



數位典藏國家型科技計畫 · 內容發展分項計畫

國家數位典藏聯合目錄系統建置計畫討論會

會議記錄

會議時間：92年7月14日 下午2點整

會議地點：中研院資訊所 101 會議室

主席：黃銘崇

出席：楊美華、王梅玲、何建明、范紀文、黃國倫、王祥安、陳亞寧、陳淑君

列席：黃慧娟、林麗虹、沈漢聰、許雅婷、洪嘉培、汪怡慧

紀錄：汪怡慧、洪嘉培

一.聯合目錄工作小組報告進度

楊美華教授報告：每週四召開例行性工作會議。

王梅玲教授報告：

依據上次會議討論結果，已根據范紀文先生的說明，擬出聯合目錄系統需求書草案，內容尚有需 Metadata 小組幫忙之處。希望系統規格需求書及程式撰寫能夠提早完成，以推進聯合目錄整體時程。目前需要 metadata 小組及早決定核心欄位，因為這關係聯合目錄的呈現結果。依上次會議結果，目前聯合目錄建置，以中研院、故宮兩單位之計畫為主，新聞、地圖、地質三個主題小組暫不包括在內。

聯合目錄系統需求書草案，為參酌國內外計畫，包括 American Memory、PANDORA、Colorado Digitization Program、AMICO(Art Museum Image Consortium)、楊英風數位美術館等，並結合去年 91 年度聯合目錄的系統功能，所提出之初步分類架構與查詢架構。希望針對整個數位典藏計畫共享整個聯合目錄的成果。

聯合目錄查詢系統主要功能為，以主題、時間、地點、典藏單位 等整合式資料，讓使用者透過 www 瀏覽與查詢(查詢包括簡易查詢與進階查詢)瞭解整個計畫(包括 12 個主題內容、9 個典藏單位、50 個典藏計畫)之內容；典藏單位的參與者，可依照其權限、著作權所有的狀況來決定目錄所要提供的資訊。

聯合目錄之內容，包括：

- (一) 目錄的簡要紀錄：主要取決於 metadata 所提供的核心欄位(依據 DC 內資料為主)，需請 metadata 小組提供。



- (二) 提供小圖示影像：研究國內外共五個網站內容，發現 AMICO、楊英風數位美術館有提供小圖示影像，因此希望聯合目錄之亦能提供此功能。可將單位的影像檔以 thumbnail 存檔作成圖示索引，讓使用者可以瀏覽圖示以增加吸引力，並連結至原典藏單位觀賞數位物件；而使用權限全在於各典藏單位，由典藏單位決定是否開放。

聯合目錄系統的功能，包括分類瀏覽與檢索功能兩項。

分類瀏覽方面，包括以下五種瀏覽架構：

- (一) <12 個主題內容與細部分類>：內容主題分 12 類，每主題數量不一，以每類數量 2000 件以下較理想，提出分類依據，並提出簡單構想：動、植物以界門綱目科屬種；人類學以南島語而言分語族、文類、語言；檔案請故宮、近史所討論合理架構；器物書畫根據故宮；金石拓片以故宮與中研院的整合、善本古籍分類型、專題、版本；考古分紀錄內容與紀錄整理，除 12 主題內容分類與第二層分類之外，需再第三層分類，需是從 Metadata 裡面產生的，而且是大家覺得合理的。
- (二) <機構>：加上 12 主題，再加上 50 個典藏計畫。
- (三) <時間>：從史前時代到民國。
- (四) <地理>：分台灣中國與其他。共計 12 個地理的瀏覽。
- (五) <數位化物件類型>：參考國家數位化典藏以往的檔案格式，因此提出數位化物件類型一項。

檢索功能方面，內容分述如下：

- (一) 提供欄位檢索、模糊檢索(含同音、同義、簡繁俗體等)、不限欄位之全文檢索；
- (二) 共通查詢模式包括簡單查詢及進階查詢，進階查詢欄位包括藏品名稱、作品類型、創作者、主題、時間、地理、典藏單位、數位典藏計畫名稱 等。
- (三) 系統具有布林邏輯查詢、切截查詢、完整查詢之功能，並以時間、藏品類型、地理、數位物件檔案格式、館藏單位等條件縮小查詢範圍，限制條件盡量以點選方式處理，不需使用者鍵入資料。
- (四) 查詢結果資料以共通欄位呈現，並可設定是否要連結更豐富的原始 metadata。
- (五) 系統可讓使用者選擇呈現資料的排序方式，並可前後翻頁顯示。可選擇排序方式，含相關度、主題、創作者、時間、典藏單位等。此外可選擇以文字清單呈現或圖示清單呈現。



二.metadata 工作小組報告進度

陳淑君小姐報告：

中研院與故宮 11 項主題計畫後設資料格式與 Dublin Core 標準對照表相關文件、以及 Dublin Core 標準對照後設資料需求規格書、主題計畫後設資料以 Dublin Core XML 文件轉出測試 等已經完成。

未來 7/15-8/20 會進行主題計畫後設資料轉出 Dublin Core 元素確認；8/21-9/5 進行 Dublin Core XML Schema 轉置軟體之測試與研析(對象：聯合目錄系統組)；9/8-30 進行 Dublin Core XML Schema 轉置軟體之測試與研析(對象：主題計畫的系統開發單位)。

三.討論

(一)對於數位化物件類型的呈現意義上並不大？因為原件數位化後的媒體格式有很多種。當初執行數位化時，就已經時間針對器物、拓片類型來數位化，如果還要再特別註明類別(type)，意義不大。(黃銘崇)

王梅玲：以 American Memory 而言，是兩者都有，黃先生建議採用數位化之前的類型。讓若要註明 tape，則需可以進行加碼與編碼，而加碼編碼系統如何還需再議。

何建明：當初在執行只針對單一類型作 metadata，現在面臨整合時，系統方面也是很大的問題。

陳亞寧：建議 DC format 的編碼，有根據的標準，若有的話，可以直接參考。

陳淑君：建議聯合目錄小組先整合有哪些資料類型？再找編碼時比較好利用。

王梅玲：難求完美，只能求最大公約數。若兩者皆考慮的話，瀏覽部分數量很重要，需要內容組提供預數位化之數量與資料類型。

楊美華：要讓使用者在總目錄上，作有層次性的瀏覽或是利用查詢作簡易查詢或進階查詢。

(二) 控制詞彙部分相關討論。

黃銘崇：如欲搜尋「蟠龍花瓶」，沒打上這四個字就找不到。

何建明：這部分可能就是控制詞彙的問題。

王梅玲：今年度聯合目錄進度未定到控制詞彙這部分，因此不考慮。

(三) American Memory 的檢索後的紀錄的必備欄位，也是用 DC，就定出了：title、creator、subject、description、identifier、date、formate 等。這些也就是查詢點，請問後設小組以未來而言 metadata 的查詢點是什麼？聯合目錄查詢的介面會呈現哪些資料？(王梅玲)

陳亞寧：format 跟 DC 配的時候，他沒有固定，依每個計畫的需求。就像依



照 OAI，所查詢的欄位也沒有一定。

陳淑君：王老師的意思是根據 DC 內 15 個元素裡面，聯合目錄大致上需要哪些必備欄位？

王梅玲：對，那這部分跟 XML DTD 有什麼關係？

陳淑君：這部分跟 metadata XML DTD 沒有關係。我想可以從一個著眼點，從最大的範圍也可以詮釋，或者是把他分成兩層，第一層 SHOW 出 5 個必備的欄位，第二層 SHOW 出剩下 10 個欄位。這是一個方式，但是如何想看更詳細的資料，就可以從 METADATA 裡面看。

(四)我們發現 METADATA 有很多種類型，有描述型、行政型、管理型、結構型等，但沒有把它區分出來。我們目前聯合目錄主要出現描述型的，結構型的、行政型的、典藏型的可能不是聯合目錄主要要出現的欄位，請問 METADATA 小組怎麼把它區分出來？（王梅玲）

陳淑君：這部分我們可能需要聯合目錄持續來做，我們剛開始幫主題計畫做的時候，無法規定她們只能有什麼，像有些 METADATA 的國際標準，有的有管理型的、有的有描述型的、有的有結構型的、有的三者都有。如果我們大家有共識的話，聯合目錄要的就是描述型的，我們可以 Mapping 到 DC，DC 就是描述型的 Metadata。其實不一定就只有 Show 出描述型的，像 Right 就是管理型的，檢索上可以宣告到什麼程度。

王梅玲：我們就是不知道聯合目錄要選哪些去呈現？剛剛簡要紀錄部分應該要空白的。

陳淑君：我們 7/15-8/20 就是要做王老師所說的這部分，來做測試。

王梅玲：那部份可能需要後設小組幫忙，那我可以將 CDP 的資料 mail 給你，她們做的還不錯。

(五)請問技術組依照這樣的系統規格需求書可以設計平台了嗎？（王梅玲）

范紀文：這邊有幾項還需補充，可以使需求書更完整：

1. 訂定聯合目錄 DTD 出來，依照現在的 DC，則訂出 DC XML DTD 與 DC XML Schema 規範。
2. 加入如 namespace 的管理性資料進去，如資料是何人建立的，何時所建 等資料。
3. 典藏資料的著錄項目與聯合目錄資料轉入格式，兩者之間的對應關係。
4. 聯合目錄分類架構，建議與各領域專家共同討論制定。
5. 在 XML 表達文件中，需要注意的項目：
 - (1)Hyperlink 的問題：建議定義一致性表達方式，如以 URI 或 URL 來表達。



(2) 分類目錄的表示法，方便大量輸入資料時，系統可以讀出的格式。

(六) 檔案主題的分類只有內閣大庫與近史所外交檔案兩種，這兩種分類不同並放在一起？（何建明）

王梅玲：目前我們所找到的檔案分類都不一樣，內閣大庫與近史所外交檔案未來有無共同點可以拉出一個分類基準，根據洪一梅小姐(中研院史語所)與王麗蕉小姐(政大圖資所檔案組畢業)討論的結果，目前是可能會利用時間來作檔案分類的基準。

何建明：當我們在建立瀏覽時，不希望太少的瀏覽，也不希望太大的瀏覽，檔案大致上算了一下，有 70 個次類別，又太大了？這幾種分類依據紀文處理教育部那邊計畫，是將這些類別編成代碼的型式，存成資源識別代碼。

黃銘崇：下次是否需要將實際在執行計畫的助理找來一起討論？

陳淑君：像考古組、人類學，她們專家為了要分類別時，發現有派別。

王梅玲：因此我們希望在分類目錄上有大一統的現象，也就說在次分類上可以容下好幾種派別，不必指規定一個次分類架構。如人類學的分類不必唯一的。

(七) 目前如何作主題分類？時間部分目前只有一層，地理是兩層，但其他的分類都至少有三層？（黃銘崇）

王梅玲：時間分類目前是依據中國圖書分類法來列出，那現在該如何訂定？去年似乎已經討論過了。目前就是要將分類與編碼準備好，與研究人員討論是嗎？

黃銘崇：目前可以找實際在執行的助理一起討論。

楊美華：還需要討論分類是要用通俗性的還是學科性的？需要與主題小組分別討論，內容組那邊可以幫忙安排時間？

許雅婷：7/24 動植物主題；考古主題會議時間尚未確定，29 日有主題小組召集人暨開放性計畫主持人聯席會可以一起討論。

王梅玲：與各主題之討論大約需要一小時，排除新聞、地質、地圖三個主題。

楊美華：希望各主題個別開會。

許雅婷：將盡量安排。

四. 決議

(一) 聯合目錄系統之分類瀏覽功能，原訂五種瀏覽架構，再加上一種瀏覽架構—『數位化前資料類型』，共為六種如下：

1. 12 個內容與細部分類
2. 9 個典藏單位與 50 餘個典藏計畫



3. 時間
 4. 地點
 5. 數位化物件類型
 6. 數位化前資料類型
- (二) 由 Metadata 工作小組，依據 Dublin Core 15 個欄位中，列出在聯合目錄系統檢索結果之呈現之核心欄位與必備欄位。
- (三) 聯合目錄各瀏覽分類架構所包含之層次數目，由聯合目錄建置計畫與各主題小組共同討論後再予決定。
- (四) 『國家數位典藏聯合目錄系統需求書』需再補齊下列內容：
1. 訂定聯合目錄 DTD 出來，依照現在的 DC，則訂出 DC XML DTD 與 DC XML Schema 規範。
 2. 加入如 namespace 的管理性資料（如資料由何單位所建立、何時建立 等）。
 3. 典藏資料的著錄項目與聯合目錄資料轉入格式，兩者之間的對應關係。
 4. 聯合目錄分類架構，建議與各領域專家共同討論制定。
 5. 在 XML 表達文件中，需要注意的項目：
 - (1) Hyperlink 的問題：建議定義一致性表達方式，如以 URI 或 URL 來表達。
 - (2) 分類目錄的表示法，方便大量輸入資料時，系統可以讀出的格式。
- (五) 由內容發展分項計畫聯絡各主題小組(新聞、地質、地圖除外)，參與討論聯合目錄建置事宜：
1. 討論時間訂於 7 月 22 日中午及 7 月 29 日（上午 10 時至 12 時、或中午 12 時至下午 2 時）。
 2. 7/24 動植物主題小組討論會，可一併進行聯合目錄討論。其餘主題小組之討論時間由內容發展分項計畫代為聯絡。
- (六) 下次會議時間暫定為 8 月 25 日(一)下午 2 點，地點暫定中研院資訊所 101 會議室。

六.散會