, Brody, Carr, & Harnad, 2005），可以有效提昇學術影響力。若是全球各學術機構都建置機構典藏系統，學術資訊的傳播可以更加直接，不必完全透過傳統的期刊出版商，更可以大幅取得學術資訊的成本。因此，無論由取得成本、傳播速率、機構定位、國際影響等各個面向思考，建置機構典藏系統都有其合理性、重要性、必要性。若進一步思考，以整個國家為主體，建構一個臺灣機構典藏共享平台系統，展現臺灣整體學術研究成果，更可以讓世界各國的使用者了解臺灣的學術研究結構，以及臺灣的學術研究發展與趨勢。本文在前述的背景與動機之下，說明機構典藏在國際的發展情形與臺灣的發展現況，並討論臺灣

學術機構典藏共享平台的建置與功能。

在2005年時，美國約有40%的大學已經開始營運機構典藏系統，53%的大學則已經著手建置機構典藏系統（Lynch & Lippincott, 2005）；歐洲地區的德國、挪威、荷蘭的機構典藏系統數量已超過其大學總數（Westrienen & Lynch, 2005），表示有單一的大學依不同目的建置多個機構典藏系統；亞洲地區的機構典藏系統數量雖然不如歐美各國普及，日本、印度、香港、臺灣等地也有為數不少的機構典藏系統。此外，已有少數國家開始建立政府層級的機構典藏政策，表示機構典藏已受到世界各國的重視，是目前學術機構積極推動的重要工作項目；也已經有許多第三機構進行機構典藏系統的評鑑工作，例如Rankings Web of World Repositories（RWWR, n.d.）以及Registry of Open Access Repositories（ROAR, n.d.）。

在圖書資訊學界與圖書館界學者與專家的呼籲與建議之下，教育部逐漸了解機構典藏的重要性，遂於2005年5月委託臺灣大學圖書館進行為期三年的「建置機構學術成果典藏計畫」，計畫目標包含：

- 建置機構學術成果典藏系統雛型；
- 規劃建立機構學術成果典藏之永續經營發展機制；
- 規劃並訂定機構學術成果典藏系統需求與標準規範；
- 規劃並訂定建置機構學術成果典藏之標準作業程序。

計畫第一年開發了以DSpace為基礎的機構典藏系統NTUR；第二年推動IR30方案，協助至少30所學術機構建置機構典藏系統；第三年持續推動IR30+方案，並開始規劃機構典藏共享平台。截至2009年12月，已有100所學術機構加入「建置機構學術成果典藏計畫」的行列，並使用臺灣大學機構典藏工作團隊開發的NTUR系統，且有65所學術機構已完成系統建置的工作。以臺灣大學、成功大學、中興大學、政治大學、清華大學、交通大學、中央大學為例，此8所大學機構典藏系統所典藏的文件數量分別為139,173、63,713、42,927、33,441、26,601、25,290、23,655筆。在學術機構安裝了相同的機構典藏系統，或是使用Open Archive Initiate的基礎下，臺灣大學機構典藏工作團隊更進一步地開發臺灣學術機構典藏系統（Taiwan Academic Institutional Repository，簡稱TAIR），其目的是在現有機構典藏系統的基礎之上，建構學術機構典藏共享平台，以便建立單一窗口，提供分類瀏覽與資料檢索的功能，呈現出臺灣整體學術研究成果的豐富面貌，讓世界各國的使用者透過此共享平台系統，共同分享臺灣各學術機構的學術研究成果，並提升臺灣學術研究的聲望與影響力。

臺灣學術機構典藏系統遵循著「分散建置、整合應用」的機制，以單一平台顯示所有參與臺灣學術機構典藏計畫之機構的文件資料，並提供

瀏覽及查詢功能，使用者可經由TAIR系統提供的文件鏈結，連結至該學術機構原有的機構典藏系統檢視文件，或是下載全文檔案。此機構典藏的共享平台提供了以下優點：

- 分散建置。讓各參與機構仍然保持其主體性，發展其研究特色，並展現其研究能量。且外，因為所有全文檔案仍儲存於各參與機構之機構典藏系統，臺灣學術機構典藏系統可節省儲存全文檔案的空間。
- 整合應用。為增加所有的參與機構與典藏文件的可見度，建立臺灣整體研究成果的索引，做為臺灣整體學術研究成果的保存、展示與利用窗口，最大化臺灣學術研究成果的產出，提升在國際的能見度。
- 資源共享。使用者可在單一平台獲取跨學術機構的產出成果，效率遠勝於至各機構典藏系統檢索資料。對所有參與機構而言，使用者利用本系統也相對減少使用電子期刊的次數，可省下部份花費於電子期刊的成本。
- 永續經營。參與此共享平台的學術機構能見度若能大大提升，對未參與機構典藏的學術機構，便提供了參與機構典藏的誘因；對參與機構典藏的學術機構，也成為其永續經營與維持機構典藏的動力。

本文將以系統開發者的角度，說明臺灣學術機構典藏系統的建置過程與實作經驗，分享機構典藏共享平台的建置技術，限於人力與篇幅，本文並沒有進行使用者調查之實證性研究工作。本文的結構說明如下：第二部份說明了目前世界各國中，類似TAIR系統的機構典藏共享平台系統的發展；第三部份將介紹臺灣學術機構典藏系統的功能；第四部份將說明TAIR系統的系統資訊與開發技術；第五部份則敘述建構TAIR系統時所遭遇到的困難與挑戰；最後是簡短的結論。

貳、機構典藏共享平台之發展

就目前所見的學術機構典藏共享平台系統雖然為數不多，但是仍然有許多研究團隊正積極建立學術機構典藏共享平台系統。為了有效地統整學術研究成果，建置機構典藏共享平台勢必是未來的趨勢。機構典藏共享平台就像是一座大傘，涵蓋眾多個別的機構典藏系統，使用者由一個入口即可取得所有參與共享平台之機構典藏系統的文件，不僅便利使用者，同時可以聯合行銷參與之學術機構的研究成果。策略層面上，達到保存研究成果、有效使用研究成果、提升學術影響等多重目標；技術層面上，可以透過書目資料索引或是全文索引，提供不同層次的檢索服務；基礎層面上，聯合各種不同的機構典藏系統，不侷限於個別的機構典藏系統。圖1展示其營運的概念模式。下文介紹重要的機構典藏共享平台系統。

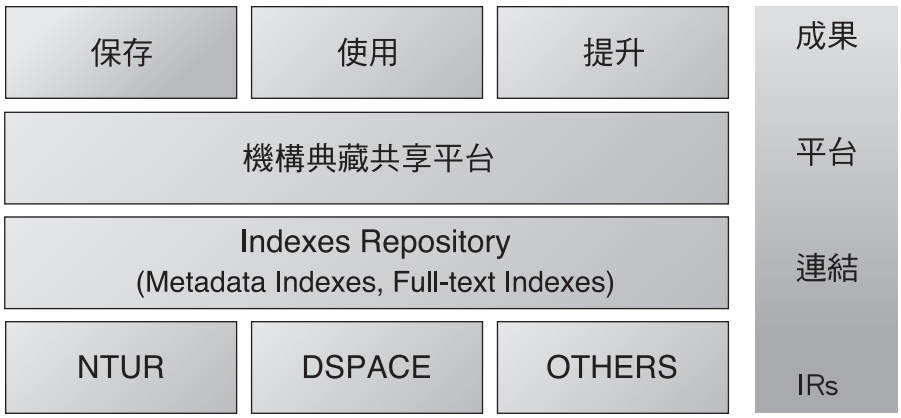


圖 1　機構典藏共享平台之概念模式

一、JAIRO

JAIRO系統（Japanese Institutional Repositories Online）是日本的機構典藏共享平台，雖然日本的學術機構不全然使用相同的機構典藏系統，但是每個機構典藏系統都支援OAI協定，雖然僅能達到書目資料的整合，但是也算是做到部分的「分散建置、整合應用」的目標。

截至2009年三月底，雖然JAIRO系統宣稱仍處於測試階段，但已有89個學術機構加入，累計603,478筆檔案。JAIRO將檔案依類型區分為十五類：期刊文章（journal article）、論文（thesis or dissertation）、學報論文（departmental bulletin paper）、研討會論文（conference paper）、簡報（presentation）、書刊（book）、技術報告（technical report）、研究論文（research paper）、文章（article）、刊前本（preprint）、教學材料（learning material）、資料（data or dataset）、軟體（software）、其他（others）。（JAIRO, n.d.）此外，JAIRO相當注重與使用者的互動，允許使用者匯出Metadata的功能，使用者可先選擇相關文件而後匯出對應之Metadata，可儲存為一般的txt格式、tsv格式、以及EndNote軟體的bib格式，在系統新增文件檔案之後，也透過RSS（Real Simple Syndication）功能提示使用者。

二、OAIster

OAIster系統是一個專門蒐集數位資源的聯盟，截至2009年五月一日前，已從1,112個典藏系統中收藏了20,928,590筆檔案，OAIster

的參與對象不一定是學術機構，其中也包括了期刊系統，也沒有地區的限制，收錄的檔案類型豐富，影音檔案也納入收錄範圍。因為所有參與的典藏系統都使用OAI-PMH（Open Archives Initiative Protocol for Metadata Harvesting）通訊協定，故取名為OAIster。（OAIster, n.d.）OAIster系統是密西根大學安娜堡分校（University of Michigan at Ann Arbor）與伊利諾大學香檳分校（University of Illinois at Urbana-Champaign）合作建置的，在二校具有足夠之技術基礎下，建立了收錄政策及作業機制，規律地抓取資料並轉換資料。自1992年起，OAIster系統不斷地改善搜尋功能，也逐步改善資料搜尋的時間或是資料呈現的格式，目前OAIster系統提供了布林檢索、資料類型檢索、以及排序的功能，而OAIster系統開發團隊仍然持續地改進搜尋功能，期望在未來還能提供全文檢索等更進階的功能。此外，OAIster系統提供了開放式的URL，允許使用者不必經由OAIster系統介面檢索資料，提供使用者做更深入的應用。

三、Ira

Ira專案（Institutional Repositories Aotearoa）乃由紐西蘭的奧克蘭大學（University of Auckland）、坎特伯里大學（University of Canterbury）及維多利亞大學（Universities of Victoria）三所學校合作進行的專案，主要目的是配合紐西蘭國家圖書館的政策，建立紐西蘭地區學術機構典藏的整合檢索平台，提供類似Google Scholar的服務，也藉由此方案向紐西蘭國家圖書館申請經費供後續的學術研究使用。（Ira, n.d.）

因為前述三所大學的機構典藏系統皆使用DSpace，Ira首先詳細介紹DSpace的各項功能，然後說明典藏文件的特色與內容。但是，Ira並不直接使用DSpace內建的檢索功能，也不自行開發檢索功能，而是直接使用Google的Site Search檢索功能，內嵌於Ira系統的使用者介面，讓使用者透過Google搜尋全紐西蘭的機構典藏系統。

參酌國際上的機構典藏的發展趨勢，對應於臺灣目前機構典藏的近程，就臺灣已經建置了機構典藏系統的大學而言，或是就整個臺灣學術研究成果的展現而言，建置整合檢索與共享平台是下一階段機構典藏的重要發展方向，本研究即是探討這項重要的課題，討論系統功能的建立、系統內容的建置、以及分享開發的技術與經驗。

參、TAIR系統之功能

臺灣大學機構典藏團隊開發的NTUR系統，是以DSpace 1.4 Beta版本為基底開發，因此繼承了DSpace系

統具備OAI機制與handle機制兩大功能（DSpace, n.d.），TAIR系統仍然善加利用以上兩種功能。TAIR系統作為各個NTUR系統的整合共享平台，在資料結構與使用介面的設計上也就承襲NTUR系統。TAIR系統的資料結構如圖2所示，延續了NTUR系統「社群、類別、檔案」的資料結構，另在社群之上加入「機構」，以彰顯臺灣各學術機構，遵循「分散建置、整合應用」的精神，既維持個別學術機構的特色，也展現臺灣學術研究機構整體的研究成果，更便於使用者整合檢索研究機構的學術成果。

TAIR系統的使用者介面如圖3所示，仿照NTUR系統的版面配置，以保留使用者相同的使用經驗。首頁上方提供了搜尋功能工具列與語言切換工具列，可供使用者切換英文、簡體中文與正體中文等介面語言，並顯示了累計造訪人次與目前線上人數。左方則為選擇瀏覽模式的功能列、相關連結、以及資料統計功能，資料統計功能分別為上傳排行、下載排行與查詢排行。右方之上方則嵌入Google Map，以展示參與臺灣學術機構典藏計畫的各個學術機構，藉由此地圖可了解臺灣各大專院校參與此計劃的數量與地理位置分佈情形；右方的下方展示簡易的資料比數分布圖，說明系統內容資料的分佈情形。畫面中央則是參與機構典藏計畫的機構清單。詳細的功能將於下文說明。

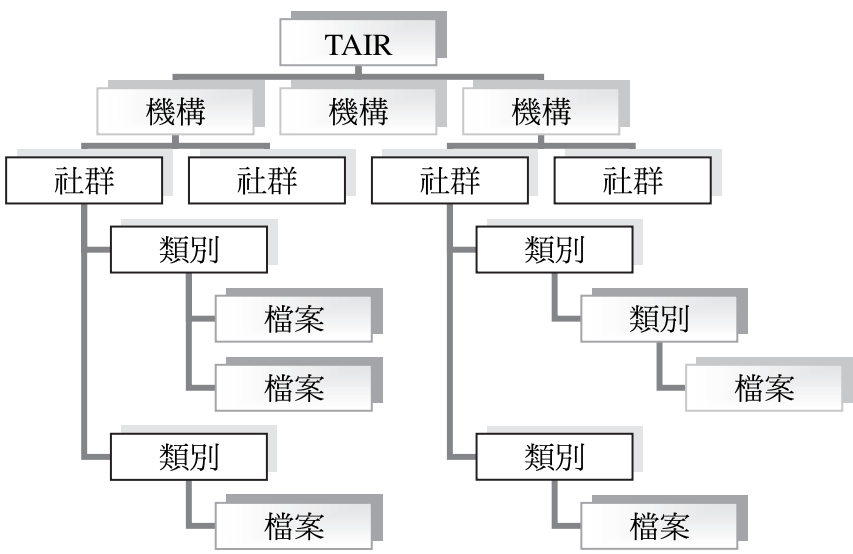


圖 2　TAIR系統之資料結構



圖 3　TAIR系統首頁

一、瀏覽模式

TAIR系統提供五種瀏覽模式，分別為依據「典藏機構」、「學術領域」、「題名」、「作者」、「日期」等模式瀏覽。依典藏機構瀏覽功能不同於NTUR系統以文件為瀏覽的單位，而是以社群與類別為瀏覽的單位，在TAIR系統所提供的樹狀圖選單中，可讓使用者於單一頁面瀏覽所有參與機構的社群與類別，了解各社群與類別之間的關聯性，點選個別連結則可連至該社群或類別的瀏覽頁面，進而瀏覽屬於該社群或類別的所有文件。使用者只要使用依「學術領域」瀏覽功能，則可以橫向地瀏覽相同學術領域的文件，而不受限於個別學術機構，可以節省開啟多個個別機構典藏系統的時間成本，這項功能是應用DSpace內建的handle機制，依據所有社群或類別的URN（Uniform Resource Name），擷取並展現相關學術領域的研究文獻。

目前個別之機構典藏系統是以學院及系所作為社群，各個學術機構對系所的命名各有差異，以工業工程學門為例，就有「工業工程學系」、「工業管理學系」、「工業工程與管理學系」等多種系所名稱，為了能在機構典藏系統打破名稱的分歧，進一步整合學術領域相關的各個社群，依照文件所屬學術領域的同質性瀏覽，以便利單一領域的使用者，可簡化大量的使用者操作步驟，因此，TAIR

系統提供依「學術領域」瀏覽的功能。在NTUR系統RC3之後的版本，可以為每個社群添加「學術領域」的標籤，學術領域的標籤是依照國科會的「學門專長分類表」作微幅修改。舉例來說，臺灣大學的機構典藏人員可以將社群「公共衛生學院」新增標籤「公共衛生及環境醫學」，屬於此社群下的所有子社群也將會繼承此標籤，當然，這些子社群也可添加其他的標籤，使得一個社群可以對應多個標籤；TAIR系統則依據這些學術領域標籤，允許使用者瀏覽特定學術領域的所有社群，TAIR系統也提供資訊篩檢的功能，使用者可以進一步篩選欲瀏覽的群組，再選擇依照題名或日期做排序，瀏覽被各個機構典藏系統認定為屬於該學術領域的文件。

依題名瀏覽、依作者瀏覽與依日期瀏覽三項功能完全參照NTUR系統，將所有參與之學術機構的文件分別依題名、依作者、依日期排序，於瀏覽頁面提供了快取工具與分頁工具，繼承了NTUR系統的優點。快取工具可讓使用者於瀏覽頁面選擇排序方式（遞增或遞減），也可以讓使用者跳轉到特定題名、特定作者或特定時間的頁面；分頁工具讓使用者可以直接選取分頁或選擇每個分頁所呈現之文件筆數，以符合使用者的個別需求。在資料結構的設計上，因為一個作者可能會在不同的機構發表著作，舉例來說，很多教授會在不同的學校授課與指導學生，自然也會在不同學校發表著作，TAIR系統於「依作者瀏覽」的介面則整合了所有作者，讓使用者可以一次瀏覽屬於此作者跨機構的所有文件。圖4為TAIR系統依日期瀏覽頁面，系統會呈現每一筆資料的所屬機構、日期、題名及作者。

二、搜尋功能

除瀏覽功能外，TAIR系統亦提供搜尋功能，讓使用者能迅速得到所需資訊。TAIR系統在系統首頁的上方設置一下拉式選單，使用者可以選擇查詢任一學術機構的文件，也可以選擇查詢所有學術機構的文件。TAIR系統會整合呈現查詢結果，使用者可以節省分別至多個學術機構典藏系統查詢的麻煩。

目前多數機構典藏共享平台系統僅提供書目資料的檢索，因為成本比較低，只需要透過OAI建構一個Service Provider即可；提供全文檢索功能則是比較少見的作法。然而，機構典藏有一個重要的精神是儘可能的提供全文，以便利使用者線上取得、閱讀、引用，以進一步地提昇機構的學術影響力。因此，TAIR系統必須保留NTUR系統原有的全文檢索以及進階檢索的功能，TAIR系統工作團隊因應這項挑戰，持續改善系統效率，在提供全文檢索功能的情形下，確保使用者的便利的使用經驗。目前使用者可針對「作者、題名、主題、摘要、其他Identifiers、語言」欄位做布林關



圖 4　TAIR系統瀏覽頁面

係查詢，也可以限定只搜尋具有全文檔案的文件。圖5為TAIR系統查詢結果頁面的一個範例，查詢結果分頁提供文件題名、作者、出版日期、所屬機構、以及簡短之摘要。

三、資料統計

TAIR系統首頁顯示了四種類型的統計資料，上傳排行、下載排行、查詢排行、以及資料筆數的分布。上傳排行會顯示臺灣機構典藏系統中上傳次數最多的機構與上傳次數最多的作者；下載排行則顯示臺灣機構典藏系統中被下載次數最多的機構與被下載次數最多的作者與文件；查詢排行將顯示TAIR系統中的熱門查詢關鍵字；資料筆數的分佈主要是以各種不同的面向分析臺灣機構典藏資料的分佈情形，例如由資料類型、機構類型、地理區分、學科類型。上述之資料統計功能希望提供使用者一些熱門資訊，讓使用者可以一目了然TAIR資料內容的特性，更希望藉由統計資料的分析刺激各個機構典藏系統的管理者上傳資料，以建立完整的臺灣機構典藏。

四、後台管理

功能完善且使用便利的資訊系統，不僅需要考量使用者的使用行為，也必須為系統管理者建構好用的後台管理功能。TAIR系統的後台管理功能，如圖6所示，TAIR系統管理後

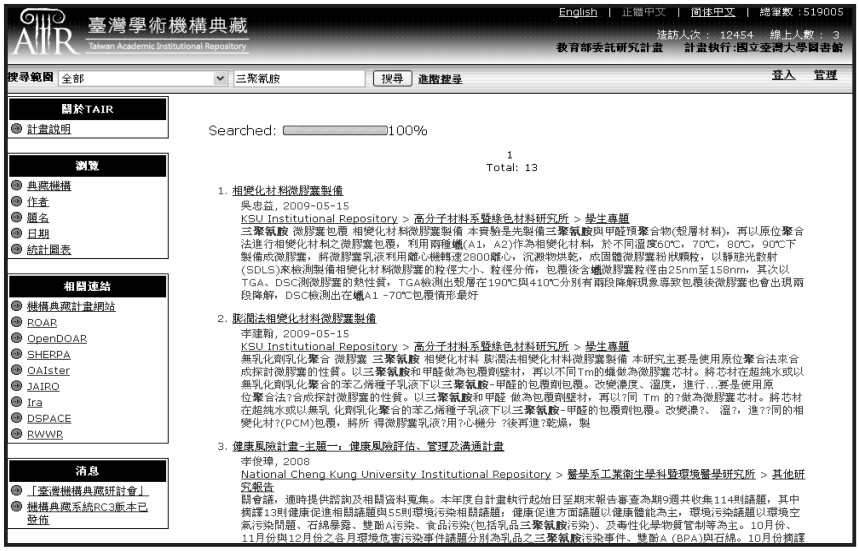


圖 5　TAIR查詢結果頁面



圖 6　TAIR系統後台管理介面

台提供了機構管理、新聞管理、文件鏈結管理、及人員管理等幾項主要的功能。管理者登入後台後，可發佈新聞及張貼相關文件鏈結。機構管理的功能較為複雜，管理者可以設定參與機構的類型分頁顯示，可以管理參與機構的中英文名稱、系統網址、OAI連結網址、下載筆數、機構位置等資訊，假設參與之學術機構的資訊改變，例如版本升級或網址變更，管理者都可在後台直接進行修正。除此之外，如果參與之機構典藏系統發生當機或進行維護，TAIR系統提供了即時性的保護措施，管理者可以在後台選擇隱藏該機構的所有資料，不在前台呈現該機構之資料，避免使用者誤用失效之連結。

肆、TAIR系統之研發

目前TAIR系統架設於Sun X4600主機，考慮TAIR系統的資料量龐大，作業系統使用Linux Red Hat，資料庫系統則採用MySQL系統，皆為耗損資源少，效能穩定的系統，為因應未來的發展，也將開發Oracle版本。至於TAIR系統功能的開發，乃是以NTUR系統的JAVA程式碼為基底修改而成，有關NTUR系統的相關細節可參考陳光華與吳哲安（2007）的論文；另一方面，NTUR系統部份程式碼也配合TAIR系統，做輸出資料格式的修正，才能大量減少開發的時程，在二個系統的維護管理工作上也兼具便利性。

對於TAIR系統而言，最重要的任務就是彙整各參與機構典藏系統的資料，且必須每日更新各參與機構典藏系統的資料。TAIR系統在處理資料傳輸的機制分為兩種：第一種為使用DSpace內建的OAI（Open Archive Initiative）機制（OAI, n.d.），可將NTUR系統內所含檔案的Metadata匯出為XML格式的資料，憑藉著XML格式具延展性的特性，不管NTUR系統裡檔案的Metadata有多少欄位，均能完整呈現，再配合專門處理OAI格式的JAVA套件XML Parser，TAIR系統可以迅速完整地讀取檔案資料並匯入系統，並於TAIR系統的伺服器端重新建立複本，圖7為NTUR系統以OAI格式輸出的檔案Metadata。由於DSpace系統仍有未轉換為OAI格式的資料，如社群資料、檢索結果等，工作團隊則必須自行修改NTUR系統，為了增進資料轉換的效率，改採JSON為輸出資料的標準格式，此為第二種資料傳輸的機制。JSON格式不同於XML格式，不具延展性，但JSON格式著重於數據交換，並支援陣列結構（JSON Array），判讀上會較XML格式節省時間。圖8為NTUR系統以JSON格式輸出的社群Metadata，分別列出了社群的名稱、編號、handle鏈結等資訊。

在匯入各機構典藏系統的資料之後，因為NTUR與TAIR系統採用的分頁模式與DSpace不同（Chen & Hsiang, 2009），若使用DSpace資料



```
- <record>
  - <header>
      <identifier>oai:140.112.114.62:246246/1726</identifier>
      <datestamp>2009-04-09T03:22:58Z</datestamp>
      <setSpec>hdl_246246_40</setSpec>
    </header>
  - <metadata>
    - <oai_dc:dc xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
      http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
        <dc:creator>崔伊蘭</dc:creator>
        <dc:date>1994-06-15</dc:date>
        <dc:date>[[accessioned]]2006-07-25T04:35:06Z</dc:date>
        <dc:date>[[available]]2006-07-25T04:35:06Z</dc:date>
        <dc:date>[[issued]]2006-07-25T04:35:06Z</dc:date>
        <dc:identifier>20060725115240859855</dc:identifier>
        <dc:language>zh_TW</dc:language>
        <dc:language>[[iso]]zh_TW</dc:language>
        <dc:publisher>臺北市：國立臺灣大學人類學系暨研究所</dc:publisher>
        <dc:rights>國立臺灣大學人類學系暨研究所</dc:rights>
        <dc:title>人類學學門人力資源的現況分析調查</dc:title>
        <dc:type>Text</dc:type>
        <dc:coverage>計畫年度：83;起迄日期：1993-10-01/1994-06-15</dc:coverage>
        <dc:description>[[note]]核定金額：204,000元</dc:description>
        <dc:identifier>[[uri]]http://140.112.114.62//handle/246246/1726</dc:identifier>
        <dc:subject>[[classification]]10</dc:subject>
      </oai_dc:dc>
    </metadata>
  </record>
```

圖 7　以OAI格式輸出的檔案Metadata



```
[{"community_id":"172","name":"教務
處","itemCnt":"0","itemCntBundle":"0","handle":"246246/34196","code":"A400000","field":""},
{"community_id":"69","name":"教務處師資培育中心","itemCnt":"48","itemCntBundle":"29","handle":"246246
/5914","code":"A407000,A407010,A407020,A407030","field":""},{"community_id":"1","name":"圖書
館","itemCnt":"94","itemCntBundle":"86","handle":"246246
/1","code":"A900000,A901000,A902000,A903000,A904000,A905000,A906000,A907000,A912000,A913000,A915000
{"community_id":"56","name":"計算機及資訊網路中心","itemCnt":"22","itemCntBundle":"19","handle":"246246
/4821","code":"AC00000,AC10000,AC20000,AC30000,AC40000,AC50000,AC","field":""},
{"community_id":"54","name":"研究發展委員會","itemCnt":"3","itemCntBundle":"2","handle":"246246
/4817","code":"E100000,E101000,E101100,E101200,E102000,E103000,E1","field":""},
{"community_id":"157","name":"體育室","itemCnt":"113","itemCntBundle":"24","handle":"246246
/16096","code":"E203000,E203010,E203020,E203030,E203040","field":""},{"community_id":"176","name":"人口與性
別研究中心","itemCnt":"7","itemCntBundle":"5","handle":"246246
/42161","code":"F200000,F201000,F202000,F203000,ZF21000","field":""},{"community_id":"95","name":"凝態科學研
究中心","itemCnt":"193","itemCntBundle":"86","handle":"246246
/8614","code":"F300000,F301000,F302000,F303000,F304000,F305000","field":""},{"community_id":"37","name":"生
物多樣性研究中心","itemCnt":"21","itemCntBundle":"16","handle":"246246
/4783","code":"F700000,F701000,F702000,F703000,F704000,F705000,F7","field":""},
```

圖 8　以JSON格式輸出的社群Metadata

庫內原有的資料表結構，在瀏覽時效率較差，因此TAIR系統需要再執行資料表重建的動作。表1為DSpace系統紀錄文件Metadata的資料表範例，所有文件的Metadata皆記載於同一張資料表，而TAIR系統在瀏覽頁面其實只會顯示題名、日期與作者三種欄位，其餘的資料欄位會拖累資料庫內部查詢的效率，因此TAIR系統利用此資料表再重建了MetadataByTitle、MetadataByDate、MetadataByAuthor三張資料表，分別記載題名、日期、作者與文件編號的關係，表2為延續表1資料重建MetadataByAuthor資料表的範例。重建前述三張資料表後，為配合依作者瀏覽功能中，以作者為單位呈現該作者的所有著作，TAIR系統必須再將MetadataByAuthor資料表轉換為單一作者與文件編號的關係資料表，以便SQL指令作查詢，承繼之前的範例，表3為把表2多位作者切分為單一作者的MetadataBySingleAuthor資料表。以上為整個資料庫的完整重建流程，TAIR系統會每日定期至各參與之機構典藏系統擷取更新資料，將Metadata存放於TAIR主機，以供瀏覽模式之用，圖9為使用者於TAIR系統使用瀏覽功能的流程。

表 1　DSpace系統紀錄檔案Meatadata的資料表

Record ID	Field	Value
1	Author	吳漢津
1	Title	內涵理論組成律對材料方向性硬化之研究
1	Date	1991
1	Publisher	臺北市：國立臺灣大學土木工程學系暨研究所
2	Author	陳光華；陳信希
2	Title	文件內容之分析--語料庫為本的模型
2	Keyword	文件內容；語料庫
2	Date	1996
2	Publisher	臺北市：國立臺灣大學圖書資訊學系

表 2　TAIR系統重建MetadataByAuthor資料表

Record ID	Author
1	吳漢津
2	陳光華；陳信希

表 3　TAIR系統重建MetadataBySingleAuthor資料表

Record ID	Author
1	吳漢津
2	陳光華
2	陳信希

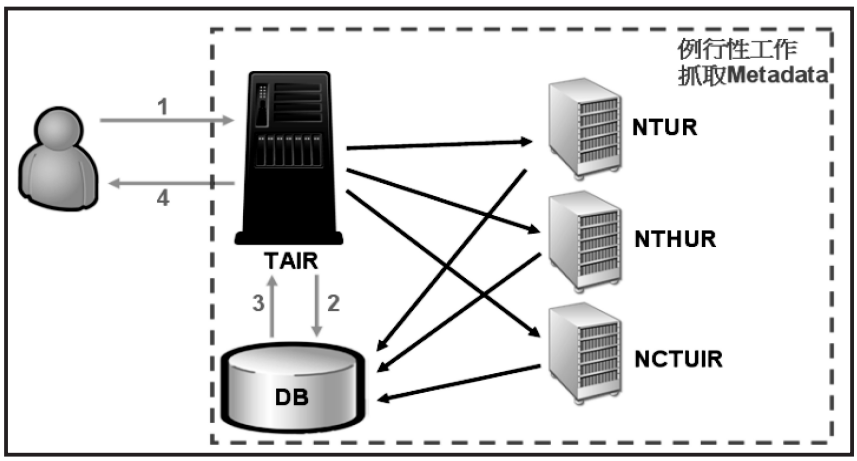


圖 9　使用者使用TAIR系統瀏覽功能流程

TAIR系統執行搜尋功能時的作業模式不同於瀏覽模式，並不是在系統本身的資料庫內做查詢，而是使用AJAX技術（Powell, 2008），即時至各參與之機構典藏系統主機查詢，彙整回傳之查詢結果後，再呈現於TAIR系統。TAIR系統不在本身資料庫做集中式索引的原因有二：

- 在「分散建置，整合使用」的原則下，TAIR系統內只存放各機構典藏系統的Metadata，並無全文檔案；
- 因為各機構典藏系統的總和資料量過大，而且系統典藏的文件會持續更新，取回全文製作集中式索引，時間效率較差。此外，各個學術機構對於將自己的文件放置於機構外部的典藏系統，仍有一定的疑慮。

因此，TAIR採用Metasearch的模式，再配合AJAX技術，可以讓版面更快呈現在使用者面前。圖10為使用者於TAIR系統使用搜尋功能的流程，TAIR系統在使用者輸入搜尋關鍵詞後，會同時傳送關鍵詞給所有參與機構的NTUR系統，當NTUR系統查詢結束後，會將查詢結果以JSON格式回傳給TAIR系統，由TAIR系統集中

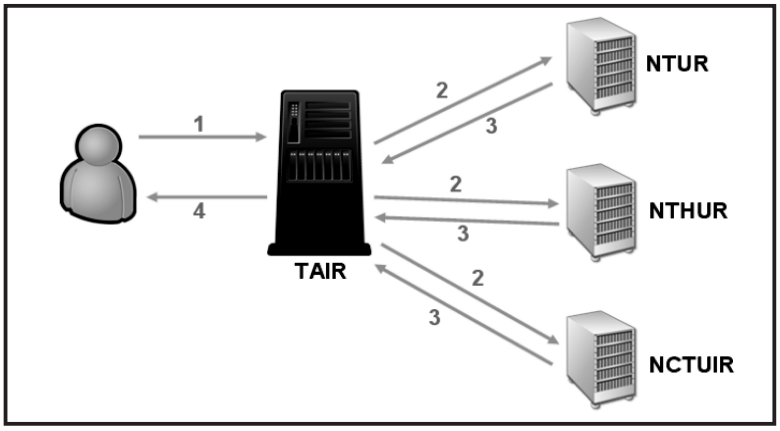


圖 10　使用者使用TAIR系統搜尋功能流程

呈現搜尋結果。此模式有一缺點，當各參與機構主機發生當機或中斷服務時，就無法取得該主機的檢索結果。

伍、問題與挑戰

TAIR系統作為一個學術機構典藏共享平台系統，有其獨特的性質，在實作過程中遭遇到一些困難，本節提出這些困難與挑戰，並分享相關技術細節與開發經驗。

一、主機效能的挑戰

TAIR系統收錄了全臺灣所有參與機構典藏計畫的學術機構文件資料，資料量勢必龐大，以臺灣大學機構典藏系統為例，截至民國98年三月，僅臺灣大學一校就典藏了十萬多筆文件，根據經驗，如此龐大的數量，會影響資料庫查詢的效能，嚴重時資料庫系統甚至發生停擺的情形，TAIR於98年四月處於測試階段時期，就擁有應付如此龐大的資料量的經驗，僅僅開啟「依題名瀏覽」頁面，就耗費40秒左右，在經過效能調校，改寫SQL語法減少資料表大量合併運算的情形後，系統已能在使用者容許的時間範圍內開啟瀏覽頁面。未來，TAIR系統勢必將處理動輒數十萬的文件資料量，面臨更大的挑戰。為因應資料量的成長，TAIR系統計畫在未來更換資料庫系統，逐步改採文件檢索技術，並建構雲端計算之軟硬體環境，或是與已建構雲端計算環境的機構合作（例如臺灣大學計算機及資訊網路中心），以因應未來資料量持續成長，以及使用流量的快速成長。

二、系統操作的便利

如同上述情形，大量的資料不

僅會造成系統效能問題，在資料的呈現也是個重要的課題。當使用者進入TAIR系統，一定會期望在符合需求的時間內找到所需的檔案。雖然TAIR系統仿照NTUR系統，於瀏覽模式建立快取工具列與分頁工具列，可讓使用者選擇排序方式與跳轉到特定頁面，但TAIR的資料量著實龐大，以98年四月測試階段時期為例，TAIR系統內十七萬餘筆檔案，就有一萬一千餘筆檔案的日期為2005，假如使用者選擇一頁檢視50筆檔案，依舊需要有兩百多個分頁顯示2005年的檔案，使得快取工具列與分頁工具列的效益大打折扣，對使用者的操作依然不具便利性。在未來該如何過濾資料、呈現資料與排序資料，讓使用者可以一目瞭然地快速找到資料，仍有很大的改進空間。

除上述問題外，TAIR系統目前不支援任何匯出功能，特別是以OAI格式匯出檔案Metadata功能，畢竟TAIR系統本質上仍是個機構典藏系統，而以OAI格式輸出Metadata已是任何一個機構典藏系統所必備的功能，往後若有其他系統需要與TAIR系統進行檔案整併或資料傳輸的動作時，OAI機制就變的不可或缺了。再者，對單一使用者來說，會希望能將所查閱到的文件Metadata以txt或bib格式匯出，另外做更深入的應用。TAIR系統在往後的任務排程上，已排定加入OAI機制與單筆資料的匯出功能。

三、參與機構的配合

為臺灣大學機構典藏團隊開發的TAIR系統與NTUR系統為互相承接的系統，除了DSpace系統原有以OAI格式輸出檔案Metadata功能外，NTUR系統自版本RC 1.0開始添增了以JSON格式輸出搜尋結果、社群Metadata等功能。換言之，若參與機構的系統版本停留在Beta 4.0以下，則僅保有DSpace原有的資料傳輸功能，在TAIR系統上，則無法支援全文檢索、依典藏機構瀏覽、依領域瀏覽等功能。隨著NTUR系統每釋出一個版本，支援TAIR系統的功能也就越趨完整，因此所有參與機構必須定期配合系統升級，若參與機構不配合升級，TAIR系統就無法自參與機構的系統取得Metadata，空有運算機制卻無法運作。

但對於參與機構而言，其實版本升級不如想像中簡單，無論在技術面與政策面，各個機構基於人力的限制或其他原因，都有其獨特的考量，不見得會在NTUR系統的新版本推出時立即升級。此外，針對TAIR系統的搜尋功能，因為是採用即時至所有參與機構搜尋的方式，所有參與機構更要保持系統的穩定性，一旦特定之機構典藏系統停擺，則搜尋時便無法得到該機構的檢索結果。TAIR系統要能穩定成長和運作，不單單是TAIR系統管理人員的責任，有賴所有的參與機構的全力配合。

四、Metadata品質的控管

在TAIR系統測試初期，對NTUR系統所輸出的資料採用照單全收的方式，但在收錄後發現各參與機構對Metadata的品質缺乏控管，出現過以下幾種狀況：

- Metadata欄位不齊全。單筆檔案的Metadata出現缺乏日期、題名等必填欄位，研判是NTUR系統開發初期上傳檔案的機制出現漏洞導致。
- 日期格式不一。若是各參與機構採用文件批次上傳時，NTUR系統對於資料格式並不進行檢查的工作，若參與機構沒有遵照統一的日期格式，TAIR系統的資料就會顯得雜亂無章，甚至出現「待查」、「印刷中」等日期格式。
- 作者標記不清。當一個文件為多位作者所擁有時，作者之間是以分號做區隔，若上傳者誤將分號鍵為逗號或其他符號，NTUR系統會將多位作者誤判為一位作者，之後再將此錯誤匯入TAIR系統。

針對以上錯誤，除了各參與機構負責人員在建立Metadata時必須注意外，TAIR系統務必建立Metadata檢查機制，與NTUR系統做資料傳輸時可進行格式的檢查，檢查出錯誤時也能及時回報結果給各參與機構以供修正。至於目前已存在於TAIR系統內的Metadata，若各參與機構對Metadata做了修改或刪除的動作，初期是採用人工修改的方式，必須等參與機構的負責人通知TAIR系統的負責人，負責人再進入TAIR系統的後台對Metadata做修改或刪除。未來，NTUR系統會記錄每日Metadata的修改與刪除結果，當TAIR系統與NTUR系統進行資料傳輸時會一併寫入紀錄，再對TAIR系統內的Metadata做修改或刪除。

五、使用非NTUR系統的機構

TAIR系統身為臺灣第一個機構典藏共享平台系統，自然肩負著於全臺灣推動機構典藏的使命，但畢竟不是所有的機構都願意使用NTUR系統，部分機構其實於機構典藏計畫推動之前就已使用DSpace系統或自行開發其機構典藏系統，在既有的系統已長時間運作之後，若要轉換為NTUR系統，不但資料移轉是個問題，使用者對新系統的操作介面又得重新學習。雖然這些機構並未使用NTUR系統，仍會想加入臺灣機構典藏計畫，TAIR系統本著整合臺灣學術研究成果的宗旨，歡迎其他非使用NTUR系統的學術機構加入，但在後續的資料處理上，必須進行資料格式的轉換工作。

首先，該機構的機構典藏系統必須支援OAI機制，能以OAI格式輸出資料至TAIR系統。再者，針對文件的Metadata格式，TAIR系統使用的是Dublin Core格式，若該機構使用的格式相異，必須先進行Metadata格式的轉換。完成上述步驟後，僅能支援「依題名瀏覽」、「依作者瀏覽」、

「依日期瀏覽」的三個基本功能，若要能於TAIR系統上執行「搜尋」、「依典藏機構瀏覽」、「依領域瀏覽」等功能，則必須另外配合TAIR系統的JSON格式輸出。因為非使用NTUR系統的機構典藏系統，其資料結構不見得相同，資料處理過程就必須靠兩方的系統開發人員溝通。

陸、結論

隨著網路蓬勃發展，數位化時代來臨，一般使用者認為只要善用搜尋引擎，就可以輕易地獲得許多資料；對圖書館而言，在保存學術資源的同時，如何讓讀者更便捷地使用學術資源，是圖書館必須思考，且面臨的極大挑戰。機構典藏正是學術圖書館推動創新服務的利器，對學術研究人員而言，透過機構典藏系統，可以取得確保學術品質的文獻；對圖書館與館員而言，透過機構典藏系統，可以開創圖書館新時代的數位化服務。學術機構藉由機構典藏系統，不僅能夠保存機構內部的學術成果，更能對外展示學術成果，放大學術能量，奠定機構的影響力。

臺灣學術機構典藏系統，賦予了機構典藏更深的一層意義，「機構」可以不只是一個學術機構，更可以是一個組織、一個聯盟、或一整個地區下所有學術機構的集合。臺灣機構典藏藉由「分散建置、整合應用」的方式，整合呈現臺灣整體的學術研究成果。學術機構典藏共享平台由一個概念，要進入實際的營運並不是一蹴可幾的，除了本文探討的技術問題與系統實作之外，還是有許多學術機構對加入臺灣機構典藏有所疑慮，但隨著TAIR系統的規模越來越大，參與的學術機構展現出加入TAIR系統的優勢，勢必會有更多的機構願意加入學術機構典藏的行列。未來最重要的課題是臺灣學術機構典藏系統必須永續經營，而不是計畫終止就曲終人散，不論是系統的維護與開發，還是文件資料的新增與維護，都必須持續地進行。分享臺灣學術研究成果不僅是臺灣學術機構典藏系統工作團隊的責任，只要身為臺灣的學術機構，就有義務經營自身的學術機構典藏系統，建立自有學術機構典藏的特色，以利於學術研究成果的傳播，才能提升臺灣學術研究的能量，展現臺灣學術研究的影響。

誌謝

作者感謝教育部為提升臺灣高等學術機構研究與教學品質所做的努力，並實際行動提撥經費支持學術機構典藏計畫。作者感謝臺灣大學圖書館前任館長項潔教授與現任館長陳雪華教授的卓越領導，以及工作同仁韓竹平編纂、張素娟主任、郭美玲編審、江玉婷小姐、吳瑟量先生、林鳳

儀小姐、孫珮慈小姐的努力奉獻。作者感謝參與機構典藏計畫的學術機構，特別是擔任種子學校的中山大學圖書與資訊處處長官大智教授、交通大學圖書館館長楊永良教授與前副館長柯皓仁教授、成功大學圖書館館長謝文真教授與副館長蔣榮先教授、修平技術學院圖書館館長楊尚霖教授與組長陳清祺教授、崑山科技大學圖書館館長閔蓉蓉教授與副館長陳國泰教授、清華大學圖書館館長謝小芩教授與副館長鍾葉青教授、暨南國際大學圖書館館長吳幼麟教授。作者感謝政治大學圖書館館長劉吉軒教授具有洞察力的意見與建議。

參考文獻

陳光華、吳哲安（2007）。臺灣大學機構典藏系統之建置。*圖書館學與資訊科學，33*（2），33-47。

Chen, K.-H. & Hsiang, J. (2009). The Unique Approach to Institutional Repository: Practice of National Taiwan University. *Electronic Library, 27*(2), 204-211.

DSpace. (n.d.). *Why use DSpace?* Retrieved May 18, 2009, from http://www.dspace.org/index.php?/Why-Use/

Hajjem,C., Gingras, Y., Brody, T., Carr, L., & Harnad, S. (2005). *Open Access to Research Increases Citation Impact*. Technical Report, Institut des sciences cognitives, Université du Québec à Montréal. Retrieved Oct. 8, 2009, from http://eprints.ecs.soton.ac.uk/11687/

Ira. (n.d.). *Overview*. Retrieved May 18, 2009, from http://www.ira.auckland.ac.nz/project/

JAIRO. (n.d.). *About JAIRO*. Retrieved May 18, 2009, from http://jairo.nii.ac.jp/help/en/about_us.html

Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature 411*, 521. Retrieved Oct. 8, 2009, from http://www.nature.com/nature/journal/v411/n6837/full/411521a0.html

Lynch, C. A. & Lippincott, J. K. (2005). Institutional Repository Deployment in the United States as of Early 2005. *D-Lib Magazine, 11*(9). Retrieved May 18, 2009, from http://www.dlib.org/dlib/september05/lynch/09lynch.html

OAI. (n.d.). *The Open Archives Initiative Protocol for Metadata Harvesting*. Retrieved May 18, 2009, from http://www.openarchives.org/OAI/openarchivesprotocol.html

OAIster. (n.d.). *What is OAIster?* Retrieved May 18, 2009, from http://www.oaister.

org/about.html

Powell, T. (2008). *Ajax: The Complete Reference*. McGraw-Hill Osborne Media.

ROAR. (n.d.). *Registry of Open Access Repositories*. Retrieved May 18, 2009, from http://roar.eprints.org/index.php

RWWR. (n.d.). *Ranking Web of World Repositories*. Retrieved August 10, 2009, from http://repositories.webometrics.info/top400_rep_inst.asp

Westrienen, G. V. & Lynch, C. A. (2005). Academic Institutional Repositories Deployment Status in 13 Nations as of Mid 2005. *D-Lib Magazine, 11*(9). Retrieved May 18, 2009, from http://www.dlib.org/dlib/september05/westrienen/09westrienen.html