

## 國立臺灣大學圖書館數位典藏 Metadata 之設計 與資料庫之建置彙整

洪淑芬

國立臺灣大學圖書館特藏組編審  
國立臺灣大學圖書資訊學研究所博士生  
Shu-Fen Hung Lin  
Executive Officer,  
Special Collections Department, National Taiwan University Library  
Doctorial Student,  
Graduate School of Library and Information Science, National Taiwan University  
E-mail : shufen@ntu.edu.tw

邱婉容

國立臺灣大學圖書館特藏組組長  
Wan-Jung Chiu  
Head,  
Special Collections Department,  
National Taiwan University Library  
E-mail : otter@ntu.edu.tw

### 摘 要

本論文主要是以國立臺灣大學圖書館典藏的《淡新檔案》、臺灣古拓本、臺灣古契書為例，論析同質異群資料整合的必要性與整合方法。本館進行典藏數位化時，考量推廣普及珍貴歷史資料之利用，故以建置便於檢索的資料庫系統為目的，為各資料群設計詮釋資料（Metadata）。然於建置資料庫系統時，則以使同質性之資料共用同一檢索介面、共用檢索結果之顯示格式、而且容許跨資料庫檢索，以便提高讀者使用上之便利性為目標。

本館就《淡新檔案》、臺灣古拓本、以及臺灣古契書等進行 Metadata 資料欄位規劃時，雖均以國際標準 Dublin Core 基本欄位為依據，然因各資料群各具有獨特性質，為充分描述資料，乃於 Dublin Core 基本欄位架構下，設置各資料群適用的欄位。本館透過 Metadata 欄位之規劃與後續的建檔規劃之配合，達成異群資料整合之最終目的。本文首先介紹《淡新檔案》、臺灣古拓本、以及臺灣古契書等的 Metadata 欄位之規劃、欄位訂定上所考量之因素，其次說明整合各資料群的 Metadata 之步驟、完整呈現各資料群獨特的 Metadata 之方法，最後剖析臺灣大學各資料群的 Metadata 及建檔規劃的優點，以及其如何有效配合聯合目錄建置之需要。希望本館 Metadata 設計與資料庫建置之經驗，可供各執行數位典藏計畫的機構以及建置數位典藏聯合目錄之參考。

## 關鍵詞 Keywords

詮釋資料、後設資料、資料庫整合、檢索介面、建檔介面、先行詞、前導語  
Metadata, Integration of Databases, Mapping of Metadata formats, Search Interface,  
Input Interface, Preceding Phrase

### 一、前言

國立臺灣大學圖書館（以下簡稱本館）開始進行館藏數位化的時間甚早；自民國 85 年「平埔族數位圖書館與博物館」的建置為始，歷年來積極參與行政院國家科學委員會「數位博物館計畫」、「數位典藏國家型科技計畫」以及行政院文化建設委員會「國家文化資料庫」之各項數位化計畫。本館於早期執行數位典藏計畫時，為建置數位化成果為可互動檢索的資料庫，並考量未來整合資料庫之需求，在詮釋資料（Metadata）建檔方面，乃選擇採用國際標準之一的都柏林核心欄位（Dublin Core）為基準，設計適用於各不同資料群之 Metadata 資料格式。（註 1）

因資料群的增加、以及未來整合之需要，本館在 Metadata 資料的建置規劃方面，乃一如都柏林核心欄位（Dublin Core）的改良更新而改版（註 2）一般，對於各資料群的 Metadata，亦在實施過程中，試圖作必要的調整，期使未來不論在支援檢索功能之開發或資料之完整呈現方面，都能適切地達成任務。

本館數位化典藏的 Metadata 資料格式之發展可分為以下二階段：

#### （一）「平埔族計畫」與「淡水河溯源計畫」時期

本館早年於「平埔族計畫」與「淡水河溯源計畫」期間，設計的 Metadata 資料格式計有《淡新檔案》、《岸裡大社文書》、《伊能嘉矩手稿》、《臺灣舊照片》等四種。各種資料群之欄位設計，均考量各種資料之特性，以能夠充分表現出資料內容與樣式，並對資料外觀作詳細描述為目標。（註 3）因此，欄位的設計，可說巨細靡遺。例如，地名的部分，依臺灣清朝與日治時期的地方行政區劃及實際地名分級，區分為「廳」、「縣」、「鎮」、「鄉」、「郡」、「街」、「堡」、「庄」等欄位。

#### （二）「國家典藏數位化計畫」（民國 90 年）、「數位典藏國家型科技計畫」（民國 91 年迄今）與「國家文化資料庫」數位化計畫（民國 90- 年）時期

本館自民國 90 年起執行「數位典藏國家型科技計畫」的先導計畫「國家典藏數位化計畫」，並於 91 年起執行為期五年的「數位典藏國家型科技計畫：臺灣文獻文物典藏數位化計畫」。該階段因為補助經費大幅提高，所需進行數位化的資料量龐大，因此，於開始大量、快速地進行數位化之前，乃就第一階段之 Metadata 建檔結果作一審視，作為適度調整各資料群的 Metadata 欄位之參考。該階段對於 Metadata 欄位之調整，在理念上可說是延續第一階段設計 Metadata 欄位時之考量，即希望對資料作詳實之描述，但亦兼顧執行上的可行性；該階段在 Metadata 格式方面的發展主要包括：

1. 將《淡新檔案》與《臺灣舊照片》之地名、人名欄位簡化，以利建檔實務之進行。

2. 開發《臺灣古拓碑》之 Metadata 欄位。(註4)
3. 透過各資料群 Metadata 欄位的對應工作，整合各資料群於一個資料庫系統中。

參與「數位典藏國家型科技計畫」之機構，如同本館一般，大多為機構本身的典藏品，由機構獨立開發、或透過有同質性典藏品的機構間之合作，針對某一類資料開發 Metadata 資料格式。自民國 90 年至今，各「數位典藏國家型科技計畫」的執行機構，各已建置相當數量之 Metadata 資料。現在「數位典藏國家型科技計畫」正規劃聯合目錄之建置並測試中；本館的「臺灣文獻文物典藏數位化計畫」在機構內的整合、或配合聯合目錄之建置而提供必要的 Metadata 欄位資料，均屬順利，因此，在此簡介本館對於 Metadata 欄位訂定與調整規劃上所考量之因素、整合各資料群 Metadata 的過程步驟、方法技巧等，以及完整呈現各資料群特有的 Metadata 欄位資料之方法、轉出 XML 格式的 Metadata 以匯入聯合目錄系統之方法、並說明臺灣大學各資料群的 Metadata 格式及建檔規劃的優點等，希望有助於各執行典藏數位化的機構間同質性之 Metadata 資料整合以及聯合目錄之建置。

## 二、Metadata 資料欄位訂定與調整規劃上所考量之因素

本館於數位化典藏的 Metadata 格式發展的第二階段期間，就未來資料庫檢索、呈現等作一具體規劃，並就 Metadata 建置之整體情形作一檢視，確認所建置之資料是否足以符合規劃上之需求。對於未來資料庫檢索的功能，預先規劃如下，再基於達成該規劃所設定的目標，進行建檔規範之擬定：

### (一) 資料庫須足以提高資料文物的「可及性」(accessibility)

希望透過有效的查詢系統，提高資料文物的「可及性」(accessibility)。資料文物的良好可及性，主要是藉適度的回收率 (recall) 與相當的精確率 (precision) 方得以達成。茲藉以下「資料庫檢索」之介面 ([表一])，針對檢索上的需求與檢索程式運作之邏輯設計，說明檢索介面與未來檢索程式擷取資料之規劃如下：(註5)

〔表一〕資料庫檢索介面與檢索程式運作之設計

資料庫檢索	
檢索詞彙參考清單	<p><u>人名</u> <u>地名</u> <u>主題</u> <u>關鍵詞</u></p> <p>進入清單中點選，所點選之詞彙將自動填入本檢索介面的欄位內。</p> <ul style="list-style-type: none"> <li>● 同一欄位點選多個詞彙，系統自動設定為 and 的關係。</li> <li>● 不同欄位內鍵有資料時，系統自動設定為跨欄位的交集檢索。</li> </ul>
題名	(完整題名或題名中的關鍵詞)
主題	(資料內容之主題)
關鍵字	(資料內容相關的詞彙，鍵入多個詞彙時，以 and (*) /or (+) 連結詞彙，系統將進行交集/聯集檢索)
作者	(資料之作者)
人名	(資料內出現之人物，鍵入多個人名時，以 and (*) /or (+) 連結，系統將進行交集/聯集檢索)
地名	(資料內出現之地名，鍵入多個地名時，以 and (*) /or (+) 連結，系統將進行交集/聯集檢索)
日期	作成年代或出版時間(西元年；可以讓查詢者鍵入特定的一年，或提供一定的年代起迄之選單，例：1701-1750/1751-1800/1801-1850/1851-1900……)
作成朝代	(例：宋/元/明/清/民國)
媒體類型	<input type="radio"/> 紙本原件 <input type="radio"/> 原件複製品 <input type="radio"/> 數位檔 <input type="radio"/> 微捲(片) <input type="radio"/> 底片 <input type="radio"/> 影音磁帶 <input type="radio"/> 圖書 <input type="radio"/> 期刊 <input type="radio"/> 網路資源
語言	<input type="radio"/> 中文 <input type="radio"/> 英文 <input type="radio"/> 日文 <input type="radio"/> 滿文 <input type="radio"/> 其他
綜合檢索	(鍵入任何詞彙，鍵入多個詞彙時，以 and (*) /or (+) 連結各詞彙，系統將搜尋題名、主題、關鍵字、摘要、人名、地名、作成朝代、來源、全集、子項內容(目次)等欄位，進行交集/聯集檢索)
依	<input type="radio"/> 時間 <input type="radio"/> 題名 <input type="radio"/> 人名 <input type="radio"/> 地名 <input type="radio"/> 遞增 <input type="radio"/> 遞減      排序

## (二) 為提高查詢之回收率 (recall) 的設計

〔表一〕表格的左欄表示提供使用者輸入檢索詞彙之欄位，表格的右欄所列說明為檢索程式撰寫之指示。為了幫助使用者提高查詢之回收率 (recall)，規劃提供「綜合檢索」的欄位；使用該欄位檢索，系統會自動擷取題名、主題、關鍵字、摘要、人名、地名、作成朝代、來源、全集、子項內容等欄位之資料，因此上列欄位均列為 Metadata 的欄位，而且規範需儘量鍵入資料。

## (三) 為提高查詢之精確率 (precision) 的設計

如果使用者由「綜合檢索」欄位開始進行檢索，而得到相當多筆資料，即檢索結果不夠精確時，希能藉由再於題名或主題、關鍵詞欄位鍵入同一詞彙或其他主要想檢索的詞彙，進行交集檢索，使檢索結果趨於精確。基於此一原理，規劃系統之檢索功能時，要求將不同欄位內之資料，設定為跨欄位的交集檢索。也因為此一檢索設計，可幫助提高檢索精確率的主題、關鍵字等欄位，亦規範應儘量正確地鍵入資料。換言之，檢索上需要使用到的欄位，包括交集檢索、聯集檢索所利用到的欄位，都列入 Metadata 之中，而且確立建檔規範。

#### (四) 資料庫須足以達成排序目的

資料庫於使用者進行查詢時，提供使用者將檢索結果加以排序之選擇；排序之依據項目包括年代時間、題名首字筆劃數、人名首字筆劃數、地名首字筆劃數等，使用者並可選擇依遞增或遞減方式排序。

依筆劃排列者，建檔格式可不作限制。但是，年代日期的排序，則有賴於著錄 Metadata 的年代日期時，以一定的格式進行著錄。

一般的古文件資料以及古刊本圖書內可能以中國紀年註記日期，該日期一般情況可視為等同於「作成日期」而將之著錄於「出版日期」或「作成日期」欄位中。為了依日期排序之需，必需將中國紀元年換算為西元年，依一定格式著錄，例如全部統一為 yyyy-mm-dd 或 yyyy/mm/dd 之樣式，而且西元年需另成立為一個獨立欄位。如此，才能透過程式的運作，將檢索結果依年代順序排列。

#### (五) 資料庫需提供人名、地名、主題或關鍵詞等清單

本館於「數位典藏國家型科技計畫：臺灣文獻文物典藏數位化計畫」中所執行的典藏數位化計畫之資料群，包含古文書、拓碑、手稿等類型之資料，該等資料內容較為艱深，一般使用者於查詢之際，往往有不知以何詞彙進行檢索之困惑，而無從下手。於資料庫的檢索介面中，提供「參考清單」，對查詢者幫助極大。所謂「參考清單」包括資料庫中（即所有資料的 Metadata）所有的人名、地名、主題或關鍵詞等清單。設計上可以讓查詢者於瀏覽清單時，點選適用之詞彙，由系統自動將查詢者所點選的辭彙鍵入檢索欄位中；俟查詢者點選足夠的辭彙，例如點選地名、主題之後，又選定年代起迄，查詢者再點選開始查詢。如此，查詢者即使對某一資料群或歷史、地名、人名等並不熟悉，或一時無法想起適用之詞彙，亦可於檢索時，得到即時的提醒而順利進行檢索。

以《淡新檔案》為例，該檔案於國立臺灣大學自台北帝國大學接收後，由戴炎輝教授主持整理計畫，將該檔案區分為行政、民事、刑事等三「編」，各「編」之下再分「類」，各「類」之下再分「款」（參見〔附錄一〕）。戴炎輝教授所編製的分類款目表，一直為學者或師生研究教學上利用《淡新檔案》時找尋檔案的憑藉。本館於第二階段規劃《淡新檔案》之 Metadata 欄位時，考量戴炎輝教授以其專業與可觀的人力時間為該檔案整理歸類之功，以及學者依戴炎輝教授之分類款目利用該檔案之習慣，乃將戴教授之「編」、「類」、「款目」等均列入 Metadata 的特定欄位。如此，在資料庫的檢索介面中，即可將「編、類」（在 Metadata 中列為「主題」欄位的內容）、「款目」（在 Metadata 欄位中列為「關鍵詞」欄位），作為提供查詢者參考用的清單，而使查詢者利用現代科技所建置完成的資料庫時，仍然可以沿用原來利用《淡新檔案》之習慣，順利查得所需之文件，並連結影像與全文。

#### (六) 為檢索結果「簡目」、「詳目」呈現之需的設計

資料庫的檢索主要是依據題名、主題、關鍵字、人名、地名等欄位內的資料進行。但是，檢索結果之呈現，可增列更多有利於資料內容辨識用的欄位如「文書形式類別」、「作成朝代」、「出版者／出版地」，以及稽核項目、資源類型、媒體類型等，讓使用者更具體瞭解所查得的資料之性質。下列〔表二〕右欄之欄位，為檢索結果「簡目」可以考慮選用的欄位（左欄是為方便了解所作的簡明歸類）：

〔表二〕古文書檢索結果簡目之例

系統管理	系統識別號（系統自動給與）
館藏管理／ OAI 需要使用到的 DOI	檔案名稱 為該筆 Metadata 的檔名（有如 MARC 書目記錄中的 .b 號碼。但該檔名最好反映資料的所在，因此，可以是 Call Number 或 Accession Number 之類具獨一性的號碼（註6）。基於跨機構整合的考量，需加上「機構代碼」（或再於「機構代碼」後加上「檔案類型代碼」），可用作 Digital Objective Identifier (DOI)。
內容相關重要訊息 （另加列出版資訊）	主要題名
	其他題名
	文書形式類別
	作者
	內容主題
	關鍵詞
	重要關係人
	相關地名
	日期（作成或出版日期）
	作成朝代
	出版者／出版地
稽核基本資訊 （Physical description）	數量／單位
	規格（例如：尺寸大小）
資源類型	原件與否
媒體類型	媒體類型
提供連結	影像檔名
	全文檔名
	影音視訊檔名

Dublin Core 的十五個基本欄位，各可依文物資料之性質，複分出相關欄位。以古文書資料為例，為古文書所設計的 Metadata 欄位格式，可多至六十個欄位左右。如果檢索結果提供「詳目」顯示，可以將所有建有資料的欄位均加呈現，或僅選擇性的較「簡目」列出較多欄位。

資料庫的檢索功能愈細緻，對使用者檢索之效益愈有幫助。另一方面，檢索結果之呈現列載愈多資訊，對使用者辨識資料也愈有幫助。但是，不論是資料庫的檢索功能或檢索結果的資料呈現，均取決於 Metadata 的欄位以及鍵入欄位內的資料。因此，本館規劃 Metadata 欄位時，事先就資料庫未來檢索功能、資料呈現等兩方面的需要進行了解，使規劃不至於遺漏必要的欄位。

### 三、整合各資料群Metadata的過程步驟、方法技巧

本館在 Metadata 的第二發展階段所進行的工作之一是將各資料群之欄位進行對應，目的是為了整合各資料群於一個資料庫系統中。整合各資料群 Metadata 的過程如下：

#### (一) 評估各資料群的性質、整合的可行性、對使用者檢索上的便利性等

經研析各資料群的性質，基於下列的結論而決定嘗試將各資料群加以彙整：

1. 《淡新檔案》為行政與司法類的檔案，《岸裡大社文書》包含土地契約書之外，亦包含司法類的案件抄錄，而《臺灣古拓碑》中包含官方頒佈的行政類告示等，三者均包含彼此性質相似之文件或碑文，均屬「古文書」類資料。
2. 《臺灣舊照片》與《伊能嘉矩手稿》性質雖然異於「古文書」，但是，除了包含書刊以及書刊的出版者等欄位之外，其他檢索上所需的基本欄位如題名、人名、地名、主題或關鍵詞等欄位，與「古文書」類資料並無大差異。
3. 《淡新檔案》、《岸裡大社文書》、《臺灣古拓碑》、《臺灣舊照片》、《伊能嘉矩手稿》等同為本館的重點臺灣研究相關資料，如能提供讀者「跨資料庫檢索」以及「選擇資料庫檢索」的不同選擇，可方便讀者依需要而決定，而提高讀者使用上的便利性。

#### (二) 選擇中文版的Dublin Core的基本架構表

雖然《淡新檔案》、《岸裡大社文書》、《臺灣古拓碑》、《臺灣舊照片》、《伊能嘉矩手稿》等五個資料群之 Metadata 均是以 Dublin Core 之基本架構為依據，但各開發出許多與各資料群性質相關的特殊欄位。在進行彙整之前，乃挑選當時國內較具共通性的「Metadata 欄位總表」，以便將各資料群的所有欄位，對應於 Metadata 欄位總表之欄位。當時選用的「Metadata 欄位總表」是「行政院國家文化建設委員會國家文化資料庫詮釋資料格式」1.3 版。(註7) 將各資料群的欄位對應至「Metadata 欄位總表」的步驟如下：

1. 確認「Metadata 欄位總表」各欄位之意義
2. 將「Metadata 欄位總表」之各欄位加以編號

Dublin Core 的基本 15 欄位編為 1-15 (下列 [表三] 左欄)，各基本欄位的次層複分欄各自編為 A、B、C、D……。(下列 [表三] 中欄)。再其次的複分欄編為 [A-1]，[A-2]，……[C-1]，[C-2]……，餘類推。下列 [表三] 為其中的一部分，該編號表請參見本論文 [附錄二]。

〔表三〕國立臺灣大學圖書館數位典藏「Metadata 欄位總表」各欄位之編號（例）

<small>identifier</small> 識別資料 10	<small>localNumber</small> 地方識別號碼 C	<small>localFileName</small> 檔案名稱 C-1 拓 岸 淡 伊 · 典藏排架號 照 · 收藏者(單位)排架號	Metadata 的檔名，為代表資料物件的唯一編號。	必備
		<small>systemID</small> 系統識別號 C-2 拓 岸 淡	系統自動產生。	必備
		<small>location</small> 館藏地 C-3 拓	目前原件資料所收藏之位置。	

**（三）將各資料群之所有Metadata欄位對應至「Metadata欄位總表」**

根據各資料群之各個欄位內容值之意義，於「Metadata 欄位總表」之各個欄位中，填入有使用到各該欄位之資料群簡稱。各資料群之 Metadata 資料庫所使用的欄位名稱相同於「Metadata 欄位總表」之欄位名稱者，只填寫資料群之簡稱，例如上〔表三〕中在「檔案名稱」欄位內填入「拓」、「岸」、「淡」，表示《臺灣古拓碑》、《岸裡大社文書》、《淡新檔案》等三資料庫之該欄位名稱均相同，為「檔案名稱」。但是，另填入「伊：典藏排架號」、「照：收藏者（單位）排架號」，表示在《伊能嘉矩手稿》資料庫中，相當於該欄位之欄位名稱為「典藏排架號」，在《臺灣舊照片》資料庫中，相當於該欄位之欄位名稱為「收藏者（單位）排架號」。

**（四）為各資料群之Metadata資料庫欄位標籤標識相當於「Metadata欄位總表」各欄位之編號**

本館各資料群之 Metadata 目前係以 Excel 軟體進行建置。為了方便未來彙整資料庫時，正確順利地選取各資料庫的相同欄位，乃於各資料庫中，為各資料群之 Metadata 資料庫欄位標籤標識相當於「Metadata 欄位總表」各欄位之編號，如下列〔表四〕：



[表四]《淡新檔案》資料庫欄位相應於「Metadata 欄位總表」各欄位之編號(相應編號見第3列)

	A	B	C	D	E	F	G
1	淡新 Metadata 資料序號	檔案名稱	主要題名	文書形式類別	主要收文者及職稱/身份	主要收文者-團體	主要關係人及職稱/身分
2	1	2	3	4	5	6	7
3		10/C/C-1	3/A	5/W	5/C/C-1 + C-2	5/C/C-1 + C-2	5/C/C-1 + C-2
4		identifier/localNumber/localFileName	title / mainTitle	description / document Categories	description / relativePerson / relativePersonName	description / relativePerson / relativePersonName	description / relativePerson / relativePersonName

	A	B	H	I	J	K	L	M	N	O	P
1	淡新 Metadata 資料序號	檔案名稱	主要關係人-文書角色	主要關係人-姓名	主要關係人-社名	原件作成地點	主要作者/發文者及職稱/身分	主要作者/發文者-團體	內容摘要	數量單位	圖像
2	1	2	8	9	10	11	12	13	14	15	16
3		10/C/C-1	5/C/C-1 + C-2	5/C/C-1 + C-2	5/C/C-1 + C-2	14/A/A-2	6/A	6/A	5/A	2/B/B-1	5/R/C
4		identifier/localNumber/localFileName	description / relativePerson / relativePersonName	description / relativePerson / relativePersonName	description / relativePerson / relativePersonName	coverage / spatial / createdPlace	creator / creatorNameTitle	creator / creatorNameTitle	description / abstract	format / extent / quantity	description / illustration

	A	B	Q	R	S	T	U	V	W	X	Y
1	淡新 Metadata 資料序號	檔案名稱	尺寸大小	附件	印記種類-章	印記種類-印：官印	印記種類-指掌印	備註	內容主題	關鍵字1	關鍵字2
2	1	2	17	18	19	20	21	22	23	24	25
3		10/C/C-1	2/B/B-2	5/R/D	5/K	5/K	5/L	5/X	4/A	4/B	4/B
4		identifier/localNumber/localFileName	format / extent / dimension	description / attachment	description / seal	description / seal	description / fingerPalm Print	description / notes	subject / subjectMatter	subject / keywords	subject / keywords

	A	B	Z	AA	AB	AC	AD	AE	AF	AG
1	淡新 Metadata 資料序號	檔案名稱	成文日期	入藏日期	編目日期	排架號	館藏代碼	來源	全集/系列名稱	相關族群
2	1	2	26	27	28	29	30	31	32	33
3		10/C/C-1	9/A	9/D	16/A	10/C/C-5	10/C/C-4	11	12/A	12/E
4		identifier/localNumber/localFileName	date / createdDate	date / acquiredDate	catalogingDate	identifier / localNumber / callNumber	identifier / localNumber / accessNumber	source	relation / isPartOf	relation / seeAlso

	A	B	AH	AI	AJ	AK	AL	AM	AN	AO	AP
1	淡新 Metadata 資料序號	檔案名稱	相關參照註	編目語文	作品語文	內容地點	時間類別	典藏單位	典藏單位國家	收藏取得方式	編目審核員
2	1	2	34	35	36	37	38	39	40	41	42
3		10/C/C-1	12/E	16/F	13	14/A/A-1	14/B	15/B/B-2	15/B/B-1	5/I	16/G
4		identifier/localNumber/localFileName	relation / seeAlso		language	coverage / spatial / contentPlace	coverage / temporal	rights / owner / ownerName	rights / owner / ownerCountry	description / acquireMethod	

A		B	AQ	AR	AS	AT	AU	AV	AW	AX	AY
1	淡新 Metadata 資料序號	檔案名稱	作品類型	媒體類型	原件與否	保存狀況	裝訂	裝讀	藏品層次	色彩	館藏地
2	1	2	43	44	45	46	47	48	49	50	51
3	10/C/C-1		1/C	2/A	1/B	5/H	5/O	5/S	1/A	2/B/B-3	10/C/C-3
4	identifier/localNumber/localFileName		type / worksType	format / medium	type / originalSurrogate	description / conditions	description / binding	description / mount	type / aggregationLevel	format / extent / color	identifier / localNumber / location
A		B	AZ	BA	BB	BC	BD	BE	BF	BG	BH
1	淡新 Metadata 資料序號	檔案名稱	其他格式	其他格式	其他格式	一致性識別名稱 URN	系統識別號	數位化類別	檔案類型	檔案描述	使用權限
2	1	2	52	53	54	55	56	57	58	59	60
3	10/C/C-1		12/D	12/D	12/D	10/B/B-1	10/C/C-2	2/C/C-1/C-1-(1)	2/C/C-1/C-1-(2)	2/C/C-1/C-1-(4)	2/C/C-1/C-1-(5)
4	identifier/localNumber/localFileName		relation / hasFormat	relation / hasFormat	relation / hasFormat	identifier / URI / URN	identifier / localNumber / systemID	format / scale / digitalScale / digitalCate	format / scale / digitalScale / fileType	format / scale / digitalScale / fileDesc	format / scale / digitalScale / permission
A		B	BI	BJ	BK	BL	BM	BN	BO	BP	BQ
1	淡新 Metadata 資料序號	檔案名稱	檔案大小	格式與版本	文件型別定義	解析度	每點位元素	圖片大小	壓縮比	評析	評等
2	1	2	61	62	63	64	65	66	67	68	69
3	10/C/C-1		2/C/C-1/C-1-(6)	2/C/C-1/C-1-(7)	2/C/C-1/C-1-(8)	2/C/C-1/C-1-(9)	2/C/C-1/C-1-(10)	2/C/C-1/C-1-(11)	2/C/C-1/C-1-(12)	5/D	5/F
4	identifier/localNumber/localFileName		format / scale / digitalScale / fileSize	format / scale / digitalScale / textFormatVersion	format / scale / digitalScale / textAssociatedOTD	format / scale / digitalScale / resolution	format / scale / digitalScale / bitPerPixel	format / scale / digitalScale / imageSize	format / scale / digitalScale / compressionRatio	description / comments	description / grade
A		B	BR	BS	BT	BU	BV	BW	BX	BZ	
1	淡新 Metadata 資料序號	檔案名稱	朝代	URL 一致性資源位置	著作財產權人	著作權授權狀態	使用限制	XML產生日期	XML產生者	影像檔 - i	
2	1	2	70	71	72	73	74	75	76	78	
3	10/C/C-1		6/B	10/B/B-2	15/A/A-1	15/A/A-2	15/A/A-3	16/B	16/D	2/C/C-1/C-1-(3)	
4	identifier/localNumber/localFileName		creator / creatorDynasty	identifier / URI / URL	rights / copyrightRestrictions / holderName	rights / copyrightRestrictions / rightsState	rights / copyrightRestrictions / accessRestrictions	XMLgenerateDate		format / scale / digitalScale	
A		B	CA		CB						
1	淡新 Metadata 資料序號	檔案名稱	影像檔 - t		全文檔名 (程式抓)						
2	1	2	79		85						
3	10/C/C-1		2/C/C-1/C-1-(3)		2/C/C-1/C-1-(3)						
4	identifier/localNumber/localFileName		format / scale / digitalScale / fileName								

#### 四、完整呈現各資料群特有的Metadata欄位資料之方法—利用「前導語」之機制

如前所述，《淡新檔案》、《岸裡大社文書》、《臺灣古拓碑》、《臺灣舊照片》、《伊能嘉矩手稿》等各資料群之 Metadata 各自具有與各資料群性質相關的特殊欄位。但是，未來各資料庫將使用統一的「資料庫檢索介面」(〔表一〕)，以及「檢索結果」之格式(〔表二〕)——該介面與簡目格式以基本欄位為主，呈現上較為簡明扼要。爲了在共用的、基本的欄位中，區隔複分的欄位，例如，「題名」欄位在建檔時建有「主要題名」(例如《淡新檔案》、《岸裡大社文書》、《臺灣古拓碑》等)以及「其他題名」、「拼音題名」、「額題」、「尾題」等(例如《臺灣古拓碑》)；「材質」方面一般是指資料原件的材質，但是，拓碑的情形則含「原碑」、「拓本」等二種物件的不同材質。此處所稱的「複分的欄位」，即「其他題名」、「拼音題名」、「額題」、「尾題」、「原碑(材質)」、「拓本(材質)」等針對特別資料群特有資料性質而設定的欄位。建檔時，對於此類特殊需要註明複分欄位性質的，即將該等複分的欄位名稱，以「前導語」(或稱「先行詞」)的格式，同時鍵入欄位之內。如此，未來檢索結果呈現 Metadata 時，欄位值即可同時呈現該特別的複分欄位之名稱，而使欄位值的意義明確化。

爲了明確釐清各資料群所使用到的「複分的欄位」(需要於建檔時加入「前導語」的欄位)，於前述〔三之(三)〕將各資料群之所有 Metadata 欄位對應至「Metadata 欄位總表」時，即詳細將各資料群的特殊欄位所需使用的「前導語」註記清楚；以下〔表五〕是自「Metadata 欄位總表」所摘錄的部分含有「前導語」的欄位：

〔表五〕需加「前導語」的欄位之例（圈記之部分）

欄位名稱		欄位說明	必備/建議 著錄欄位
title 作品名稱 3	mainTitle 主要題名 (A) 拓 岸 淡 伊 題名 照 題名 alternativeTitle* 其他名稱 (B) 拓 拼音題名 額題 首題 陰題	該古文書的主要名稱，為正式名稱、或多數人所熟知的名稱、作者所命名的名稱等。 如：宋學士文粹 副題名、封面題名、劃一題名、全集題名、翻譯名稱、其他常用或少用的名稱...等，著錄時需加入前導語。	必備
coverage 時空涵蓋範圍 14	spatial 空間類別 A contentPlace 內容地點 A-1 岸 淡 內容涵蓋地點 伊 地名 照 作成地點-省 拍攝/繪製地點	該古文書內容涵蓋的地點。	
	createdPlace 原件作成地點 A-2 淡 作成地點 issuedPlace 出版/發行地點 A-3 recordedPlace 錄製/製作地點 A-4 拓 原物件原址 A-5 拓 現碑地點 原物件現址 A-6 拓 原碑地點	該古文書成文之地點。 該古文書出版/發行之地點。 該古文書資料錄製或複製之地點。	
	temporal 時間類別 (A) 岸 淡 本件涵蓋時間 照 事件時間-圖像/照片內容時間	該古文書的時間涵蓋範圍；此時間涵蓋範圍是指該古文書的相關時間，而非該古文書產生或成文時間（後者應著錄於「日期時間」欄位），著錄時通常採用時期名稱（如：日據時期），或是與「日期時間」欄位建議之相同的格式著錄。	
	material 材質 (C) 拓 原碑 拓本	該古文書的材質。	

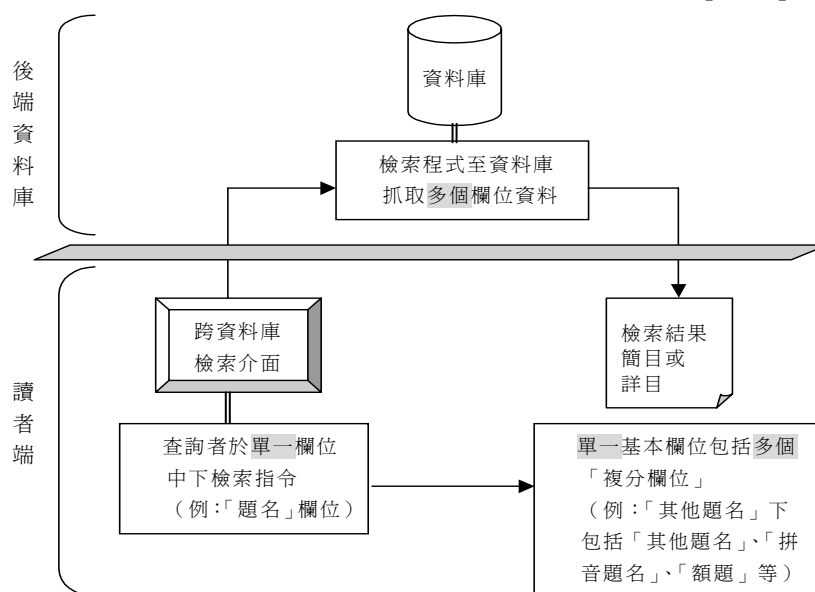
完整地將各資料群之所有 Metadata 欄位對應至「Metadata 欄位總表」，可幫助了解各資料群之間欄位內容值意義相同卻使用不同欄位名稱的情形，有助於資料庫欄位名稱之對應。

## 五、整合資料庫檢索程式之資料擷取與呈現

在規劃檢索程式資料擷取與讀者端檢索結果呈現時所考量的基本原則如下：

1. 共用相同的檢索介面。
2. 讀者端檢索介面之欄位採用讀者熟悉、常用的欄位名稱。
3. 讀者端檢索介面之欄位數以精簡為原則。(目前本館所完成的檢索介面如 [附錄三])
4. 共用相同的書目呈現介面。(目前本館所完成的檢索結果書目呈現格式如 [附錄四]，分「簡目」與「詳目」)
5. 檢索結果的書目資料呈現保留最完整的 Metadata 資訊，但必須使欄位內容值之意義明確化。

基於上述之考量，對於檢索程式之運作，規劃如下列 [圖一]：



圖一 資料庫檢索程式之資料擷取與呈現

基於上述之理念，以及提供具體確實之資料以供資料庫程式撰寫之需，在完成「Metadata 欄位總表」各欄位之編號 ([表三]) 以及為各資料群之資料庫欄位進行相應於「Metadata 欄位總表」各欄位之編號 ([表四]) 後，再進行各資料群相對應欄位之表格製作，該種表格包括二部分：一個表格是指示檢索程式於查詢者下查詢指令之後，至資料庫所要抓取的欄位 ([表六])，一個表格是指示檢索程式在讀者端的各特定欄位所要呈現的資料庫對應欄位 ([表七])：(完整的 [表六] 與 [表七] 請參見 [附錄五])

〔表六〕拓碑、岸裡、淡新、伊能、照片\_資料庫 INPUT 欄位對應表

(指示檢索程式於查詢者下查詢指令之後，至資料庫所要抓取的欄位)

資料庫 對應欄位 INPUT 欄位	拓碑	岸裡	淡新	伊能	照片
題名	3/A 3/B	3/A	3/A	3/A	3/A
主題	4/A	4/A	4/A		4/A(主&關)
關鍵字	4/B	4/B	4/B		4/A
作者	6/A 7	6/A ↕	6/A 7	6/A ↕	6/A 7
人名	5/C/C-1 6/A 7	5/C/C-1 6/A ↕	5/C/C-1 6/A 7	5/C/C-1 6/A ↕	5/C/C-1 6/A 7
地名	14/A/A-4 14/A/A-5 14/A/A-6	14/A/A-1 ↕	14/A/A-1 14/A/A-2 ↕	14/A/A-1 ↕	14/A/A-1 ↕

〔表七〕拓碑、岸裡、淡新、伊能、照片\_資料庫 OUTPUT 欄位對應表

(指示檢索程式在讀者端的各特定欄位所要呈現的資料庫內的欄位)

資料庫 對應欄位 OUTPUT 欄位	拓碑	岸裡	淡新	伊能	照片
檔案名稱	10/C/C-1	10/C/C-1	10/C/C-1	10/C/C-1	10/C/C-1
主要題名	3/A	3/A	3/A	3/A	3/A
其他題名	3/B	↕	↕	↕	↕
文書形式類別	5/W	5/W	5/W	5/W	5/W
作者	6/A 7	6/A	6/A 7	6/A	6/A 7
內容主題	4/A	4/A	4/A	↕	4/A
關鍵詞	4/B	4/B	4/B	↕	4/B
重要關係人	5/C/C-1	5/C/C-1	5/C/C-1	5/C/C-1	5/C/C-1

## 六、轉出XML格式的Metadata以匯入聯合目錄系統之方法

數位典藏國家型科技計畫後設資料工作組就各參與「數位典藏國家型科技計畫」的機構所建置之 Metadata 進行調查，而於民國 92 年 4 月將調查之結果製作為「檔案後設資料核心元素清單與調查結果」(為「檔案主題小組『後設資料核心元素』彙整結果」的 [附件一])，以利「數位典藏國家型科技計畫聯合目錄」

之建置。本館配合該聯合目錄之測試需要而將已完成的《淡新檔案》之部分 Metadata 轉檔為 XML 格式，提供匯入該聯合目錄，轉檔過程之步驟如下：

1. 詳細了解「數位典藏國家型科技計畫檔案主題小組『後設資料核心元素』」之欄位定義。
2. 於前述「國立臺灣大學圖書館數位典藏『Metadata 欄位總表』」中，找尋與「數位典藏國家型科技計畫檔案主題小組『後設資料核心元素』」相當之欄位，並將本館對於各 Metadata 欄位之編號（參見 [表三]），記錄於檔案主題小組訂定的「後設資料核心元素」之旁。
3. 決定要轉出至聯合目錄的欄位（為下列 [表八] 加網底之欄位）。
4. 將原有的《淡新檔案》資料庫另存，將要轉出至聯合目錄的欄位名稱改為檔案主題小組訂定的「後設資料核心元素」之欄位名稱。換言之，欄位名稱之變更，只在另存的檔案作變更；本館持續使用與建檔的檔案，仍然保持本館原訂定之欄位名稱。
5. 將已改好欄位名稱的資料庫與數位典藏國家型科技計畫後設資料工作組所訂定的 DTD 二者一併交由本館系統資訊組工作同仁進行轉檔為 XML 格式之檔案，繳交予數位典藏國家型科技計畫的聯合目錄工作同仁。

[表八] 國立臺灣大學圖書館《淡新檔案》欄位與「檔案主題小組『後設資料核心元素』」對應表（該表格請參見 [附錄六]）

說明：

1. 該表格係摘錄自「檔案主題小組『後設資料核心元素』彙整結果：附件一：檔案後設資料核心元素清單與調查結果」。(註8)
2. 元素名稱以網底標示者，表該元素為「必備元素」，以粗黑字體標示者表臺灣大學圖書館 2004 年匯出之欄位。

核心元素		英文名稱	定義
中文名稱			
15/B/B-2	<b>1. 典藏單位</b>	Repository	負責典藏資料的機關單位
10/C/C-1	<b>2. 館藏號</b>	Collection number	辨識資料的唯一號碼，或所屬層次的唯一號碼。
5/W	<b>3. 類型</b>	TYPE	資料的形式或類別
1/A	<b>4. 描述層次</b>	Level of Description	該筆資料的層次
3/A	<b>5. 題名</b>	Title	資料的題名或資料分類層級之分類名稱

臺灣大學圖書館對於各 Metadata 欄位之編號

## 七、臺灣大學圖書館各資料群的Metadata格式及建檔規劃的優點與問題

本館對於 Metadata 之建檔，不僅事前為檢索上的需要有所考量，也確立建檔規範，並舉行建檔的相關訓練講習；以確實建檔、俾使資料庫能如規劃所預期的理想而運作為目標。以下為對於 Metadata 欄位設計與執行建檔規劃所達成的具體實效：

### (一) 優點

#### 1. 對欄位內容值的定義明確

以「國立臺灣大學圖書館數位典藏『Metadata 欄位總表』」([表三])為基準，確立各 Metadata 欄位內容值的定義，進而比對已建置的各資料群之 Metadata 欄位的內容值，將各資料群相同的欄位加以對應之外，對於內容值略有出入者，亦得以逐步調整，使趨於一致，而利於共用同一資料庫檢索程式。

#### 2. 建檔格式具有一致性

基於檢索需求之考量，例如排序之需要，以及當初預計將提供主題詞、關鍵詞、人名與地名等之瀏覽清單的規劃，在建檔規範中，對於相關欄位，特別規範必須以一定的格式建檔。因此，除了有利於未來的運用，例如，日期欄位之利於排序之外，主題詞、關鍵詞、人名與地名等欄位亦隨即可運用以建立瀏覽清單。

#### 3. 欄位設計完整充足

為了提供使用者更多的資料找尋途徑，除了提供前述的欄位檢索方式之外，亦可提供階層式瀏覽查詢的方式。本館因為對於欄位的設計完整充足，因此，建置階層式之瀏覽查詢方式時，有適切的欄位，足以提供建置由「廣」(broad)而「專」(specific)的資料組織層次。例如，《淡新檔案》的階層可以由 Metadata 中擷取以下欄位：

- 第一層：「集叢名」(淡新檔案)
- 第二層：「主題」(相當於戴炎輝教授擬定的「編—類」)
- 第三層：「關鍵詞」(採戴炎輝教授擬定的「款目」)
- 第四層：「集叢號」(為《淡新檔案》的「案號」，為該檔案的第一層編號)

此外，《臺灣古拓碑》的階層式顯示可用的欄位如下：

- 第一層：「集叢名」(臺灣古拓碑)
- 第二層：「主題」(仿戴炎輝教授擬定的「編、類」而建置)
- 第三層：「關鍵詞」(仿戴炎輝教授擬定的「款目」而建置)
- 第四層：「北、中、南部」或「年代」

一般古契書的階層式顯示可用的欄位如下：

- 第一層：「集叢名」(資料群名稱)
- 第二層：「主題」(仿戴炎輝教授擬定的編類款目方式而建置的「編—類」)
- 第三層：「文書型式類別」(例如「地契」、「鬮書」、「賣子契」等)
- 第四層：「年代」或「集叢號」(如「簿冊編號」)



目前本館的《淡新檔案》已由國立臺灣大學數位典藏計畫第七子計畫以及數位典藏國家型科技計畫之「數位典藏聯合目錄」等，順利利用現有之欄位，分別建置階層式瀏覽檢索之使用介面（請參見 [附錄七]）。

## （二）問題：人名與地名的權威控制

本館目前進行數位化的各資料群，其中所出現的人名、地名、官銜職稱等，不易利用現有的相關權威檔進行權威控制，原因如下：

1. 文件內出現的一般事件關係人多為平民百姓，且人名有的怪異奇特，並未為現有的權威檔所收錄。
2. 人名有的是親友或鄰人之間的綽號稱呼。
3. 不同年代可能出現同名同姓者。
4. 相同的官名職稱，在不同的原件資料內，可能使用全稱、簡稱等各種不同的稱謂，而同一種官名職銜的簡稱方式，又可能有數種之多。因為 Metadata 是分別由不同人建置，建置者未必有工具書可幫助確認，有的一時也難以確定其是否是簡稱。
5. 有的地名為土名，文件未詳細註明其所屬上一層級之地名，因此難以判斷。

因為判斷人名、地名、官銜職稱等之權威標目，需要具有相當的知識，以及足夠的工具。目前分散建置 Metadata 的情形下，權威控制不易達到相當信度，因此，目前的建檔係依原件內容出現的人名、地名等照錄。建檔者如對某人名或地名等判斷為需要做權威控制，則於另一特闢的「人名註解」欄位，記述人物的真正姓名或別名，未來再批次處理權威控制之問題。

## 八、結論

本館從「淡水河溯源」時期開發各種不同資料群的 Metadata，到目前大量建置 Metadata 以及整合異群資料 Metadata 的過程，雖然大致可劃分為前後二階段，但是基本理念和執行上，前後二個階段是互相貫通連續的，亦即均以完整詳盡的欄位、充分描述資料為原則。

因為各資料群原有的欄位堪稱充足，因此，建置整合性的資料庫時，主要的工作是欄位的對應、小幅度的欄位名稱調整、以及部分欄位內容值的異動。而此類修正性的異動，均屬小幅度更動，因此，整合資料庫查詢系統的建置、階層瀏覽式使用介面的開發、匯入「數位典藏聯合目錄」的過程等，均屬順利。由本館數位典藏 Metadata 之設計與資料庫之彙整建置經驗可知，Metadata 的欄位完整、各欄位的意義與內容建檔方式前後保持一致性，才能裨益資料庫檢索功能的發揮。而且，Metadata 詳盡，可以使未來具有相當的運用彈性。

但是，要建置詳盡的 Metadata，十分費時，使得進度緩慢。相反的，Metadata 過簡，未來在檢索上，得到雷同資訊之比率高，無法幫助使用者篩選出真正適用的資料。如何在二者之間取得平衡，在考驗著執行典藏數位化的機構。

資料庫系統不論是透過欄位檢索或階層式瀏覽而提供查詢使用，查詢上是否便捷有效，其實與 Metadata 欄位的規劃以及內容的確實建置，均息息相關。由

於典藏數位化將是未來許多機構會陸續從事的工作，爲了幫助正在執行中的機構，以及新加入數位化行列的機構，都能使 Metadata 的欄位規劃以及建檔品質達到一定的標準，由對 Metadata 實務較具經驗的聯合目錄小組之類的組織，基於協助之立場，主動適時地、及早地、定時地對執行機構進行建檔欄位設計以及建檔品質之評鑑，並提供改善建議，實有其必要。

## 致 謝

國立臺灣大學圖書館執行數位化計畫多年，從 Metadata 的欄位規劃、建檔、以至於後續的資料庫程式撰寫，均歷經多屆工作同仁的接棒努力。為了達成跨資料庫檢索之目標，館長、前特藏組主任陸續督導，而內容建置單位（特藏組）與後續的資料庫程式撰寫單位（系統資訊組）相關同仁亦無不竭盡心力。如果沒有這麼多工作夥伴的合作協助，多個資料群的整合與資料庫建置的大工程是無法完美達成任務的。謹在此向所有幫助完成現階段工作的主管和工作夥伴們深致謝忱。

### 附錄一 《淡新檔案》總目（戴炎輝教授編製）

編	類	款	各款所含案數
第一編 行政 (共 575 案)	第一類 總務 (共 101 案)	第一款 祀儀	8
		第二款 吏務	17
		第三款 吏紀	27
		第四款 薪津	11
		第五款 外事	11
		第六款 移交	10
		第七款 雜事	17
	第二類 民政 (共 91 案)	第一款 學務	4
		第二款 鄉保	44
		第三款 聯庄	4
		第四款 保甲	6
		第五款 厚俗	22
		第六款 義倉	6
		第七款 救恤	5
	第三類 財政 (共 100 案)	第一款 田賦	11
		第二款 清賦	24
		第三款 稅契	16
		第四款 給照	7
		第五款 釐金	8
		第六款 錢幣	4
		第七款 征收	10
		第八款 解庫	14
		第九款 雜事	6
	第四類 建設 (共 72 案)	第一款 民業	6
		第二款 鹽務	35

編	類	款	各款所含案數	
		第三款 樟腦	14	
		第四款 礦產	10	
		第五款 工程	7	
	第五類 交通 (共 47 案)	第一款 驛站	14	
		第二款 船政	18	
		第三款 鐵路	2	
		第四款 義渡	13	
	第六類 軍事 (共 61 案)	第一款 軍政	8	
		第二款 兵餉	16	
		第三款 城工	7	
		第四款 息庄	18	
		第五款 私火	12	
	第七類 撫墾 (共 103 案)	第一款 社務	15	
		第二款 社租	13	
		第三款 隘務	41	
		第四款 屯務	34	
	第二編 民事 (共 224 案)	第一類 人事 (共 13 案)	第一款 失蹤	1
			第二款 結婚	7
			第三款 離婚	2
第四款 收養			2	
第五款 監護			1	
第二類 田房 (共 138 案)		第一款 租借	7	
		第二款 抗租	27	
		第三款 霸收	3	
		第四款 霸佔	43	
		第五款 爭界	27	
		第六款 爭財	15	
		第七款 公業	10	
		第八款 用水	4	
		第九款 抄押	2	
第三類 錢債 (共 68 案)		第一款 買賣	4	
		第二款 典當	10	
		第三款 胎借	13	
		第四款 借貸	24	
		第五款 委寄	3	
		第六款 討物	4	
		第七款 抗吞	6	
		第八款 抗算	3	
		第九款 匿契	1	
第四類 商事 (共 5 案)		第一款 郊行	1	
		第二款 合股	1	
		第三款 倒閉	1	
		第四款 侵權	2	
第三編 刑事 (共 365 案)		第一類 總務 (共 120 案)	第一款 法令	6
			第二款 冊報	18

編	類	款	各款所含案數	
		第三款 指揮	7	
		第四款 互助移解	22	
		第五款 通緝	42	
		第六款 驗屍	6	
		第七款 定讞	3	
		第八款 執行	14	
		第九款 其他	2	
		第二類 人身自由 (共 61 案)	第一款 人命	7
			第二款 毆傷	17
	第三款 勒贖		1	
	第四款 誘拐		10	
	第五款 略賣		7	
	第六款 擄禁		12	
	第七款 擄搶		7	
	第三類 財產侵奪 (共 123 案)	第一款 竊盜	39	
		第二款 搶奪	19	
		第三款 強盜	30	
		第四款 強盜殺傷	10	
		第五款 海盜	5	
		第六款 恐嚇	3	
		第七款 詐欺	6	
		第八款 盜賣	7	
		第九款 毀棄	4	
	第四類 公共秩序 (共 19 案)	第一款 匪徒	9	
		第二款 侵害水源	2	
		第三款 騷擾	7	
		第四款 謠言	1	
第五類 風化 (共 42 案)	第一款 忤逆	5		
	第二款 毀墳	13		
	第三款 通姦	6		
	第四款 拐姦	5		
	第五款 賭博	13		

共 1164 案

## 附錄二 Metadata欄位總表：淡新、岸裡、拓碑、照片、伊能 Mapping

(說明：本表格係根據「行政院國家文化資料庫詮釋資料格式」1.0 版進行修正。)

欄位名稱		欄位說明	必備/建議 著錄欄位
type 資源類型  <b>1</b>	aggregationLevel 藏品層次  (A) 拓· ----- 岸· ----- 淡·	說明所描述的物件為個別作品(work；單件)或全集作品(collection；合集)。著錄之內容值僅有「單件」或「合集」兩種選擇。	必備
	originalSurrogate 原件與否  (B) 拓· ----- 岸· ----- 淡· ----- 照· 藏品類型-3	說明所描述的物件為原件或重製品。著錄之內容值僅有「原件」或「重製」兩種選擇。	必備
	worksType 作品類型  淡新檔案 拓· ----- 岸· ----- 淡· ----- 照· 藏品類型-1	說明作品的類別，因本表為古文書詮釋資料格式，所以著錄的內容值為「史料」。	必備
format+ 資料格式  <b>2</b>	medium 媒體類型  (A) 拓· ----- 岸· ----- 淡· ----- 伊· 藏品形式: ----- 照· 藏品類型-2	若所描述物件為數位資料，則此欄內容值為「數位檔案」。反之，描述物件為非數位資料，則依其媒體形式、種類著錄，如唱片、盤式錄音帶、匣式錄音帶、CD-ROM、VCD、LD...等。內容值請參見： <a href="http://km.cca.gov.tw/download/rule/format_table.pdf">http://km.cca.gov.tw/download/rule/format_table.pdf</a>	必備

欄位名稱		欄位說明	必備/建議 著錄欄位
B extent? 數量/尺寸	quantity? 數量單位 <b>B-1</b>	所描述的物件之數量及其單位，如圖書之冊數。	
	拓· ----- 岸· 淡· 伊· 稽核資料-頁數 照· 稽核資料-頁數		
	Dimension? 尺寸大小/重量 <b>B-2</b>	所描述的物件之最大體積及重量，註明其名稱、數據、所用的測量單位、位置。尺寸類型包括：高度、長度、寬度、深度、厚度、直徑、重量等，如圖書的尺寸大小。	
	拓· 尺寸大小 岸· 尺寸大小 ----- 淡· 尺寸大小 ----- 伊· 稽核資料-高廣尺寸 ----- 照· 稽核資料-高廣尺寸		
C scale? 規格	色彩 color <b>B-3</b>	原件資料的色彩	
	拓· ----- 淡· ----- 照· 插圖		
	duration? 播放/演出時間長度 <b>B-4</b>	該古文書資料放映或演出時間之總長度。(所描述的物件為錄音或錄影資料時才需填寫)	
	digitalScale* 數位化規格 (數位檔案才需要填寫) <b>C-1</b>	digitalCategory 數位化類別 <b>C-1-(1)</b>	所著錄之內容值為：文字、影像(圖片)、聲音、視訊。
	拓· ----- 岸· ----- 淡· -----		必備

欄位名稱			欄位說明	必備/建議 著錄欄位
		fileType 檔案類型 <b>C-1-(2)</b> 拓· 岸· ----- 淡· -----	說明檔案格式，例如 tiff 檔、jpeg 檔。	必備
		<input type="checkbox"/> ilename 檔案名稱 <b>C-1-(3)</b> 拓· 岸· 淡·	Cca110001-od-ilt0001-0001.jpg	必備
		fileDesc? 檔案描述 <b>C-1-(4)</b> 拓 岸· ----- 淡·	檔案的描述，例如該檔案的用途。 如：預覽檔	
		permission? 使用權限 <b>C-1-(5)</b> 拓· ----- 岸· ----- 淡·	合法使用檔案的範圍。 如：限網路瀏覽	
		fileSize? 檔案大小 <b>C-1-(6)</b> 拓· 岸· 淡· ----- 照· 檔案資訊-大小	檔案的大小。	
		TextFormatVersion? 格式與版本 <b>C-1-(7)</b> 拓· ----- 岸· ----- 淡·	數位化類別內容值為文字時才須輸入，如內容相同但呈現格式不同(繁體字型及簡體字型)所產生的版本可於本欄位著錄說明。	

欄位名稱		欄位說明	必備/建議 著錄欄位
	TextAssociatedDTD? 文件型別定義 <b>C-1-(8)</b> 拓· ----- 岸· ----- 淡·	數位化類別內容值為文字時才須輸入，僅針對 SGML 及 XML 語法所寫之文件，以驗證全文結構的正確性。	
	Resolution? 解析度 <b>C-1-(9)</b> 拓· ----- 岸· ----- 淡· ----- 照· 品質	數位化類別內容值為影像(圖片)時才須輸入。 解析度是指每英吋所含像素點的數目，以 dpi 為單位。	
	BitPerPixel ? 每點位元素 <b>C-1-(10)</b> 拓· ----- 岸· ----- 淡·	數位化類別內容值為影像(圖片)時才須輸入。 例如：灰階-每像素 8-bits；彩色-每像素 32-bits。	
	ImageSize ? 圖片大小 <b>C-1-(11)</b> 拓· ----- 岸· ----- 淡·	數位化類別內容值為影像(圖片)時才須輸入。 圖片長與寬的大小，以 pixels 為單位，若無完整數值，填入長邊即可。 如：長邊 500pixels	
	compressionRatio ? 壓縮比 <b>C-1-(12)</b> 拓· ----- 岸· ----- 淡·	數位化類別內容值為影像(圖片)時才須輸入。 影像(圖片)壓縮前後的比例。 如：10：1	
	color? 色彩  <b>C-5</b>	說明所描述的物件之色彩，如黑白影片、彩色影片。 著錄之內容值僅有「黑白」或「彩色」兩種選擇。 (當作品翻拍為多媒體資料時，才需填寫此欄位)	



欄位名稱		欄位說明	必備/建議 著錄欄位
title 作品名稱  3	mainTitle 主要題名 (A) 拓 · ----- 岸 · ----- 淡 · 伊 · 題名: 照 · 題名:	該古文書的主要名稱，為正式名稱、或多數人所熟知的名稱、作者所命名的名稱等。 如：宋學士文粹	必備
	alternativeTitle* 其他名稱 (B) 拓 · 其他題名： ----- 尾題： 拼音題名： 額題： 首題： 陰題： 伊 · 集叢名	副題名、封面題名、劃一題名、全集題名、翻譯名稱、其他常用或少用的名稱...等，著錄時需加入前導語。 如：全集題名：淡新檔案	
subject ? 主題與關鍵詞  4	subjectMatter* 內容主題 (A) 拓 · 類目： ----- 內容主題： 岸 · 內容主題： ----- 淡 · 內容主題： ----- 關鍵詞 1： 照 · 主題或關鍵詞： ----- 相關事件名稱：	描述該古文書之內容主題或所傳達之意念。	建議
	keywords* 關鍵詞 (B) 拓 · ----- 岸 · 淡 · -----	請填入被描述資料之相關重要詞彙，以利日後檢索之用。 如：宋濂	建議
description ? 簡述  5	abstract* 摘要/說明/導言 (A) 拓 · 內容摘要： ----- 岸 · 內容摘要： ----- 淡 · 內容摘要： ----- 伊 · 內容摘要： ----- 照 · 內容摘要：	該古文書之內容大綱。	建議

欄位名稱		欄位說明	必備/建議 著錄欄位
tableOfContents*	內容目次 <b>B</b>	該古文書之目次、章節等細目。 如：卷第一 皇明寶訓序 大明日曆序...	

(以下略)

### 附錄三 檢索介面

網址(D) <http://140.112.113.6/diglib/QryPage.aspx>

您所選擇的是 淡新檔案，伊能嘉矩手稿，臺灣古拓碑 資料庫檢索

---

**綜合檢索**

學務

● 資料庫會將您使用的檢索詞彙在題名、主題、關鍵詞、作者、人名、地名、年代、媒體類型及語言等欄位中進行檢索。

---

**進階檢索**

檢索欄位個數:

後,請按Enter鍵

<input type="button" value="請選擇"/>	<input type="text"/>	<input type="button" value="請選擇"/>	
<input type="button" value="請選擇"/>	<input type="text"/>	<input type="button" value="請選擇"/>	
<input type="button" value="請選擇"/>	<input type="text"/>	<input type="button" value="請選擇"/>	

紙本     原件複製品     微捲(片)  
 底片     影音磁帶檔     圖書  
 期刊     網路資源     全部  
**媒體**

中文     英文     日文  
 滿文     其他     全部  
**語言**

● 請先選取您的檢索欄位，再輸入檢索詞彙和布林邏輯(And, Or, Not)。  
● 調整檢索欄位個數後,請按Enter鍵。(檢索欄位個數範圍為3到20)

## 附錄四 檢索結果書目呈現格式 (分「簡目」與「詳目」)

說明：主要的基本欄位下，可能包含「複分欄位」(例如圈記之處)。如此既可呈現完整資訊，又可使所呈現的欄位值意義明確。

網址(D) <http://140.112.113.6/diglib/Abstracts.aspx?ID=29&DBName=RT3>

檢索結果(簡目) 顯示詳目

系統識別號	29	 <p>連結全文 <span>到</span> <input type="text"/> 頁 <input type="button" value="Go!"/></p> <p>第 1 個(共 1 個)影像檔</p>
檔案名稱	ntul-nr-rt00029	
主要題名	獎勵楊志申捐獻學田碑記	
其他題名	額題: 皇清; 其他題名: 楊志申捐獻學田碑記; 拼音題名: jiang li yang zhi shen juan xian xue tian bei ji	
文書形式類別	紙本	
作者		
內容主題	教育--學務; 楊志申	
關鍵詞	楊志申; 捐輸; 學田; 臺灣縣學; 楊廷璋; 曹繩柱; 記	
重要關係人	曹繩柱(發文者--福建布政使); 王瑛曾(立碑人--台灣縣知縣)	
相關地名	(原碑地點) 原碑在臺灣縣學(遺址在延平郡王祠附近); (現碑地點) 臺南市中西區南門路 大南門碑林	
作成日期/出版日期	乾隆28年9月(1763年)	
數量/單位	一幅	
規格	234x77 (cm)	
原件與否	原件	

網址(D) [http://140.112.113.6/diglib/Detail\\_RT.aspx?ID=29&DBName=RT3](http://140.112.113.6/diglib/Detail_RT.aspx?ID=29&DBName=RT3)

檢索結果(詳目) 顯示簡目

系統識別號	29
檔案名稱	ntul-nr-rt00029
藏品層次	單件
原件與否	原件
主要題名	獎勵楊志申捐獻學田碑記
其他題名	其他題名: 楊志申捐獻學田碑記; 拼音題名: jiang li yang zhi shen juan xian xue tian bei ji; 額題: 皇清
作品類型	史料 -- 拓本原件
文書形式類別	記
內容主題	教育--學務; 楊志申
關鍵詞	楊志申; 捐輸; 學田; 臺灣縣學; 楊廷璋; 曹繩柱; 記
內容摘要	本件碑記為獎勵貢生楊志申捐獻學田，年收歲粟折價充入臺灣縣學，為師生月課、飯食、紙筆之需的義學，除頒發「好義可風」匾額，並勒石臺灣縣學前，以垂永久。
重要關係人	曹繩柱(發文者--福建布政使); 王瑛曾(立碑人--台灣縣知縣)
地名	(現碑地點) 臺南市中西區南門路 大南門碑林; (原碑地點) 原碑在臺灣縣學(遺址在延平郡王祠附近); 國立臺灣大學圖書館五樓特藏組

網址(D) [http://140.112.113.6/diglib/Detail\\_RT.aspx?ID=29&DBName=RT3](http://140.112.113.6/diglib/Detail_RT.aspx?ID=29&DBName=RT3)

**附錄五 ([表六]) 拓碑、岸裡、淡新、伊能、照片\_資料庫INPUT欄位對應表**

(指示檢索程式於查詢者下查詢指令之後，至資料庫所要抓取的欄位)

資料庫 對應欄位 INPUT 欄位	拓碑	岸裡	淡新	伊能	照片
題名	3/A 3/B	3/A	3/A	3/A	3/A
主題	4/A	4/A	4/A		4/A(主&關)
關鍵字	4/B	4/B	4/B		4/A
作者	6/A 7	6/A	6/A 7	6/A	6/A 7
人名	5/C/C-1 6/A 7	5/C/C-1 6/A	5/C/C-1 6/A 7	5/C/C-1 6/A	5/C/C-1 6/A 7
地名	14/A/A-4 14/A/A-5 14/A/A-6	14/A/A-1	14/A/A-1 14/A/A-2	14/A/A-1	14/A/A-1
案件號	10/C/C-1 的檔 名前二段(從 RT 開始算)	10/C/C-1 的檔 名前二段(從 AL 開始算)	10/C/C-1 的檔 名前二段(從 TH 開始算)	10/C/C-1 的檔 名前二段	10/C/C-1 的 檔名前二段
成文日期 or 出版時間	9/A	9/A	9/A	9/A	9/A
媒體類型	2/A	2/A	2/A	2/A	2/A
語言	13	13	13		13
綜合檢索	3/A 3/B 4/A 4/B 5/A 5/C/C-1 6/A 7 14/A/A-4 14/A/A-5 14/A/A-6 11	3/A 4/A 4/B 5/A 5/C/C-1 6/A 14/A/A-1 11 12/A 12/B	3/A 4/A 4/B 5/A 5/C/C-1 6/A 7 14/A/A-1 14/A/A-2 11 12/A	3/A 5/A 5/C/C-1 6/A 14/A/A-1 11 12/A	3/A 4/A 5/A 5/C/C-1 6/A 7 14/A/A-1
排序	9/A 3/A 3/B 5/C/C-1 6/A 7 14/A/A-4 14/A/A-5 14/A/A-6	9/A 3/A 5/C/C-1 6/A 14/A/A-1	9/A 3/A 5/C/C-1 6/A 7 14/A/A-1 14/A/A-2	9/A 3/A 5/C/C-1 6/A 14/A/A-1	9/A 3/A 5/C/C-1 6/A 7 14/A/A-1

(製表：趙秋萍)

## 附錄五 ([表七]) 拓碑、岸裡、淡新、伊能、照片\_資料庫OUTPUT欄位對應表

(指示檢索程式在讀者端的各特定欄位所要呈現的資料庫內的欄位)

資料庫 對應欄位 OUTPUT 欄位	拓碑	岸裡	淡新	伊能	照片
檔案名稱	10/C/C-1	10/C/C-1	10/C/C-1	10/C/C-1	10/C/C-1
主要題名	3/A	3/A	3/A	3/A	3/A
其他題名	3/B			3/B	
文書形式類別	5/W	5/W	5/W	5/W	5/W
作者	6/A 7	6/A	6/A 7	6/A	6/A 7
內容主題	4/A	4/A	4/A		4/A
關鍵詞	4/B	4/B	4/B		4/B
重要關係人	5/C/C-1	5/C/C-1	5/C/C-1	5/C/C-1	5/C/C-1
相關地名	14/A/A-4 14/A/A-5 14/A/A-6	14/A/A-1	14/A/A-1 14/A/A-2	14/A/A-1	14/A/A-1
成文日期 or 出版日期	9/A	9/A	9/A	9/A	9/A
出版者/出版地					
數量/單位	2/B/B-1	2/B/B-1	2/B/B-1	2/B/B-1	2/B/B-1
規格(尺寸大小)	2/B/B-2	2/B/B-2	2/B/B-2	2/B/B-2	2/B/B-2
原件與否	1/B	1/B	1/B		1/B
媒體類型	2/A	2/A	2/A	2/A	2/A
影像檔名	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)
全文檔名	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)	2/C/C-1/C-1-(3)
影音視訊檔名					

(製表：趙秋萍)

## 附錄六 數位典藏國家型科技計畫檔案主題小組「後設資料核心元素」與國立臺灣大學圖書館〈淡新檔案〉欄位對應表

- 註：1. 該表格係摘錄自「檔案主題小組『後設資料核心元素』彙整結果：附件一：檔案後設資料核心元素清單與調查結果」。(註9)
2. 元素名稱以網底標示者，表該元素為「必備元素」，以粗黑字體標示者表示其為臺灣大學圖書館 2004 年匯出之欄位。

核心元素		定義	
中文名稱	英文名稱		
15/B/B-2	<b>1.典藏單位</b>	Repository	負責典藏資料的機關單位
10/C/C-1	<b>2.館藏號</b>	Collection number	辨識資料的唯一號碼，或所屬層次的唯一號碼。
5/W	<b>3.類型</b>	TYPE	資料的形式或類別
1/A	<b>4.描述層次</b>	Level of Description	該筆資料的層次
3/A	<b>5.題名</b>	Title	資料的題名或資料分類層級之分類名稱
下階段給	<b>6.內容描述</b>	Content Description	對資料內容的提要、摘要或補充說明。
9/A	<b>7.時間</b>	Date(s)	資料產生日期，或業務處理過程中所彙積資料的起迄日期
6/A	<b>8.產生者</b>	Name of creator	資料的創造者或產生者 (可能是個人或團體組織)
14/A/A-2	<b>9.產生地</b>	Geographic name	資料創造或產生的地點
	<b>10.傳記歷史註</b>	Administrative/ Biographical	1、資料的行政歷史或傳記資訊，如：組織之起源、發展演變或與工作有關的重要資訊；或個人之一生或工作等傳記資訊。 2、資料所有轉移權、權責與保管之沿革或相關活動，對現有檔案的結構與編排有關的資訊。
	<b>11.範圍與內容</b>	Scope and content	資料的範圍(時期或地理)與內容(文件形式、主題性質、行政處理過程)的摘要
4/A	<b>12.標題資訊</b>	Subject	資料內容中所提到的重要人名、地名、團體名稱、家族名稱、主題名稱等，可用以做內容標目者。 (可依館內需求決定是否以此進行權威控制)

(以下略)

## 附錄七 《淡新檔案》之階層式檢索

1. 由國立臺灣大學數位典藏計畫第七子計畫建立之《淡新檔案》的階層式檢索介面

<http://140.112.30.244/darc2/>

The screenshot shows the website interface for the Digital Archives Project of National Taiwan University. The search results are displayed in a table format, listing various historical documents related to currency and finance. The search criteria used are ((subject:"錢幣") AND (subject:"財政")).

Thumbnail	Document Title	淡新檔案: 第	內容主題: 行政 -- 財政
	據本城郊舖金長和稟不法奸徒私往內地採買錢幣來臺發兌叩請示禁由	13601 案	
	郊舖金長和為利弊攸關不平則鳴懇恩曉諭嚴禁以資地方事	13601 案	
	署新竹縣正堂劉為曉諭嚴禁事	13601 案	
	簡述: 署新竹縣正堂劉元陞出示郊商舖戶民人等買賣須用制錢嚴禁擅用小錢		
	署新竹縣正堂劉為出示曉諭事	13601 案	
	簡述: 署新竹縣正堂劉元陞出示各郊商舖戶民人等遵用制錢一律公平交易不得混雜錢錢抬高物價		
	署新竹縣正堂施為出示嚴禁事	13601 案	
	簡述: 署新竹縣正堂施錫衛出示各郊商舖戶民人等嚴禁私鑄務須公平交易概用制錢		
	署新竹縣正堂施為出示嚴禁事	13601 案	
	簡述: 署新竹縣正堂施錫衛出示嚴禁小錢務須公平交易概用制錢不准拱抬物價		

2. 由數位典藏國家型科技計畫所建置之「數位典藏聯合目錄」中的《淡新檔案》之階層式檢索介面

<http://catalog.ndap.org.tw/System/Catalog/Catalog.jsp>

The screenshot shows the Union Catalog of National Digital Archives Program interface. It features a hierarchical search structure on the left side, allowing users to navigate through various categories. The current search results are displayed in a grid format, showing the number of items in each category.

Category	Number of Items
保甲	6分項
厚俗	22分項
救恤	5分項
鄉保	44分項
義倉	6分項
學務	4分項
聯庄	4分項

## 註釋

- <sup>1</sup> 數位典藏國家型科技計畫數位典藏技術彙編〈2002年版〉第1冊，「淡新檔案欄位及其與臺灣史料 metadata 欄位對應表」，(民國91，頁1.5.2.11.1-1.5.2.11.5)。
- <sup>2</sup> Dublin Core 由最初的十二欄位陸續增加為十五欄位、十六欄位。
- <sup>3</sup> 完整的 Metadata 記錄，可說是為典藏資料作成詳實記錄，有利於未來本詮釋資料與實體物件間之查核比對（例如，典藏單位可據以盤查典藏品是否曾被毀損或調換），以及不同實體物件間之比對（例如，不同單位典藏名稱相同之典藏品的情況，可根據對資料之詳細描述例如原件與否、大小尺寸、其他特色等，而確定哪一單位的典藏品為原件，或確定二者是否為同一時代、同一版印製的出版物）。如果將各典藏單位所建置的所有 Metadata 加以彙整，建置為聯合目錄，則足以反應國家文化財之內涵與全貌，並顯示各文化財之所在位置。（洪淑芬，文獻典藏數位化的實務與技術，數位典藏國家型科技計畫數位典藏叢書 E001（台北市：數位典藏國家型科技計畫訓練推廣分項計畫，民國93），頁55。）
- <sup>4</sup> 本館訂定《臺灣古拓碑》之 Metadata，曾參酌中央研究院歷史語言研究所傅斯年圖書館當時之金石拓片 Metadata 欄位，加上針對本館館藏實況之必要考量而擬定。（洪淑芬，「國立臺灣大學臺灣文獻文物典藏數位化計畫之臺灣古拓碑數位化工作流程簡介」，在數位化工作流程—金石拓片主題小組，數位典藏國家型科技計畫 內容發展分項計畫，（台北：數位典藏國家型科技計畫 內容發展分項計畫，[2005]）。
- <sup>5</sup> 洪淑芬，文獻典藏數位化的實務與技術，數位典藏國家型科技計畫數位典藏叢書 E001（台北市：數位典藏國家型科技計畫訓練推廣分項計畫，民國93），頁56。
- <sup>6</sup> “DC-Library Application Profile,”  
<<http://dublincore.org/documents/library-application-profile/index.shtml#Identifier>> 該 Profile 的作者 Rebecca Guenther 對 Identifier 的說明是：“Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Example of formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).”另外，對於實務上的做法則建議加上館藏機構代碼：  
“*Unique-resource identifier*: Provide applicable identifiers assigned to one-of-a-kind resources (such as accession numbers assigned to items in a museum collection) if one or more of this class of identifiers have been assigned to the resource and are known to the metadata agency. If the structure of the identifier (e.g., all numerals) is potentially ambiguous it is recommended that the associated agencies' name be included as part of the identifier element.”
- <sup>7</sup> 當時本館對於欄位的內容值之詮釋，部分有異於「國家文化資料庫」的「詮釋資料格式」之表格內容值說明。兩者相異時，仍維持本館對於欄位內容值之詮釋，僅採用「國家文化資料庫」之「詮釋資料格式」之欄位表格以幫助彙整的



完成。「國家文化資料庫」之「詮釋資料格式：古文書」目前已更新至 1.4 版(「行政院國家建設委員會國家文化資料庫知識管理系統：相關規範——詮釋資料：古文書」，< <http://km.cca.gov.tw/> > (retrieved 2005/2/19)

<sup>8</sup> 數位典藏國家型科技計畫 後設資料工作組，「檔案主題小組『後設資料核心元素』彙整結果」，(民國 92 年 4 月 23 日)。

<sup>9</sup> 同註 8