

臺灣大學機構典藏系統之建置

The Construction of National Taiwan University Repository

陳光華

國立臺灣大學圖書資訊學系副教授；國立臺灣大學圖書館系統資訊組組長

Kuang-hua Chen

Associate Professor, Department of Library and Information Science, National Taiwan University

Head, Division of Information Technology Services, National Taiwan University Library

E-mail: khchen@ntu.edu.tw

吳哲安

國立臺灣大學圖書館研究助理

Che-an Wu

Research Assistant, National Taiwan University Library

E-mail: cheanwu@ntu.edu.tw

關鍵詞 (Keywords)： 館藏發展(Collection Development)；機構典藏(Institutional Repository)；
系統開發(System Development)

【摘要】

機構典藏的真實意涵是典藏機構內部的研究產出，以展現機構的研究成果與學術影響，目前建置機構典藏系統已經受到廣泛的重視。臺灣大學擁有豐富的研究成果及大量的學術出版，如何善用並促進這些研究能量的呈現、取用與交流，是臺灣大學機構典藏系統計畫的精神目標；為使用者建構一個易於使用的機構典藏系統是該計畫的技術目標。本文說明臺灣大學機構典藏系統的建置過程與經驗，並討論開發團隊遭遇的各種技術問題與因應措施，涵蓋軟體評估、客製化系統、資料徵集與中文化等課題。

【Abstract】

The meaning of Institutional Repository is to preserve the research outputs of an Institution. The preserved contents as a whole will demonstrate the achievements and influences of this institution. As a result, the construction of institutional repository systems has been receiving attentions of research institutions. National Taiwan University (NTU) has formally or informally been publishing many research results. How to make these researches available to the public is the spiritual goal of NTU Repository (NTUR) project; how to implement a user-friendly system is the technical goal. This article first describes the construction and the experience of NTUR. The

difficulties and the possible solutions including the selection of system platform, system customization, item acquisition, and feature localization are discussed as well.

一、緒論

自從 WWW 瀏覽器成為網路應用的共同使用平台，透過瀏覽器呈現各式各樣的數位文件，已成為一股潮流，讀者與作者的藩籬已漸趨消逝，也可以完全忽略出版商的中介角色。圖書館作為典藏知識的寶庫，因應網際網路的發展，擬定適切的館藏發展政策與策略，方能確保圖書館的永續發展。項潔、洪筱盈（2007）探討大學圖書館由近而遠、由內而外的典藏策略，極具參考價值，也是在網際網路時代，大學圖書館掌握網路的特性，充分發揮圖書館功能的具體案例。大學圖書館的角色與功能，其實是很艱鉅的，尤其是在一所研究型大學的組織內部。圖書館必須提供有助於研究的學術文獻，創造有助於研究的學術資源，典藏機構內部的歷史文獻與成員創造的學術文獻，建置親和好用的學術資訊系統，因為學術研究是一所研究型大學持續發展的最重要課題，也是建立學術聲望與學術地位的基礎性工作。

1665 年第一份學術期刊 “Journal des Savants” 由 Denis de Sallo 創刊後（Hallam, 1970, p.406），學術研究成果的展現透過期刊進行學術傳播的模式已經有數百年的歷史了。大學作為學術研究的重鎮，教授、學者們採取這種傳播模式，將個別的研究成果展現於目前數以百萬計的學術期刊（ISSN International Centre, 2006）；各領域的學者專家們在自己的學術領域內，研讀所屬領域的學術期刊，似乎已經是既定的學術活動模式。然而，就提供學術研究環境的主體--大學或研究機構，這裡存在三個重要的議題：第一，研究機構無法瞭解自身整體的研究成果；第二，研究機構難以記錄歷來的研究成果；第三，研究機構難以承擔價額越來越昂貴的學術期刊。面對這三個議題，目前有二個概念引起廣泛的討論，其一是開放存取（Open Access，簡稱 OA），其二是機構典藏（Institutional Repository，簡稱 IR）。開放存取可以扭轉前述第三個議題提及的問題；機構典藏則提供了第一個與第二個議題的解答。簡而

言之，一個機構典藏系統勢必是開放取用的，而且一定是建構於網際網路環境的學術資訊系統，透過有別於傳統學術期刊文獻的傳播通路，展現並分享學術機構的整體研究成果，提升機構的影響力與學術聲望。

機構典藏的真實意涵是典藏機構內部的研究產出。目前學者專家的研究產出多數是數位文件（Born in Digital），但是卻還沒有可確保長期且安全的方式來保存其研究成果（Lynch, 2003），過去將紙本論文數位化的成本已大幅降低，現在更沒有理由不積極進行機構典藏的工作。以大學的機構典藏系統而言，其服務的對象主要為校內各單位的成員，重點在於數位內容的保存、使用與散佈（Lynch, 2003）。目前國外已有許多研究機構著手建置機構典藏系統（Westrienen & Lynch, 2005），而臺灣大學擁有豐富的研究成果及大量的學術出版品，如何善用並促進這些研究能量的呈現、取用與交流，是臺灣大學機構典藏系統（National Taiwan University Repository，簡稱 NTUR）計畫的精神目標；為使用者建構一個易於使用的機構典藏系統是 NTUR 計畫的技術目標。因此，NTUR 計畫特別強調系統對於全文檢索、瀏覽的應用與支援，以及檢索的後分類，以因應使用者不同的瀏覽需求。

本文說明臺灣大學機構典藏系統的建置過程與實作經驗，討論開發團隊遭遇的各種技術問題與因應措施，涵蓋軟體評估、客製化系統、資料徵集與中文化等課題。NTUR 計畫是由臺灣大學圖書館系統資訊組所組成的工作團隊負責進行系統的開發與建置，採用 MIT 與 HP 開發的 DSpace 1.4 Beta 軟體套件作為發展的基礎平台，改寫為 NTUR 系統，持續新增並改善系統功能。目前 NTUR 機構典藏系統收錄的文件共有 32,533 筆，包含國科會計畫、期刊論文及各式校內出版品，其中具有全文的共有 15,885 筆。此外，臺灣大學圖書館也同時進行臺灣機構典藏計畫的 OAI 聯合檢索測試，期望將目前已建置機構典藏系統且支援 OAI 協定的學校透過此一機制連結起來，能夠展現臺灣整體的學術研究能量，提升國際能見度。

本文第二部分說明目前世界各國機構典藏的發展現況；第三部分討論臺灣大學圖書館 NTUR 軟體系

統平台的選擇與建置；第四部分探討 NTUR 開發過程遭遇的問題與挑戰；最後是簡短的結論與建議。

二、機構典藏發展現況

(一) 歐美各國機構典藏現況

至 2005 年中為止，美國已有 40% 以上的大學擁有營運中的機構典藏系統，而尚未建置機構典藏系統的大學中，也有 88% 的學校正著手進行建置 (Lynch & Lippincott, 2005)，也就是說，有 92.8% 的學校正計畫或進行機構典藏系統的建置與運作，認可推行機構典藏的重要性。在對歐洲等國的調查結果中指出 (如表一)，挪威、德國、荷蘭等國的大學建置比率已達到百分之百 (Westrienen & Lynch, 2005)，顯示一個學術機構不僅建置一套機構典藏系統，會因應不同的目的，而發展不同的機構典藏系統；此外，66% 以上國家的建置比率超過 40%。表二則更進一步以保健與生物科學為涵蓋主題，廣泛調查世界各國建置機構典藏系統的情形，可以發現這兩個學術領域在整個機構典藏系統的建置與發展，佔有非常重要的角色，也可以再次看到機構典藏系統受到世界各國重視的情形。這些統計資料都顯示出建置機構典藏的重要性已逐漸為已開發國家的高等教育機構所認可，雖然目前僅有少數的國家嘗試建立政府層級的學術機構典藏政策，但是，機構典藏觀念的推廣，機構典藏系統的建置，以及運用機構典藏系統作為學術研究成果保存與交流的平台，已經變成一股全球化趨勢。

(二) 亞洲各國機構典藏現況

亞洲地區國家的機構典藏發展現況可參考表三，目前以日本、印度、香港、臺灣等地的建置數量較多。中國大陸的情況比較特殊，依表三所列僅有一個 (表二則顯示有 2 個，各個資料來源的調查結果不盡相同)，廈門大學已建置機構典藏系統 (<http://dspace.xmu.edu.cn/dspace/>)。此外，另有由國家提供資源來協助建置 CNKI 中國知識資源總庫 (<http://www.cnki.net/>)，就比例而言，CNKI 涵蓋 300 所大學約佔其總數的 16%，雖然就機構典藏的定義而言，個別機構的角色消逝不見，不能算是機構典藏系統，但是就使用者使用資源的角度而言，具有

整合各機構典藏資源的優點；而從國家的觀點，可以做國家整體研究能量的展現窗口。除了中國大陸之外，其餘亞洲等國的建置比例則在 4~10% 之間，而香港地區大學總數較少，因此建置的比例偏高。CNKI 已經建置大量的學術論文全文資料，對外展現其華文學術資料，臺灣若不積極建構機構典藏系統，會造成臺灣學術界要用本國的學術研究成果，可能比使用英文學術資源或大陸的學術資源更加不易，而國際學術界要引用臺灣學術研究成果就更加困難。

(三) 臺灣機構典藏現況

相較於國外的發展，臺灣機構典藏觀念的推廣與機構典藏系統的建置，起步都比較晚，目前僅有臺灣大學、政治大學、交通大學、清華大學、逢甲大學、崑山科技大學等校建置機構典藏系統，請參閱表四。其中政治大學、交通大學、清華大學採用臺灣大學修改 DSpace (<http://www.dspace.org/>) 的 NTUR 系統，逢甲大學與崑山科技大學採用 DSpace，進行機構典藏計畫的測試或營運機構典藏系統。此外，臺灣各大學都建置博碩士論文的提交與典藏系統，如中山大學採用 ETD-db 系統作為博碩士論文的提交與典藏系統，然而僅收錄博碩士論文。由於機構典藏必須以機構的觀點，典藏並展現機構的整體研究成果，由比較嚴格的機構典藏的定義，僅收錄博碩士論文的系統並非真正的機構典藏系統。依據教育部的統計，全國大專院校總數為 163 所 (教育部，2006)，所以目前臺灣大專院校建置機構典藏系統的比例僅有 7/163，比例仍然偏低。

臺灣大學經過仔細的討論與考量，決定使用 HP 與 MIT 合作開發的 DSpace 系統，並新增許多新功能，以符合臺灣使用者的需求，改進的新系統為 National Taiwan University Repository (NTUR, <http://ntur.lib.ntu.edu.tw>)。NTUR 的開發團隊陸續於全國大學校長會議與臺灣許多的機構典藏研討會上展示 NTUR 機構典藏系統，並於臺灣大學內部推廣 NTUR 的使用，以及推動臺灣各大專院校使用 NTUR 建置各校之機構典藏系統。臺灣大學圖書館 NTUR 開發團隊亦將持續開發 NTUR 系統的各项新功能。

三、NTUR 之建置

目前機構典藏相關的解決方案種類很多，有必須付費取得的商業軟體、也有開放原始碼的免費軟體，每種方式都有其優缺點。當決定要建置機構典藏系統之後，欲導入的單位必須考量其機構典藏系統應具備的功能需求，並衡量自身擁有的相關資源，再決定採用何種方式建置系統。觀察目前世界各國已建置的機構典藏系統，主要採取下列三種方式。

- 採購商業軟體，如 ProQuest 的 Digital Commons。Columbia University 採用這種做法，但是 Columbia University Library 表示未來將改用 DSpace。
- 使用自由軟體，如 Eprints、DSpace。MIT、UM at Ann Arbor 採用 DSpace。
- 使用自由軟體，但是進行大幅度的改寫，如臺灣大學使用 DSpace，但是改寫成 NTUR。

但是無論採用何種方式，都必須仔細考量建置系統的總成本，包含開發、導入、教育訓練、維護及人事等成本。臺灣大學圖書館 NTUR 開發小組，經過仔細思考，決定採取第三種方式，主要的考量是希望導入開放原始碼的系統，以獲得自行客製化系統的彈性，而最後決定採用惠普公司和麻省理工學共同研發的 DSpace。在眾多的自由軟體中選擇 DSpace 的主要原因有下列幾點：

- DSpace 採用 Java 撰寫系統，符合臺灣大學圖書館系統開發應使用 Java 語言的原則。
- DSpace 可以跨平台安裝，且可利用 Java API，自行開發擴充功能。
- DSpace 有完善的使用者社群可以相互切磋討論。
- DSpace 已經被 200 個以上的學術機構採用，作為其機構典藏系統平台 (DSpace, 2007)，所以它在實作上已經被證明了具有一定效能和穩定性，也可支援大量資料的運作。
- DSpace 支援 OAI 機制，縮短後續發展整合查詢的功能。
- DSpace 整合 handle 機制，系統內的數位文件具有唯一的 URN。

- DSpace 前後台功能架構完整，可大幅縮短系統建置時程。

實際上 DSpace 是持續發展的軟體，而臺灣大學圖書館 NTUR 系統的開發決定採用 DSpace 1.4 Beta 作基礎，而且不會再隨著 DSpace 的更新而改版。換言之，NTUR 的開發由 DSpace 1.4 Beta 之後，就與 DSpace 的開發分道揚鑠，臺灣大學圖書館 NTUR 團隊將自行開發各種功能，並依照本地的語言特性及使用需求進行客製化，使其功能更加完備，使用更加便利。

NTUR 使用 DSpace 作為基礎，其系統架構與 DSpace 相似，由儲存層、商業邏輯層、應用層三個階層組成，如圖一所示，各階層則分別具有多個元件負責各項功能。儲存層負責管理實體儲存設備中的書目資料及上傳文件；商業邏輯層處理帳號管理、權限設定、與工作流程；應用層則包含與外部系統進行溝通的相關元件，例如使用者介面與 OAI 通訊協定等。(Tansley, et. al, 2006) 臺灣大學機構典藏系統為因應本地的需求與使用方式，新增或修改的客製化功能如圖一粗線框線所示，下文將說明各個模組的改進作法或新增功能。

(一) 使用者介面 (Web UI)

為了改善原來 DSpace 系統使用者介面的友善程度，NTUR 在流程和介面上都增加了許多符合使用者操作習慣的相關功能。例如資料查詢的後分類處理、各式統計數據、檔案格式判定、多語言系統介面等等。NTUR 不斷強化使用者介面一致性，讓系統的操作性更為直覺，並加強各種輔助措施，以協助使用者，期望即使使用者是第一次使用 NTUR，對於各種功能仍然能夠一目了然，能夠進行流暢的操作，並且快速找到所需的資料。

NTUR 系統前台網站首頁如圖二所示。首頁右上方顯示[英文 | 正體中文 | 簡體中文]的多語言介面選項、全文筆數與總筆數、造訪人次與線上人數等訊息，使用者可以直接切換不同語言的介面，方便不同使用者的需求。首頁右方則列有文件的上傳排行與下載排行，使用者可以知道哪些作者或社群的論文經常被下載，哪些社群或作者上傳的文件最多。首頁正中央是學院的列表，並附上各學術的文件筆數，使用者可以一目了然系統典藏文件的分佈

情形；使用者還可以直接展開，檢視學院之下各社群的文件分佈情形。首頁左方羅列社群與類別、題名、作者、日期，以供使用者瀏覽，例如使用者點選日期，則會顯示如圖三所示之依日期排序的文件列表，可以遞增排序（由舊到新排序）或遞減排序（由新到舊排序）。事實上，這樣的時間資訊會帶使用者額外的訊息，如 NTUR 系統目前收錄最早的文件是 1963 年出版的一篇物理論文，這種檢視方式加上了時間的脈絡資訊，可以追溯臺灣大學學術研究的歷程，甚至學者可以使用 NTUR，撰寫臺灣大學學術研究發展史。

改善使用者介面的目的，在於期望提供的功能更符合使用者的使用需求，以提升使用滿意度，讓典藏於 NTUR 系統的學術成果可以有效地透過網際網路分享給全球潛在的使用者，以發揮其學術價值與增進機構的學術聲望。

(二) 過濾器 (Media Filter)

過濾器負責過濾各種類型檔案內容的工作，目前我們的改進方式主要是調整其中剖析器的功能。DSpace 內建的剖析器為 pdfbox，其處理中文的能力並不理想，成功率僅約三成左右而已，致使早期的 NTUR 的檢索結果有許多的亂碼，造成使用者的困擾。就一個機構典藏系統而言，必須典藏各種類型的檔案，可能是 Microsoft 的 DOC 檔案、PPT 檔案，可能是 Adobe 的 PDF 檔案，也可能是一般的 TXT。目前多數的檔案格式是 PDF，雖然 PDF 是廣泛使用的檔案格式，也是所謂的「可攜式文件格式」，但是 PDF 檔案卻有不同的產生方式，可能是由 Postscript 轉成 PDF 檔案，可能是由 LaTeX 產生再轉成 PDF 檔案，可能是自由軟體產生的 PDF 檔案，可能是各種格式的文件經由 PDF 虛擬印表機列印而得，不同方式產生的 PDF 檔案不見得完全遵守 Adobe 對於 PDF 檔案制訂的規範，中文的文件更容易出現這種情形，以致於相容性不高，造成過濾器的負擔很沈重，必須能夠處理各式各樣的問題，因之，過濾器的失敗率也很高。NTUR 開發團隊決定更換 DSpace 內建之剖析器以提升中文資料處理的準確度。

(三) 搜尋 (Search)

NTUR 特別增加資料搜尋後分類與關鍵字標示等

功能，以增強搜尋結果的可使用性。圖二的粗線框線是 NTUR 的檢索功能，共分簡易檢索與進階檢索，NTUR 大幅度改寫 DSpace 原提供的檢索功能，新增檢索後分類的功能，相較於其他的機構典藏系統，這應該是 NTUR 最具特色的功能。以檢索「原住民」為例，會顯示圖四所示的檢索結果，使用者可以再依作者、類別或日期瀏覽，就是一種檢索後分類的功能。例如點選以作者瀏覽，使用者可以瞭解有哪些學者從事原住民研究，自然而然，可以透過 NTUR 形成跨學院的研究社群；若點選以年代瀏覽，使用者可以參照時間歷線，瞭解臺灣大學歷年從事「原住民」研究的情形，研究的高峰發生在哪一個時間點。為了維持系統的查詢效率，NTUR 應用 Ajax 的技術，在增加各項新功能後，使用者仍然可以快速獲得所需的查詢結果，也就是在使用者瀏覽資料的同時，系統則繼續進行運算，將統計後的資訊傳回前台。此外，亦新增異體字查詢功能，是為了解決中文化方面的問題所做的設計調整。

(四) 瀏覽工具 (Browse Tools)

原來 DSpace 在設計上是為了在大量資料裡找出瀏覽頁面上需呈現的資料，故使用類似滑動視窗的作法，把使用者畫面直接帶到整個資料庫的某個區段。由於是滑動視窗的作法，所以在資料呈現時不會有明顯的分頁頁次，這跟使用者平常在網路上的使用方式並不相符。因此，在增加分頁功能時，需要分別對題名、作者及日期的瀏覽清單與索引等進行修改。修改後的分頁器除了提供一般的分頁功能，如跳到某一頁、到最前頁、到最後頁等功能外，也增加了每頁顯示數目的選項，使用者可以分別選擇每頁顯示 10 筆、25 筆或 50 筆資料，以配合每個使用者不同的使用習慣。

另外，在使用原來的 DSpace 系統時，若使用者需要檢視某文件有什麼全文檔時，需要再次點選，以進入文件頁面逐一檢視。改進後的 NTUR 瀏覽介面在題名與日期瀏覽列表，會依據全文檔的格式，顯示對應的全文檔圖示，讓使用者更容易判別文件是否內含全文檔，以及全文檔的格式為何；而在作者瀏覽的部份則列出該作者在系統內的文件數量與有全文檔的文件數量（以 [全文數/總筆數] 表示），務求讓使用者可以在單一頁面上得到最多的資訊。

(五) 資料庫 (Database)

臺灣大學圖書館開發的 NTUR，未來會釋出供其他學校使用，考慮到不同的學校有其個別的考量，因此 NTUR 必須可以安裝於 Windows 系列作業系統以及 Linux 系列作業系統，所以釋出的版本必須使用開放的免費資料庫系統軟體，同時也可以執行於前述二種作業系統的資料庫系統軟體。原來 DSpace 採用 PostgreSQL 資料庫，但是臺灣的使用者與系統建置者較常使用 MySQL，NTUR 決定改用 MySQL 進行系統的開發；此外，為了確保 NTUR 於臺灣大學校內運轉的效能，NTUR 開發團隊也開發使用 ORACLE 資料庫的 NTUR 版本。

原來 DSpace 套件的資料庫架構非常簡單，僅有四個資料表 (Tables or Relations)，基本元素資料表；基本元素從屬關係資料表；物件瀏覽資料表；其他元素資料表。NTUR 系統則新建了許多資料表以支援系統新增的許多功能，如權威控制資料表、詞彙對應資料表、作者文件對應資料表；或是在既有的資料表，新增資料欄位，如計數欄位、使用者及系所屬性欄位、檔案連結欄位、存取權限控管欄位、以及第二語言欄位。

四、問題與挑戰

由於 NTUR 是建構於 DSpace 1.4 Beta 版的基礎之上，改寫既有的系統，並考量本地的使用需求，進行客製化的改善功能與新增功能是無可避免的，當然會面臨一些問題與挑戰，以下簡要說明。

(一) 系統架構的調整

NTUR 系統在開發初期，由於有時程上的壓力，往往以新需求的處理為優先考量，因此使得整個系統架構有如疊床架屋一般，嚴重影響系統的執行效率。例如分頁的功能設計，原來 DSpace 系統架構在瀏覽資料的設計並不適合使用分頁的功能，因此必須選取全部資料以計算分頁呈現的頁數，導致拖慢整理分頁的速度。另外，在關鍵字查詢的後處理上，也需彙整額外的統計數字，使得結果呈現的速度受到影響，每次查詢平均約需費時 8~10 秒。

比較適當的作法應該是容許系統開發人員有足夠的時間去熟悉原來 DSpace 套件的系統架構，且改善

或新增功能需求時，應該力求整合於原有的程式系統架構，並參考原系統之程式撰寫原則來開發系統。目前經重新設計調整系統架構之後，使用瀏覽全部的功能來瀏覽系統內典藏的全部文件時，只需費時 1~2 秒即可打開相關頁面，已大幅改善網站瀏覽效率。

(二) 系統資源的管理

在 NTUR 系統的開發初期，系統經常在多位使用者同時進行瀏覽或全文檢索之後出現記憶體方面的錯誤，進而導致系統產生服務中斷的情形。這種問題的來源有兩方面：第一，資訊工程師尚未熟悉原來 DSpace 的架構，在功能開發上難以兼顧需求與效能兩者；第二，系統參數的調校未達最佳化。針對第一方面的問題，在工程師充分掌握系統架構之後，目前已逐漸調整架構以配合新需求之開發，避免重複及無效率的資訊庫存取運算；在第二方面的效能調校上，則藉助系統效能監控工具，進行流量與存取效率測試，逐步調整參數以最佳化 NTUR 前台網站伺服器、NTUR 後台資料庫與 NTUR 後台系統伺服器三者間的整體效能。在完成以上的調整之後，目前已無服務中斷或效能不佳的情形出現。

(三) 資料檢索的效率

DSpace 資料庫的設計方式乃以彈性為主要考量，可支援多種語言、提供自訂欄位，但其規劃方式違反資料庫正規化之相關原則，因此當資料量變大的時候，資料庫的存取效能將受到嚴厲挑戰。目前多數採用 DSpace 套件建置的機構典藏系統，其收錄的資料量並不多，如表五所示，80%以上的機構典藏系統其資料量均不足 5,000 筆，因此，檢索及瀏覽效率不佳的問題較少會被觸及到，尤其是 NTUR 在改寫 DSpace 之後，新增了許多功能，更需要特別考量檢索效率是否受到新增功能的影響。確實在 NTUR 發展初期，新增的許多功能造成系統效能的下降，進行關鍵字查詢時，動輒需花費 8 秒左右的時間才能讓結果顯示出來。但事實上，使用者有耐心等待的時間是在 6~10 秒 (King, 2003)，如果超過此一範圍，使用者的滿意度就會大幅下降，也有可能因為不耐煩而不願意繼續等待。因此，我們更改了索引目錄存放的欄位，如文件的系所與類別等資料，另

外建立摘要文字的索引，讓系統以更快的速度顯示檢索結果。

如前文所述，NTUR 系統在查詢結果的後分類上亦下了許多功夫，這些新增功能也可能會降低檢索結果呈現的效率，因此，NTUR 採用 Ajax 技術，先以最快速度回傳查詢結果給使用者，再逐步呈現各種統計數據。此外，為因應大幅成長的資料量，在第二階段資料徵集數量預計將達到五至十萬筆，屆時資料庫中的資料筆數也會逼近兩百萬筆，因此將配合更換現有資料庫，以符合未來的資料成長需求。

(四) 資料徵集的困難

機構典藏在系統導入初期，多半會遇到不知從何匯入文件的問題，如何從無到有、建立一個可讓使用者體驗機構典藏系統的效益，以說服其提供學術文件典藏於機構典藏系統，是機構典藏系統成功營運的重要關鍵。NTUR 獲取文件資料的方式，主要透過國科會計畫、期刊論文、校內出版中心等三個管道，蒐集歷年已發表之博碩士論文，以及校內之相關出版品。此外，也主動蒐集在臺灣大學網域內公開分享的電子文獻，經由人工查核後收錄於機構典藏系統。目前 NTUR 典藏的文件數量，結至 2007 年六月底為止已達 32,533 筆資料，全文資料計有 15,885 筆，其中全文比例接近 50%，請參見圖五，其中期刊論文幾乎皆有全文，國科會計畫成果報告則因早期繳交紙本報告，全文比例約為 40%，但是近期的報告皆為數位全文，可預期隨著時間遞移，全文比例將越來越高。至於長期的推動規畫上，則透過學校行政程序設定為學術出版品的典藏系統，建立學科館員與系所的聯絡窗口，協助發函後繼續追蹤系所窗口聯絡人事宜，聯繫指標系所老師及聯絡人參與說明會，應主動邀集師生響應機構典藏系統內容的建置，並持續宣導機構典藏的相關優點，鼓勵全校師生上傳其研究成果。

(五) 資料呈現的問題

Dspace 雖可支援多種語言顯示，但並未考慮到所有情況，NTUR 仍須針對許多功能將其中文化，例如書目資料欄位名、授權書、線上說明、社群與類別的中英對照等。此外，其 Dspace 內建之全文剖析器 PDFBox 對中文的支援亦相當有限，以 NTUR 第

一批匯入測試的一萬筆資料為例，約有 70% 以上的資料無法正確解析，以致產生大量亂碼資料。由於亂碼對於使用者而言，會產生極大的困擾，也會降低使用者的使用意願，甚至懷疑系統的品質，因此，NTUR 開發團隊必須嚴謹認真地處理這個問題。開發團隊為了節省開發系統的時程，決定採用廠商提供的產品，而不自行開發 PDF 剖析器，經過測試了各不同廠商提供的解決方案後，發現由 Snowtide Informatics 開發的 PDFTextStream 效果最佳，其正確率可達 95% 以上，請參見圖六，所以 NTUR 系統採用其為全文檢索的核心剖析器。然而，PDFTextStream 是商業產品，必須付費使用；因此，未來 NTUR 系統的釋出版本仍然會使用 PDFBox，欲使用高品質的 PDF 剖析器，個別學校仍須自行付費使用。

(六) 學術文獻的授權

在 Davis & Connolly 的文章中提到，對著作權定義混淆不清、擔心侵犯著作權，是使用者拒絕採用機構典藏系統的主因之一。(Davis & Connolly, 2007) 為消除使用者在版權方面的疑慮，臺灣大學機構典藏系統採用項潔於「臺灣機構典藏發展芻議」一文提出的三權分立的概念(項潔、洪筱盈, 2005)，也就是釐清學術文獻的自我典藏權利、全文索引權利、全文取得權利。目前大多數的期刊出版社至少允許前二種權利，亦即作者及其母機構仍保有自我典藏與進行全文索引的權利，唯在著作權已讓渡給出版社的情形下，不得在機構典藏系統進行全文取閱，而是透過連結引導使用者連回原出版社，依照其規定進行取閱。

英國 SHERPA/ROMEO 計畫詳細調查各期刊出版社的授權現況，根據其調查結果，目前已有九成以上的出版社允許作者以自我典藏的方式將論文儲存在母機構的典藏系統(SHERPA/ROMEO, 2006)，代表注重作者著作權的議題已逐漸成為趨勢，出版社也不再獨自壟斷著作權，共同尋求雙贏的方案。

五、結論與建議

機構典藏由一項議題到實際建置機構典藏系統，已經逐漸受到學者專家以及學術機構的重視，藉由機構典藏系統來保存機構的學術產出，並作為對外

展示研究成果的平台，以強化機構的學術影響力，除此之外，結合全球的機構典藏系統，達成學術論文的開放取用，降低學術研究資源取用的成本，加速學術研究資源的傳播。

大學圖書館做為大學的資訊與知識的典藏中心與傳播中心，自應負擔推行機構典藏的觀念，並且建置機構典藏系統。目前許多國家也逐漸正視其重要性，開始以國家的力量投入資源、建置機構典藏系統。中國大陸推動的『中國知識資源總庫』，雖然不能算是機構典藏系統，或許可稱為國家學術研究典藏系統，涵蓋 300 所大學的 8000 種以上期刊，研究者能夠很容易地取得中文學術研究論文，我們應正視其所帶來的衝擊與影響，如果臺灣的中文研究文獻取閱相對困難，將直接對臺灣學術界的研究曝光程度造成嚴重的影響，導致在國際學術研究地位逐漸被邊緣化。

臺灣大學圖書館 NTUR 開發團隊從眾多的選擇方案中，決定使用 DSpace 系統套件，並大幅新增許多客製化功能，有其本身的考量因素，並非是唯一的建置機構典藏系統的解決方案，仍然有許多不同的作法，可供擁有不同資源的學術機構選擇，最重要的是機構典藏的觀念推廣與系統建置。

機構典藏概念的推廣，仍然存在著許多障礙，有使用者對於建置機構典藏系統有所疑慮，例如重複的發佈管道、學習曲線、著作權認知不清、資料品質不一致等問題存在 (Davis & Connolly, 2007)，但隨著機構典藏的觀念逐漸成熟，機構典藏系統的建置越來越多，由點到線到面，形成完整的、開放的、整合的學術研究資源取用環境後，前述的問題將僅僅是這個過程的小插曲。

誌謝

感謝 NTUR 計畫團隊成員：項潔館長、張素娟主任、郭美玲編審、江玉婷、何庭輝、高巧倫、劉秋菊、洪筱盈、何浩洋、陳詩麗等人的貢獻，以及臺灣大學圖書館學科館員：林鳳儀、劉雅姿、葉欣怡、劉益君、李百珣、劉仁翔、楊韻蓉、何燦君、翁雅芳、黃文琪、黃滢芳、劉應琳、林郁婕、陳鳳如、廖敬華、蔣禮芸、杜宜凌等同仁的協助。

參考文獻

- Davis, P. M. & Connolly, M.J.L. (2007). Institutional Repositories: Evaluating the Reasons for Non-use of Cornell University's Installation of DSpace. *D-Lib Magazine*, 13(3/4). Retrieved June 4, 2007, from <http://www.dlib.org/dlib/march07/davis/03davis.html>.
- DSpace (2007). *Who's Using DSpace?* Retrieved June 4, 2007, from <http://wiki.dspace.org/index.php/DspaceInstances>.
- Ghose, S. B. & Das, A. K. (2006). Open access and institutional repositories – a developing country perspective: a case study of India. In *Proceedings of the 72nd IFLA General Conference and Council*, Seoul, Korea. Retrieved July 12, 2007, from http://www.ifla.org/IV/ifla72/papers/157-Ghosh_Das-en.pdf.
- Hallam, Henry (1970). *Introduction to the Literature of Europe in the Fifteenth, Sixteenth, and Seventeenth Centuries*. New York: Frederick Ungar.
- ISSN International Centre (2006). *ISSN Statistics*. Retrieved July 12, 2007, from <http://www.issn.org/en/node/330>.
- King, A. B. (2003). *Speed up your site: Web site optimization*. Indianapolis, Ind: New Riders.
- Lynch, C. A. & Lippincott, J. K. (2005). Institutional Repository Deployment in the United States as of Early 2005. *D-Lib Magazine*, 11(9). Retrieved July 12, 2007, from <http://www.dlib.org/dlib/september05/lynch/09lynch.html>.
- Lynch, C. A. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. *ARL*, 226, 1-7. Retrieved July 12, 2007, from <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>.
- Markey, K., Rieh, S. Y., et al. (2007). *Census of*

Institutional Repositories in the United States: MIRACLE Project Research Findings. Retrieved June 4, 2007, from <http://www.clir.org/pubs/abstr/act/pub140abst.html>.

Sherpa/Romeo (2006). *Publisher copyright policies & self-archiving*. Retrieved July 12, 2007, from <http://www.sherpa.ac.uk/romeo.php>.

Tansley, R., M. Bass, et al. (2006). *DSpace System Documentation*. Retrieved July 4, 2007, from <http://www.dspace.org/technology/system-docs/>.

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 July 12, 2007, from <http://www.dlib.org/dlib/september05/westrienen/09westrienen.html>.

教育部 (2006)。大專院校概況統計。Retrieved June 4, 2007, from http://www.edu.tw/EDU_WEB/EDU_MGT/STATISTICS/EDU7220001/data/serial/u.xls?UNITID=139&CATEGORYID=260&FILEID=130550&open

項潔、洪筱盈 (2005)。臺灣機構典藏發展芻議。《教育資料與圖書館學》，43(2)，頁 173-189。

附錄

表一 歐美等十三國建置機構典藏比例

國家	IR 數量	大學總數	建置百分比	平均文件量
Australia	37	39	95	<i>not reported</i>
Belgium	8	15	53	450
Canada	31	<i>not reported</i>	-	500
Denmark	6	12	50	<i>not reported</i>
Finland	1	21	5	<i>not reported</i>
France	23	85	27	1000
Germany	103	80	100	300
Italy	17	77	22	300
Norway	7	6	100	<i>not reported</i>
Sweden	25	39	64	400
The Netherlands	16	13	100	3,000 / 12,500
United Kingdom	31	144	22	240
U.S.A.	<i>not reported</i>	261 / 2364	-	<i>not reported</i>

資料來源：van Westrienen, G. & Lynch, C.A. (2005). Academic institutional repositories: deployment status in 13 nations as of mid 2005. *D-Lib Magazine*, 11(9). Retrieved June 4, 2007, from <http://www.dlib.org/dlib/september05/westrienen/09westrienen.html>.

表二 世界各國建置機構典藏與保健/生物科學領域的機構典藏一覽表

項目 國家	IR 數量	生物科學類	保健類	研發佔國民所得比重
1. Australia	17	9	12	1.54
2. Austria	3	1	1	2.19
3. Belgium	9	4	2	2.33
4. Brazil	10	5	5	1.04
5. Canada	18	5	8	1.87
6. Chile	2	0	2	0.57
7. China	2	1	1	1.22
8. Columbia	2	1	0	0.10
9. Denmark	3	3	0	2.52
10. Finland	3	2	2	3.46
11. France	21	10	5	2.26
12. Germany	39	18	16	2.50
13. Greece	1	0	0	0.65
14. Hungary	3	0	0	0.95
15. India	8	3	1	N.D.
16. Ireland	2	1	0	1.13
17. Israel	1	0	0	4.90
18. Italy	15	4	3	1.11
19. Japan	4	2	2	3.12
20. Mexico	1	0	0	0.39
21. Namibia	1	0	0	N.D.
22. Netherlands	14	5	8	1.88
23. New Zealand	1	1	1	1.16
24. Norway	3	2	2	1.67
25. Pakistan	1	0	0	N.D.
26. Portugal	3	1	0	0.94
27. Russia	1	0	0	1.28
28. Singapore	1	0	0	2.15
29. Slovenia	1	0	0	1.53
30. South Africa	4	3	3	N.D.
31. South Korea	0	0	0	2.64
32. Spain	5	3	3	1.03
33. Sweden	18	9	11	4.27
34. Switzerland	4	2	2	2.57

表二 世界各國建置機構典藏與保健/生物科學領域的機構典藏一覽表（續）

國家	項目	IR 數量	生物科學類	保健類	研發佔國民所得比重
35.	United Kingdom	56	23	18	1.87
36.	United States	102	35	34	2.67
37.	Venezuela	2	1	2	N.D.

資料來源：Oliver, K. B. & Swain, R. (2007). Directories of Institutional Repositories: Research Results & Recommendations. In Proceedings of World Library and Information Congress: 72nd IFLA General Conference and Council. Retrieved June 4, 2007, from http://www.ifla.org/IV/ifla72/papers/151-Oliver_Swain-en.pdf

表三 亞洲各國建置機構典藏系統一覽表

國家	IR 數量	國家	IR 數量
China	1	Russia	2
Japan	26	Pakistan	1
India	25	Philippines	1
Taiwan	7	Singapore	1
Hong Kong	3	Indonesia	1
Israel	2	Korea	1

資料來源：作者整理自 Eprints、OpenDOAR、OAister、DSpace、以及 GHOSH, S. B. & DAS, A. K. (2006). Open access and institutional repositories – a developing country perspective: a case study of India. In Proceedings of World Library and Information Congress: 72nd IFLA General Conference And Council. Retrieved June 4, 2007, from http://www.ifla.org/IV/ifla72/papers/157-Ghosh_Das-en.pdf

表四 臺灣各校機構典藏

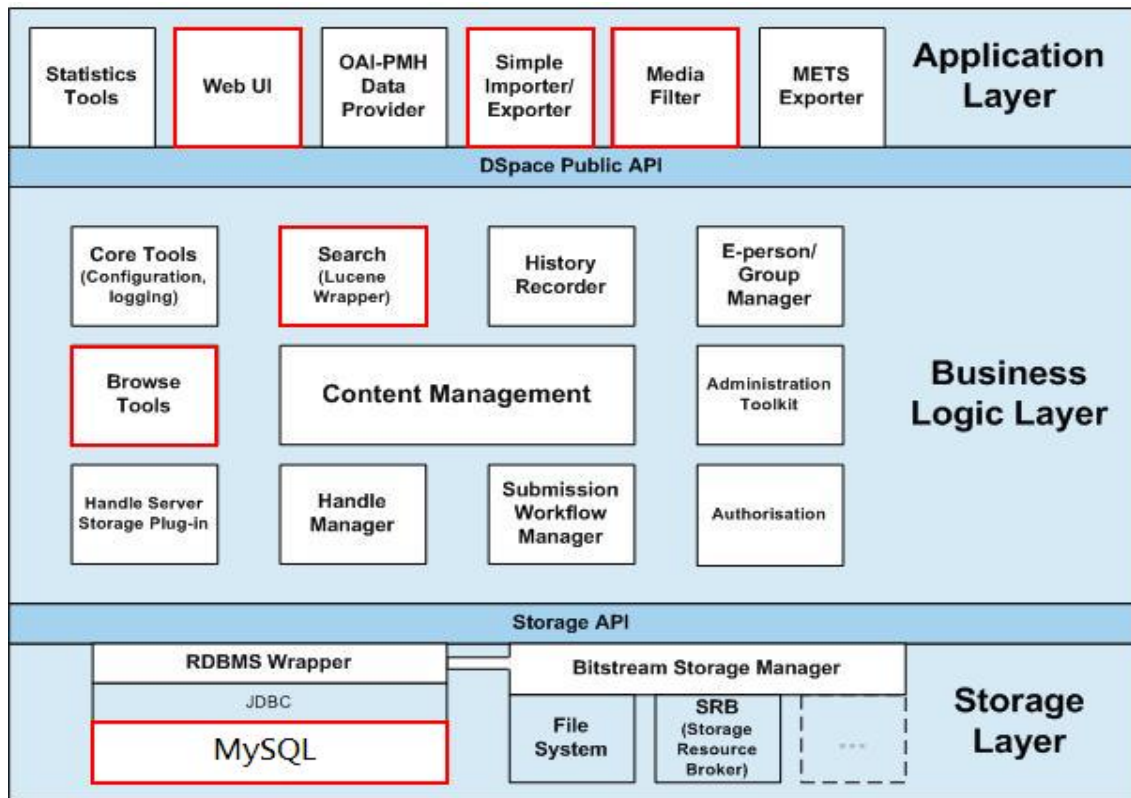
學校	資料筆數	採用系統	網址
臺灣大學	32,525	NTUR	http://ntur.lib.ntu.edu.tw/
逢甲大學	3,319	DSpace	http://140.134.132.124:8080/dspace/
崑山科技大學	2,619	DSpace	http://dspace.lib.ksu.edu.tw:8080/
清華大學	4,214	DSpace NTUR	http://irsite.lib.nthu.edu.tw:8080/ http://140.114.72.31/dspace/
交通大學	測試中	NTUR	http://140.113.39.213/dspace/
政治大學	測試中	NTUR	http://140.119.115.26/

資料來源：作者整理自各大學機構典藏系統網站。由於系統的資料會持續增加，統計數據僅供參考；崑山科技大學也將採用臺灣大學改寫之 NTUR 系統。

表五 現有機構典藏系統典藏文件之概況

文件數量	實際上線	規畫與測試中
<501	41.9%	68.7%
501~1,000	8.3%	10.4%
1,001~2,000	11.1%	4.2%
2,001~3,000	11.1%	2.1%
3,001~4,000	5.6%	2.1%
5,001~5,000	2.6%	4.2%
>5,000	19.4%	8.3%
總計	100%	100%
機構典藏系統數量	48	70

資料來源：Markey, K., Rieh, S. Y., et al. (2007). Census of Institutional Repositories in the United States: MIRACLE Project Research Findings. Retrieved June 4, 2007, from <http://www.clir.org/pubs/abstract/pub140abst.html>.



圖一 NTUR 系統架構

資料來源：修改自 http://www.dspace.org/index.php?option=com_content&task=view&id=145

English | 正體中文 | 簡體中文 | 全文筆數/總筆數: 15885/32533
造訪人次: 20870 線上人數: 7
© Powered By DSPACE, MIT. Enhanced by NTU Library IR team

搜尋範圍 全部NTUR [搜尋] 進階搜尋 登入 · 上櫃 · 說明 · 關於NTUR · 管理

瀏覽全部
 社群與類別
 題名
 作者
 日期

臺大機構典藏 NTUR > 主頁

臺灣大學 機構典藏 NATIONAL TAIWAN UNIVERSITY REPOSITORY

NTUR中的社群 [15885/32533] [全文筆數/總筆數]

- + 文學院 [765/1369]
- + 理學院 [1828/4109]
- + 社會科學院 [353/726]
- + 醫學院 [3272/5884]
- + 工學院 [2330/5446]
- + 生物資源暨農學院 [1336/2964]
- + 管理學院 [557/1012]
- + 公共衛生學院 [230/475]
- + 電機資訊學院 [4198/5921]
- + 法律學院 [155/323]
- + 生命科學院 [444/1143]

上傳排行 下載排行

社群

- 電機工程學系 [2695/3693]
- 資訊工程學系 [940/1262]
- 物理學系 [637/998]
- 醫學院附設醫院 [611/652]
- 機械工程學系 [511/1216]

作者

- 陳良基 [181/210]
- 郭華仁 [139/146]
- 傅立成 [128/174]
- 貝蘇章 [100/138]
- 郭斯彥 [86/107]

圖二 NTUR 系統前台網站首頁

English | 正體中文 | 簡體中文 | 全文筆數/總筆數: 15885/32533
造訪人次: 20897 線上人數: 4
© Powered By DSPACE, MIT. Enhanced by NTU Library IR team

搜尋範圍 全部NTUR [搜尋] 進階搜尋 登入 · 上櫃 · 說明 · 關於NTUR · 管理

瀏覽全部
 社群與類別
 題名
 作者
 日期

臺大機構典藏 NTUR > 依日期瀏覽

跳至: (選擇年份) (選擇月份) 搜尋
 或輸入年份:
 從最近的開始 由舊到新排序

顯示項目 1-25 / 30055. (共1184頁)
 1 2 3 4 5 6 7 8 9 10 >>>
 每頁顯示 [10 | 25 | 50] 項目

日期	題名	作者
1963-04	Neutron Total Cross Section of Arsenic at 14 Mev	Hsu, Y.C.; Huang, C.Y.; Tseng, P.K.; Tai, Y.K.; 許垂基; 黃家怡; 鄭白昆; 戴運軌
1963-04	New Scheme for Calculating Nuclear Radii	Hwang, Jenn-Lin; 黃振麟
1963-10	Absence of Zero Sound in Superfluid Fermi Systems	Kobe, Donald H.
1963-10	Counting Efficiency of Nuclear Multiplate Camera	Huang, Chia-Yi; Hwang, Jenn-Lin; Hsu, Yuin-Chi; Tai, Yuin-Kwei; 黃家怡; 黃振麟; 許垂基; 戴運軌
	Radii, Surface Diffuseness, and Binding Energies of	Hwang, Jenn-Lin; Yang, Yeh-

圖三 依日期排序檢視文件

[類別瀏覽](#) | [作者瀏覽](#) | [年代瀏覽](#) | 臺大機構典藏 NTUR > [搜尋結果](#)

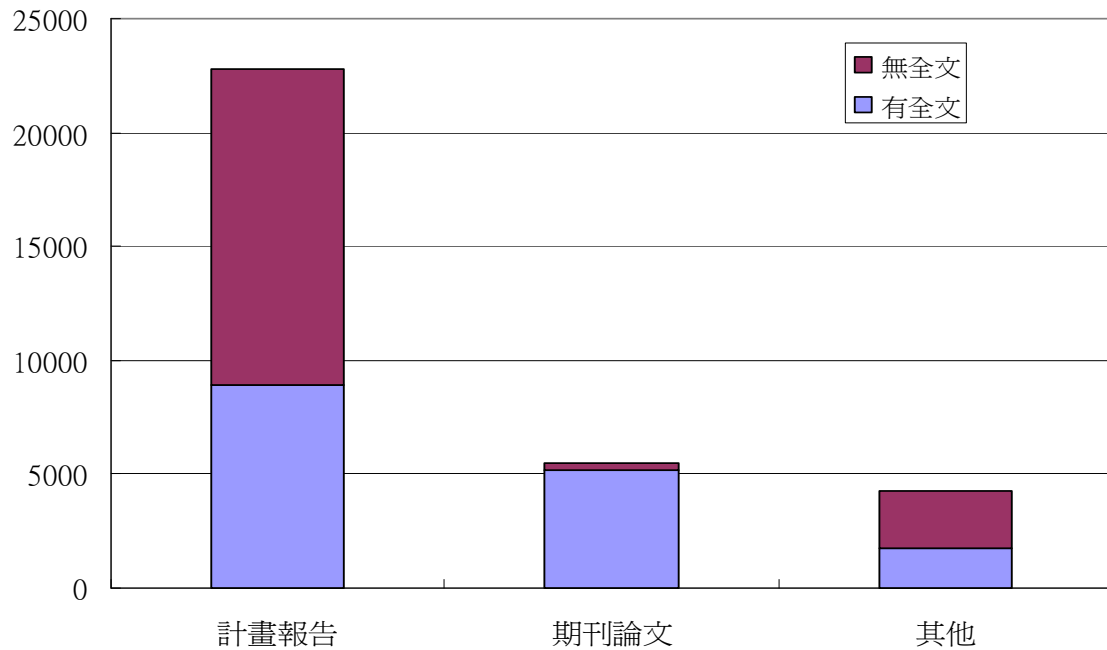
[依文章數排序](#) | [依學院瀏覽](#)

[只搜尋有全文項目 | 搜尋所有項目]

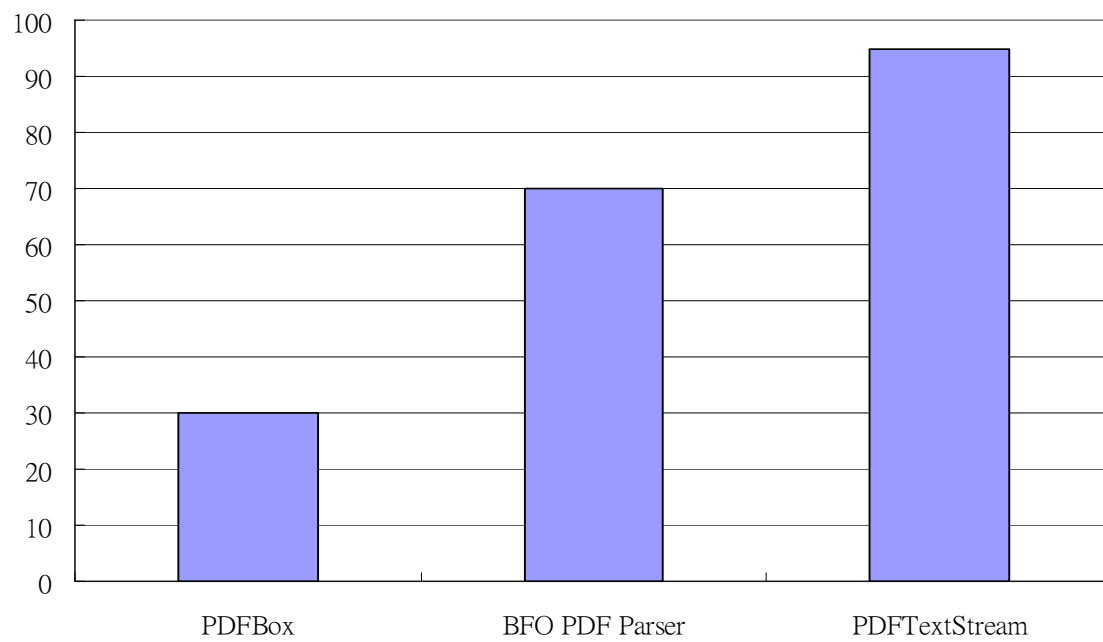
結果 1-25 / 32.
 1 2 下一頁
 每頁顯示 [10 | 25 | 50] 項目

- 1. 台灣地區原住民飲酒問題研究-台灣地區原住民飲酒問題之家族遺傳研究**
 陳為堅, 1996-07-31
 [公共衛生學院流行病學研究所] 國科會計畫
 陳為堅 台灣地區原住民飲酒問題研究-台灣地區原住民飲酒問題之家族遺傳研究 1996-07-31 臺北市: 國立臺灣大學公共衛生學院流行病學研究所
- 2. 臺灣原住民基本法之回顧與展望**
 王泰升, 1997
 [法律學系] 國科會計畫
 王泰升 臺灣原住民基本法之回顧與展望 1997 臺北市: 國立臺灣大學法律學系暨研究所
- 3. 台灣原住民族資源利用與永續發展之研究-總計畫暨子計畫五: 原住民族地區環境變遷脆弱性分析-新竹縣尖石鄉之個案(1)**
 張長義, 2006
 [地理環境資源學系] 國科會計畫
 張長義 台灣原住民族資源利用與永續發展之研究-總計畫暨子計畫五: 原住民族地區環境變遷脆弱性分析-新竹縣

圖四 檢索結果後分類--依作者瀏覽



圖五 NTUR 典藏之文件全文比例



圖六 中文資料剖析正確率之比較