

## 知識寶庫掘井人：「2017 臺灣學術機構典藏研討會」與會札記

數位資訊組 曾昱博、王鑫一

時間：106 年 11 月 24 日(五)13:00 至 17:00

地點：中山大學圖資大樓 B1 視訊研討室

主辦單位：國立臺灣大學圖書館、國立中山大學圖書與資訊處

大專院校建置機構典藏系統，是各單位了解整體機構研究能量、行銷研究成果的重要管道之一，臺灣學術機構典藏(Taiwan Academic Institutional Repository, 簡稱 TAIR)為臺灣大學圖書館受教育部委託建置之臺灣學術成果入口網，建構至今已歷時十年，透過分散建置、整合使用，參與單位已日益茁壯。

會議引言人臺灣大學圖書館陳光華館長提及：目前全臺灣共計有 138 所單位參與 TAIR，全文量占資料總比數的 66%。隨著各項引文數據、使用者介面、使用者體驗逐漸受到重視，該系統也在去年底發表新版本－RC7 版，以提供更適切的服務。然而，根據團隊服務案件統計及分析各單位之反饋意見，了解到現有版本在系統發展上所遭遇之瓶頸，團隊多方探索與研究，找到如何解決擴充功能受限之方案。

本次研討會除邀請幾所參與館進行經驗分享外，同時 TAIR 計畫團隊亦發表機構典藏系統的新發展，各講者分享之主題詳見如下表：

題目	講者
淡江大學機構典藏經驗分享	淡江大學圖書館數位資訊組組長 林泰宏
知識匯 Knowledge Hub	國立清華大學科技管理學院服務科學研究所教授兼圖書館館長 林福仁
南華大學機構典藏建置經驗分享	南華大學圖書館數位支援組組長 柯婉懿
IR 可以做得更多－NTUScholars 初登場	國立臺灣大學圖書館系統資訊組編審 江玉婷

以下將針對機構典藏現況與困境、清華大學知識匯，以及臺灣學術機構典藏系統新契機三個面向分別說明之。

### 一、機構典藏現況與困境

茲根據講者分享内容，歸納出各校執行機構典藏作業過程之共通點。

機構典藏收錄的資料類型，大致可區分為期刊論文、會議論文、專書、專書之單篇、研究報告、學位論文、計畫報告、演講等常見類型，有些單位同時收錄視聽著作、電腦程式、專利、考古題、課程教材、特色館藏等資料。部分則會嚴格把關及區分所收錄之資料類型。

綜觀各校執行機構典藏過程所遭遇之困境，可分為以下幾個面向：

### 1.多套相似系統並行導致資料同步問題

校內同時有多個管理著作目錄之系統，例如：教師歷程檔、研究獎勵系統等，甚至圖書館也有機構典藏、學術庫等系統並行，多套系統增加教師維護上的負擔，也造成人員作業重工問題。雖可透過系統串連進行資料同步，但可能導致時間差，例如：教師建檔、館員補正、再經過教師修改，加深資料同步問題的惡性循環。此外，系統間互匯資料，亦可能導致資料複本問題，需要進一步查重。面對不同系統，各有其核心功能，台灣大學圖書館構想出的解決方案則是採用同一資料庫存放各系統聯集的資訊，分別設計使用介面，各取所需。

### 2.建檔資料補遺

前述提及校內同時有多套相似系統，可分別匯入書目資料至機構典藏，然而，部分系統因本身不具備上傳全文檔功能，或未要求教師或研究者上傳全文。故自其他系統批次匯入書目資料後，圖書館員需透過人工查找全文或外部連結加以補足。而在教師自我建檔時，欄位建檔狀況不一，圖書館員仍會逐筆檢視及補正。因此，書目資料品質問題，也是機構典藏作業過程中，遭遇到的重要問題之一。

### 3.人員不足與異動頻繁

機構典藏資料的徵集需花費不少時間與人力，一旦人員異動頻繁，交接過程即會耗費時日，而部分作業如彙整著作、搜尋全文、勘誤、推廣行銷、聯繫與跟催資料等，也是需要長時間、循序漸進式地執行。

面對資料蒐集與建檔問題，講者紛紛提出希望透過開放型研究者及投稿者識別碼(Open Researchers and Contributor ID，簡稱 ORCiD)解決，淡江大學採取要求新任教師提供 ORCiD ID 的方式追蹤與連結教師個人研究歷程。清華大學林館長建議與會單位可嘗試加入 ORCiD member，此舉可將同為 ORCiD member 的資料庫廠商，其收錄之本校研究者的產出資料匯入，但同時需要考量資料複本問題。台灣大學也已加入 ORCiD member，其推廣方式係為研究人員批次申請 ORCiD 帳號，運用 API 與校內系統互匯資料，並且由被授權人員為研究者管理帳號與 ORCiD 研究資訊，透過系統直接抓取所需清單、數據、個人資料。

在全文資料徵集部分，淡江大學特別補充該館會仔細確認出版社對典藏政策的規範，例如要求典藏作者原稿且尚未經過審核之版本(pre-print)、作者原稿經過審核與修改的版本(post-print)或可典藏出版社排版後之版本(publisher's version)，為上傳的全文做一道把關。

## 二、清華大學知識匯 Knowledge Hub

清華大學為整合校內系統、提供分析統計與完整呈現全校性研發成果，在圖書館、研究發展處、計算機與通訊中心的合作下，共同建置「知識匯 Knowledge Hub」系統。合作團隊以研究者知識創造週期之 SECI 模型理念發想，其意旨研究社群間「隱性知識」及「顯性知識」相互轉化的過程，歷經了四個階段：1.社會化(socialization)；2.外顯化(externalization)；3.組合化(combination)；4.內隱化(internalization)。研究者的隱性知識，藉由社交、合作、撰寫等外顯化，外顯之成果諸如 paper、專利等，可供查找、下載或透過交流等方式，將其整合成各種新的顯性知識，再被利用、內化為研究者的內隱知識，藉此不斷循環著。而從知識管理的角度而言，如何獲取機構中個人的知識產出、如何呈現整體機構的外顯知識、知識創造的過程為何、如何評斷知識產出、何為知識傳播的範圍等，也是建置知識匯過程中考量之重點。

林館長分享清華大學 Knowledge Hub 設計理念出自於以研究者(即知識產出者)為重心，如同每位藝術家皆有各時期的代表作品，科學家、人文社會學者亦可，此概念的延伸即創造「個人品牌」、展現教師特色。長遠下來，期望讓教師、研究者從被動服務到主動參與，同時，可協助圖書館員掌握學科知識發展的脈絡，加值學科服務內涵，進階參與學術社群活動。

清華大學知識匯建置過程，考量現有廠商已開發之系統，如 Academic Hub / Pure Experts，無法符合校方、教師、研究者所需，故採用與廠商合作開發之方式進行，透過教師訪談了解需求，進而設計系統。知識匯共分為四大重點：1.「匯聚」校內/行政教學系統、教師個人簡歷、提供校內研發成果趨勢分析，讓機構典藏變成一項服務系統，而非監控教師研究產出的統計系統；2.建構知識系統、個人書房功能，運用 text mining 技術主動推薦文獻，透過組織、典藏及分享，「掌握」研發新知；3.建立合著者、研究計畫、指導關係網路圖，及學者、領域知識地圖，以有效「連結」相關領域概念；4.運用上述網路圖、知識地圖，挖掘潛在合作對象，促成由校內至全球化之跨域研發新契機，將機構研究成果「行銷」國際。

## 三、臺灣學術機構典藏系統新契機

臺灣學術機構典藏係以 DSpace 1.4 版為基礎進行中文化的系統，當前由於機構典藏的資料類型日益多元化，資料度用的概念逐漸受到重視，該版本架構較老舊，不便於擴展功能，故已不敷使用，即使新版本 DSpace 雖已具備所需功能，考量語系無法直接移植使用，需有足夠開發人力方可完備。因此，

TAIR 計畫團隊經過研究，調查相關開源系統，將選擇以 DSpace 為基礎的擴充系統“DSpace-CRIS”作為系統改版的規劃。當前已採用此系統的單位將近 100 個，如香港大學、香港理工大學、新加坡大學等，該系統不僅能整合 ORCID、與研究資料互連，亦可與外部評核數據(如 Altmetric、PlumX、Scopus) 連結，提供引文計量數據。茲摘錄會議中介紹 DSpace-CRIS 之功能特色如下：

1. 彈性物件管理：物件包含 DSpace 核心資料原件(item)及研究相關實體(entities)，管理面則包含物件/欄位存取權限設定、物件詮釋資料自動參照及互連設定、個人資料下細分巢狀架構。
2. 相關物件連結及群聚：採用語意及多重性關聯，可設定 item(如研究成果)及 entities(如研究人員、研究計畫、期刊)之間的關聯性及權威資料參照，進而在研究成果頁面顯示相關的研究人員及期刊，或在研究人員主頁面呈現研究成果、計畫或期刊。
3. 詮釋資料改善：打破扁平資料模式，建立詮釋資料階層式巢狀欄位，可透過權威資料連結機制，將研究成果連結至外部物件，如透過 ORCID 連結研究者資料與產出，並且提供各詮釋資料欄位輸入格式之驗證。
4. 與各系統合作：透過 API 介接方式串聯與匯入外部資料，如書目、研究者資料、控制詞彙等。
5. 外部評核指標數據：利用研究物件著錄之識別碼，如 DOI、PMID、EID 及 UT，自動抓取 SCOPUS、WOS、PubMed、Altmetric、Google Scholar 之書目計量資訊，獲取之數據存於資料庫中，以進一步演算及分析。
6. 重複資料處理：透過查重機制，在單筆資料上傳時，支援重複資料提醒，亦提供可能相似之清單。此外，系統記錄使用者回饋之重複資料，以供管理者判斷。
7. 統計模組：除對單一研究物件統計其瀏覽、下載、被引用次數外，可彙整單位、系所整體產出物件之檔案總數量、瀏覽、下載次數，使統計數據更具階層化分析。此外，統計模組亦支援前述提及之互聯資料功能統計。
8. 研究資料(Dataset)處理：研究資料除了單一物件、出版品的 publication 層次外，同時涵蓋研究數據集的 data 層次，或不同資料類型的層次，該系統可支援的檔案格式較為多樣化。
10. HTML meta tag 標註：透過標註部分欄位，如作者資訊 ORCID、VIAF records、物件識別碼 DOI、出版品頁面等，可提升資料能見度，並且

揭示物件之間的關聯性。

11. ORCiD 整合：透過 ORCiD 之 member API，可雙向傳遞資料，達成客戶端主動提取(pull)及系統端主動推播(push)的運作機制，相較於舊版，新版系統可透過 API 自動介接國際出版社之期刊文獻及批次匯入 ORCiD 研究者個人資訊、研究成果、計畫資料。

12. 研究人員資料的管理與推廣：使用者登入後，系統將自動建立研究人員資料，並且透過 ORCiD 自動同步已存在的資訊。

當前 TAIR 計畫團隊已加入官方開發團隊，執行 DSpace-CRIS 共同開發，將考量本地所需進行客製化、建立新系統雛型、介面視覺化設計、新舊系統資料轉移等，初期採用新舊系統並行方式，未來將研擬是否可嘗試建立雲端系統，屆時亦會徵詢各單位之意願及相關意見進行評估。

#### 四、結語

校務研發能量除仰賴各研究者的知識結晶外，更需要有效系統協助研究者在知識創造週期中更便利蒐集新知、彙整產出，乃至跨域整合，筆者認為建構知識網路地圖、促成跨域協同合作，是近年來機構典藏逐步前進的里程碑。從分享者的經驗談，可窺見清華大學知識匯有別於往常以資料主體為重心管理，反觀採取以人為本的思維在建置系統與提供服務，站在教師及研究者的立場思考，方能達到雙贏、讓機構更向前邁進一大步，值得每位從事機構典藏的人員效法。

會中探討到多種相似資料庫的並行及機構典藏定位之問題，近年來隨著各式需求的衍生，紛紛發展出學術庫、特色典藏網站、磨課師網站等，而前述也提到部分單位會收錄特色館藏、考古題、課程教材、電腦程式等，此亦不乏是機構及教師在教學、研究、典藏的成果。因此，機構典藏收錄的範圍界定為何？關乎是否整合以上系統之書目資料及物件連結，以及未來是否能透過物件關聯性，串連機構的知識網絡，是個值得深思的議題。

#### 參加「2017 年台灣 OCLC 管理成員館聯盟會員大會」心得報告

採編組 王曉萍

時間：民國 106 年 12 月 6 日

地點：國立台灣大學集思台大會議中心 B1 蘇格拉底廳

主辦單位：國立台灣大學圖書館

協辦單位：文崗資訊(股)公司、飛資得知識服務(股)公司