

# 數位典藏建置規劃與管理

賴忠勤 國立臺中圖書館資訊小組執行秘書

【摘要】本文為「宗教史料數位典藏研習會」之講稿整理，以國立臺中圖書館數位典藏的經驗談起，內容包括數位化的目的、數位典藏的架構、建置數位資源要點、數位化工作流程、數位檔案命名、數位版權、資訊儲存及數位檔案保存，希望可以對想做數位典藏的單位，提供參考。

關鍵詞：數位化；數位典藏；Metadata；數位版權

## 一、為何要數位化

當圖書館或典藏單位考量要進行數位化時，首先要先思考為何要數位化？因為數位化是一項高成本的作業。以國立臺中圖書館數位化的幾個計畫為例，數位化相關費用如下：一般圖書掃描（全彩）：9-15 元/頁，Metadata 建檔：30-100 元/筆，全文建檔（含校對）：0.2 元/字，數位化管理、檢索系統：從免費到 120 萬，數位典藏系統主機和儲存設備：40 萬。若以年度經費為例，國立臺中圖書館一年一個普通的數位化個案，經費就高達五百萬元以上。（見表一）因此是否進行數位化作業，必須認真考量其價值性，畢竟所費不貲。

## 二、進行數位化應達到之目標

數位化的高成本，假如初步規劃進行數位化前，請先認真思考以下幾個問題：

### （一）保存文化資產

進行數位化的資料，是否達到具有文化資產價值來保存。例如：許多照片數位化，如果只是近期的照片，而不是歷史照片，就尚未達到值得數位化的時間點。

### （二）減少原件管理成本及對原件取用的損害

典藏的資料若取用頻繁，為了減少原件管理成本及對原件取用的損害，數位化之後，這些典藏資料只需在必要時才予以取用，一般只要直接查看數位檔即可。

### （三）便於取用及分享

典藏的物件若使用頻率不高，短時間也沒有損壞之虞，在經費有限的情況下，可以加強保存環境的改善，暫緩數位化的進行。



表一、日文舊籍數位化作業經費概估範例

經費項目及內容	單價	單位	數量	總價	備註
<b>【委外項目】</b>					
1. 影像掃描作業	\$11	頁	337,319	\$3,710,509	
2. 影像檔轉存 (JPG 檔)		批		\$100,000	
3. Metadata 資料建檔及校對費	\$100	冊	692	\$69,200	
4. 圖書裝訂修復作業	\$400	冊	692	\$276,800	
5. 儲存與備份媒體製作：		份	1	\$250,000	
(1) DVD 光碟片、光碟燒錄與拷貝、光碟片面與外盒標籤製作、塑膠單片盒。					
(2) 硬碟一份					
6. 資料庫檢索系統功能加強			1 式	\$100,000	
7. 電子書主題及冊數增加			1 式	\$300,000	
<b>委外項目合計</b>				\$4,806,509	
<b>【自辦項目】</b>					
1. 學者專家出席討論或審查之出席費	\$2,000	人數*次數	18	\$36,000	3 人 x6 次
2. 外聘學者專家差旅費 (交通費、膳雜費)		式	1	\$18,000	
3. 僱請工讀生 (含日薪及勞健保費用)	\$25,000	月	10	\$250,000	
4. 雜支 (文具、郵資、餐點及其他支出等)		式	1	\$200,000	
<b>自辦項目合計</b>				\$504,000	
<b>總計 (委外+自辦)</b>				\$5,310,509	



#### (四) 數位典藏加值應用

近年來數位加值應用為數位化作業進一步延伸的要求，尤其花費公務預算所進行的數位化作業，其成果就是要讓全民分享，不能做完就限在單位內部使用，而且許多數位化的物件是比較歷史研究相關的範圍，如果開放並吸引大眾使用，就必須做一些加值性的設計及功能開發，讓使用者可以檢索及瀏覽到更有用的內容，甚至設計結合遊戲及設計因素，進行進一步加值性利用。

#### (五) 整合、檢索，產生新發現

許多數位化的史料，經過建檔、整合及檢索之後，往往會產生新的發現，若更能支援服務的使用者，協助其研究發展，便具有數位化的價值。

假如有達到以上幾個目的，就可以開始準備數位化，開始進行規劃作業，撰寫計畫書爭取經費。

### 三、數位典藏一般架構

從國立臺中圖書館進行數年數位化的經驗，個人認為一般的數位化作業並非屬於高技術性的工作。以圖書資料的數位化作業為例，掃描、建檔、文字辨識（如果有需要）、數典系統購置與匯入資料，即可完成基本的數位化要件。因此建立管理 Metadata 和電子全文、多媒體（如：影像、聲音、圖形、視訊、3D）物件的系統。（見圖一）不過同樣的掃描作業，也可以做到非常精細的解析度要求、校色、放置色卡等作業。如果一般典藏機構對其藏品經評估後值得進行，就可以開始依相關的規範著手進行。

### 四、數位化要點

進行數位化作業有幾項要點，有助於將數位化工作做好：

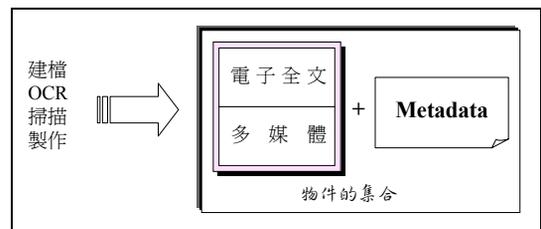
#### (一) 遵循、採用標準

數位典藏國家型科技計畫及文建會國家文化資料庫網站，已制訂非常多有關數位化作業 Metadata 或數位化物件等規範，因此日後執行的數位化作業性質類似，原則上儘量採用已有的規範，一方面可以減少再新擬訂數位化作業的規格所耗費的精力；另一方面日後如要進行交換，比較不會有格式不同的問題產生。例如掃描檔案一般用 tiff 檔作為備存，jpg 檔作為線上瀏覽之用，但如果自訂採用比較特別的 png 檔，未來就較難分享給其他系統使用。

#### (二) 避免重複

不諱言現在國內有一些數位典藏計畫有重複的情形，大部分是性質相近的資料，分別在不同數典計畫執行，結果在不同的數典系統上運作。例如國內有許多機構藏有日治時期留下大量且珍貴的日文圖書，卻有多個數位化作業在不同的機構或大學圖書館執行。會有這種現象，有時是經費來源不同，而國內兩大數典計畫又缺乏橫向的行政連結與分工，造成重複的情形。

圖一、數位典藏簡要作業內容架構



### （三）採用共同系統（著錄系統、檢索系統）

數位典藏計畫曾發展免費的數位典藏系統，例如 Metalogy 系統即能提供給機關著錄 Metadata、檢索等功能。此系統雖然功能不是很強，但是具有基本著錄、儲存、檢索等功能，對於經費不甚充足的典藏單位可以考慮採用，除非經費充裕，或想開發特殊的功能，才需要特別委外設計開發一套符合典藏單位需求的客製化系統。

### （四）分散建置、集中管理

數位典藏聯合目錄運作方式，是將物件的典藏是分屬各個機構之系統，透過 OAI 等協定，將各系統數位典藏 metadata 集中至聯合目錄提供查詢。另外國家文化資料庫在各附屬機關建置完成後，要求上傳一份完成的資料（包含 metadata 及數位化物件）到國家文化資料庫典藏，不過各機關數位化的資料容量已非常龐大，若全部集中到國家文化資料庫系統，其支援的儲存空間必定非常驚人，也可能會產生超出管理負荷問題。因此分散建置，集中管理是目錄或全部物件資料，需考量集中管理單位本身的支援能力。

### （五）資料交換、備存

如同圖書館同樣採 MARC 格式為圖書資料的建檔格式及標準，有利於資料的交換及保存。進行數位化作業的機構若能依循既有的 metadata 格式建檔，不自創 metadata 格式，對日後資料的交換是相當有幫助的。

### （六）數位物件檔案轉換

隨著電腦軟硬體設備的演進，數位化物件檔案未來一定會遭遇必須進行檔案格式的轉換問題，才能在未來的電腦設備運用。可惜到目前為

止，仍無法有一個較好的解決方案。因為目前電子資源一般可以掌握的不過十多年的資料而已，實在很難去驗證未來二十年後的電腦能否讀取現在產生的檔案。而檔案轉換也需要經費支應，而且又要避免品質減損問題，因此各種不確定因素非常大。

### （七）著作權

著作權是數位化作業最不容易處理的問題，若機構屬圖書館性質，就可以保存典藏的需要「複製」一份，但也僅能在館內使用，無法透過網路供外界連結使用。

### （八）永久保存議題

與數位資料轉換關係密切，可以設定一個議題：「五十年後，這些數位檔案能否可以繼續使用？」再視資訊科技的演進，隨時調整儲存、轉換策略，或許就可以達到此一目標。

## 五、數位化工作流程

數位化作業之流程仍需視各機構的特性，規劃出適合機構作業流程的方式。以下主要參考「數位典藏國家型科技計畫—內容發展分項計畫」所出版，由林彥宏所著《文書檔案數位化工作流程參考標準》，為各位說明數位化作業時共通的工作流程。

### （一）前置作業

不論是經過審慎評估，或是政策上就已經決定進行數位化作業了，著手進行的前置作業包括下列數項：



### 1. 確認數位化史料範圍

確認欲數位化資料的範圍，如數量、性質等。尤其數量一定要加以估算，因為後續撰寫計畫書時將影響整體經費的精確度。

### 2. 史料檔案整編

為數位化作業所做的史料檔案整理工作，必須更為仔細，尤其之前整理就不是非常理想化的館藏資料，整理的項目如建立目錄清冊、記錄史料檔案狀況是正常、蟲蛀霉蝕、散落、修裱、遺失、大尺寸、浮貼……等。以國立臺中圖書館進行日文舊籍數位化作業為例，其中有些插頁是比較大的尺寸，整理人員會特別標註，日後要求廠商特別處理這些特殊規格的資料，相對會影響到整個成本的估算；又如浮貼，古書有很多浮貼，這在數位化時也是一個很重要的史料之一，所以要把它標註清楚，才不會發生資料數位化後散失或忽略數位化的情形。

### 3. 數位化方案選擇

#### (1) 自製、委外

執行數位化作業二種主要方式：自製與委外。自製由館員或工讀生進行掃描、建檔的工作，並且需要品質控制；還要有資訊技術的背景，架設伺服器（server），安裝數典系統及設定，將數位化 metadata 及物件放上去；委外作業則是爭取一筆經

費，然後委託專業廠商執行。以國立臺中圖書館而言，均採委外方式辦理，因為自製所需要的人力、產量較難掌握，而且經費科目也不適合用在雇請工讀生。委外作業若能選擇到作業品質管控優良的廠商，將能為業主（圖書館）節省許多精力，不過委外廠商的選擇就變得非常重要。

#### (2) 數位化規格

數位化規格的訂定，需要視資料的性質及使用的目的，加以適當的區分。以國家圖書館進行圖書、期刊數位化作業為例（見表二），將影像檔分為三種規格，解析度最高，作為備存用的 tiff 檔，600 dpi 是非常高的，如果只有文字而沒有圖片、照片，一般只要 300 dpi 就足夠了。第二種 jpg 格式，主要是放在網路上提供瀏覽下載，有些單位 jpg 格式的壓縮比若未調大，即使下載用的 jpg 格式也可以用在出版印刷上，原則上失真也不會很大。當然用於出版最佳的格式是 tiff 檔，但檔案非常大，例如報紙一版約 A2 大小的數位化檔案，用全彩掃描就有 50、60MB，檢視檔案的時間就要等 30、40 分鐘以上。第三種影像檔為 gif 檔案，作為預覽用。就個人觀察，有些數典系統並不一定把預覽格式轉成 gif 檔，有的只是將 jpg 檔縮小應用，因為在寬頻網路上並不會影響影像顯示的速度。

表二、數位化規格（國家圖書館期刊報紙數位化作業）

檔案格式	建議規格	說明
數位化標的以文字為主		
資料永久保存 格式	檔案格式：TIFF 色調深度：黑白；灰階-每像素 8-bits；彩色-每像素 24-bits	將資料數位化典藏，保持原有風貌。提供使用者作重



檔案格式	建議規格	說明
	壓縮：不壓縮 解析度：300~600（或更高）dpi（依原始資料品質及重要性選擇適當解析度，一般印刷品可採 300dpi）	製、壓縮處理或其他圖像處理交換之用。
網路下載格式	檔案格式：JPEG or JPEG 2000 色調深度：黑白-每像素 8-bits；彩色-每像素 24-bits 壓縮：JPEG（灰階壓縮比約 10:1，彩色壓縮比約 10:1） 解析度或影像大小：150dpi~300 dpi，或影像大小從 500×400 至 1000×700pixels	提供使用者網路上觀看及列印用。
預覽影像	檔案格式：GIF 色調深度：每像素 8-bits 壓縮：原生影像至 GIF 解析度或影像大小：72dpi，或影像大小從 150×100 到 200×200 pixels	提供使用者預覽及選擇欄位用。
數位化標的以影像為主		
資料永久保存格式	檔案格式：TIFF 色調深度：灰階-每像素 8-bits；彩色-每像素 32-bits 壓縮：不壓縮色彩濃度值 4.0D 以上(color)，3.2D(B&W) 解析度：300~600（或更高）dpi（依原始資料品質及重要性選擇適當解析度，一般印刷品可採 300dpi，美術品供複製畫使用建議採 600dpi，供印刷出版使用採 350dpi）	將資料數位化典藏，保持原有風貌。提供使用者作為重製、壓縮處理或其他圖像處理交換之用。
網路下載格式	檔案格式：JPG（JPEG 交換格式） 色調深度：灰階-每像素 8-bits；彩色-每像素 24-bits 壓縮：JPEG（灰階壓縮比約 10:1，彩色壓縮比約 20:1） 解析度或影像大小：150~300 dpi，或影像大小從 500×400 至 1000×700 pixels	提供使用者網路上觀看及列印用。
預覽影像	檔案格式：GIF 色調深度：每像素 8-bits 壓縮：原生影像至 GIF 解析度或影像大小：72dpi，或影像大小從 150×100 到 200×200 pixels	提供使用者預覽及選擇欄位用。



#### 4. 工作規劃、計畫書撰寫

工作規劃以計畫書的形式呈現，是獲得上級支持及爭取經費的重要步驟。

##### (1) 計畫書格式

不同計畫案自有其不同計畫書格式要求，以國立臺中圖書館申請國家文化資料庫所撰寫的計畫書格式為例，省略部分較特殊的項目，大致包括如下內容：

##### I. 前言

以三分之二的頁面對計劃內容作一簡單概述，包含為何要規劃本項計畫（緣由）等內容，可供上級長官或審查委員很快了解計畫內容概要。

##### II. 計畫目標

以條列方式，簡明扼要列出本計畫欲達成之目標。

##### III. 數位化內容及範圍

說明欲數位化資料的內容、範圍和重要性，尤其若要送審爭取經費，就必須強調這批資料的重要性，以獲得審查委員的青睞。所呈現的內容包括數位化數量，最好能以較精確的數量呈現，例如圖書幾冊，頁數約幾頁。

##### IV. 資料所有權及使用權

因牽涉資料重製與公開傳送使用，因此必須釐清著作財產權問題。國中圖向文建會爭取經費補助，製作完成的數位化物件均須上傳至國家文化資料庫，因此特別重視著作權問題。

##### V. 執行內容、工作項目及步驟

對於執行內容加以說明且具體呈現。尤其主要工作項目，必須分項呈現，如掃描數量、Metadata 建檔數量，甚至包含資料修復的數量

等。實施步驟及期程可以用甘特圖呈現，有時上級單位會要求預估未來三、四年的期程，所以要將本年度之時程、未來年度之期程一起列出來。

##### VI. 資源分配

在人力方面，要求將現有人力配置、需求人力分配清楚。如：在做計畫時，有時會要求聘請工讀生，或單位內部有多少人力可以來支援這個專案，所以就必須把人力交待清楚，這樣上級審查時才能感受到人力欠缺，才能同意聘請數位化助理或工讀生。

另外，經費是非常重要的項目，由於整體經費有限，有時評審會特別挑經費項目檢視是否合理，如果經費編列不合理，甚至會有整項都被刪減的危險。所以原則上要詳列計價方式，並可略為高估一成左右，可供其他未及列入但必須支用的項目使用。

##### VII. 預期效益及影響

預期效益及影響應盡量予以量化呈現。

##### VIII. 附件

詳細數位化資料清單或其他補充項目。

其他項目內容視需要再予以增加，例如國中圖有時在計畫書上列出目前使用的主機設備、系統及儲存空間，作為審查參考。

##### 5. 規格需求書

數位化作業若要採委外作業，因招標的需要，必須撰寫「規格需求書」，將執行細節詳列清楚。

##### 6. 委外管理

進行委外管理的目的，主要在檢視廠商是否依進度進行，數位化成品是否有不符需求書內



容，以期能及時發現並要求改善。否則一直到最後驗收時才發現問題，就會非常不易解決。國中圖一般會規劃期中審查和期末審查，期中審查的目的是要監控廠商是否有照進度進行及品質控制，並可作為分期驗收付款的依據。有時也會派員到廠商的工作現場勘察，檢視掃描環境的防盜、消防安全等；另外，在運送時，也派車跟隨，怕運送車會有狀況發生。

## (二) 物件數位化程序

若數位化作業決定採自製方式，就需要了解以下物件數位化程序的工作流程。

1. 建立色彩管理模式，包括顯示器校色和掃描器校色。顯示器校色（見圖三），一個星期要做一次螢幕校色；掃描器校色，是將 IT8 校色卡（見圖四）放到掃描器上配合軟體使用，把不符合的色彩調校回來。好的機器會自動校色，也可以要求廠商每天進行校色作業。
2. 數位化資料掃描程序，包括典藏資料點交及歸還、設備檢查、準備資料、進行掃描等作業。

典藏資料點交及歸還作業，保管人員要將資料交給進行數位化人員時，要確實做到歸還並簽名，並確認先前整編發現的問題。

其次是掃描機器設備檢查，要在掃描前 20 分鐘開機予以熱機，因為若開啟後馬上進行掃描，可能有色彩會跑掉。

對於準備掃描資料，通常開始掃描之前會先準備一批資料，不要把掃描的資料一次全拿到掃描器旁邊，採分批拿取，約半小時作業時間左右的數量，才不會在掃描時造成混亂的情形。

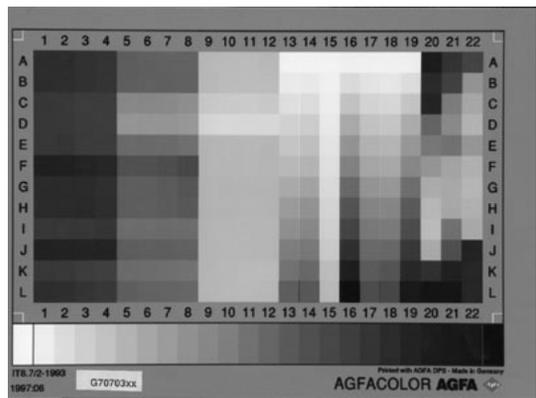
3. 數位影像後製

### (1) 校驗

圖三、顯示器校色



圖四、IT8 校色卡



包括檔案能否開啟，是否符合命名原則，檢查影像大小、解析度、檔案格式、色彩模式，歪斜度，重複、漏頁、摺頁、皺折，留白要求，雜訊或模糊不清，進階校驗可檢視 256 色階分配。

校檢作業最好在掃描完成後，就馬上進行。一般歪斜的情形，會要求最大不能超過 3 度或 5 度；掃描後要檢查檔案是否可以開啟，因為用 tif 檔儲存的檔案非常大，檔案愈大相對的損壞率也就愈高；有的檔雖然可以開啟，但影像有一半以黑色呈現或像打馬賽克一



樣，這可能是在存檔或複製時，檔案已經毀損。再來是檢查命名是否符合規範，如果命名錯誤，就會整批錯誤，將來以 Metadata 作連結時會發生問題。另外，可以開啟看圖軟體來看檔案的大小（size）、解析度、檔案格式、色彩模式是否符合要求。還有重複、漏頁、摺頁、皺折等。掃描影像應要求留白，文字約 1-2 公分的空間，圖片約 3-5 公分的空間。

#### (2) 影像轉檔

校驗最高解析度的檔案沒問題後，就可以進行影像轉檔，轉為不同解析度、壓縮比的 jpg 檔或 gif 檔。可以採用軟體整批轉檔，為了避免轉檔出錯，最好還是檢查一遍。

#### 4. 數位影像備份及管理

將製作完成的數位物件轉錄至儲存媒體，如 DVD 或外接式硬碟。轉錄至 DVD 光碟注意事項，以圖書而言，同一卷（冊）之影像檔案儘量在同一片，還要預留光碟空間，勿將整片光碟全部燒滿，說明文件要在同一片光碟。此外，儲存媒體要標示清楚；儲存媒體的保管，要多份備存於不同位置，還有保管政策，如中研院數位典藏有一個保管政策，同時做五份資料備存於中研院不同的研究所，並定期輪替更新。轉錄至硬碟可以方便上傳至系統主機，速度較快，也可以節省換片的時間。

#### (三) Metadata 與資料庫建置

Metadata 與資料庫建置內容，包括確認 Metadata 欄位、資料著錄方式二種。其中，資料著錄又可分為訂定著錄規範、資料建檔、校對三項。

Metadata 的建置，視數位化典藏資料的內容，有時需要尋求學科專家協助，使 metadata 紀錄品

質更完善。Metadata 著錄相關規範，可以參考國家圖書館《圖書、期刊制訂 Metadata 著錄規範》（<http://catweb.ncl.edu.tw/metadata/mici00070510.PDF>）。另外 Metadata 建檔後，至少應該進行一次校對作業，避免建檔輸入錯誤情形。

#### 1. 何謂 Metadata

Metadata 在國內有幾種講法，例如：詮釋資料、元資料、後設資料、資源描述格式、超資料等。中研院數位典藏採用「後設資料」的稱呼。詮釋資料是對資料（資源）的描述格式，目的在促進資訊系統對資料（資源）的檢索、管理與分析。具有傳統目錄「著錄」功能，目的在使資料的管理維護者及使用者，可透過詮釋資料了解並辨識資料，進而去利用和管理資料。可以提供系統上一個更便利的擷取管道。雖然目前資訊檢索系統可以達到將文獻內容經過文字辨識，並提供全文檢索的方式，不過就檢索的精確度上，仍需要利用 Metadata 上的著錄資料。

#### 2. Metadata 的重要性

- (1) 詮釋資料紀錄比原始資料（典藏品）之檔案小很多。
- (2) 結構化之紀錄格式，易於處理、檢索；尤其對於非文字物件的搜尋。
- (3) 可提高檢索的精確度。
- (4) 可提供不存在於原始資料（典藏品）本身之資訊，例如：主題、稽核資料、附加說明、授權範圍等。
- (5) 可增進原始資料（典藏品）數位化管理的容易度。

#### 3. Metadata 與全文資料檢索

Metadata 經由人工分析建置，將原始資料（典



藏品)中的重要資訊擷取或標示出來,原則上檢索 Metadata 可獲得較精確的結果。全文資料檢索的結果數量,但檢索結果不精確;不過透過適當的關鍵字輸入,也可以獲得較為精確的結果。

國中圖之前進行的日文舊籍、舊版報紙及古文書等數位化作業,也都有予以全文處理的規劃,不過目前全文化的成本仍高。之前評估舊報紙進行全文資料化,可惜用文字辨識(OCR)技術,僅能成功辨識約 20%,還不如交給廠商送往大陸進行全文建檔。未來全文化還是必定要進行的加值作業,只是成本必須加以考量。

#### 4. Metadata 之種類

- (1) MARC 格式,就是一般書目格式。如國中圖日文舊籍數位化,就直接採用既有的 MARC 格式建檔。
- (2) CIMI (Computer Interchange of Museum Information),供博物館藏品、圖書館特藏等之描述格式,採用 SGML DTD。
- (3) EAD (Encoding Archival Description),供檔案、手稿特藏的描述格式,網路上有現成格式可供參考。
- (4) GILS (Government Information Locator Service),由美國政府所訂定,除描述文獻式資料外,亦可描述簡單的地理空間性及時間性。
- (5) TEI (Text Encoding Initiative),原應用於人文學及語言學電子文件的描述格式,目的在促進資源的交換,並應用到其他學科範疇的電子文件上。
- (6) Dublin Core,描述網路上電子資源的簡單格式,目前最常被應用,許多 Metadata 格式都

由 DC 延伸應用。

Dublin Core 由初始的 13 個項目欄位,後來又增修為現在的 15 個欄位:主題與關鍵詞 (Subject & Keywords);題名 (Title);創作者 (Author & Creator);簡述 (Description);出版者 (Publisher);其他貢獻者 (Other contributors);日期 (Date);資源類型 (Resource type);資料格式 (Format);識別資料 (Resource identifier);關連 (Relation);來源 (Source);語文 (Language);時空涵蓋範圍 (Coverage);權限範圍 (Rights management)。

Dublin Core 的特色為:

- I. 簡單,為一組描述網路資源之簡單欄位,各項目欄位易懂易學。
- II. 語意可互相轉換。
- III. 具擴張性,欄位具可重複性、可延伸性、可檢索性。
- IV. 網路上的標準,科際整合、國際化。

以「國家文化資料庫」為例,有針對不同種類數位典藏物件建立詮釋格式,如老照片 XML DTD,下載到數位典藏系統後,就可以直接產生系統 metadata 著錄欄位,若有需要,再加修正。文字性的版本可在 <http://km.cca.gov.tw/download/metadata.html> 獲得。

## 六、數位檔案命名格式

個人發現國內不少機構在進行數位化時,並不是很重視數位典藏物件檔案命名。有時只要求在建構的系統上能順利擷取顯示,而忽略了未來系統發展或與外界交換時,會造成數位檔案名稱一樣的問題。



數位檔案命名要具有唯一性、可辨識、連結性，其原則主要包含：

1. 可由檔名辨識此資料是由公立、私立單位或個人所提供。
2. 此命名方式可支援同一物件有多種檔案格式及使用目的。
3. 此命名方式在整個系統中，每一數位資源皆有其唯一之檔名。
4. 檔案名稱與 Metadata 相結合。

另外，還必須符合下列規則：

1. 使用 ASCII code 命名。
2. 檔案名稱一律使用半形英文小寫字母及數字構成，不可使用中文或全形。
3. 檔名不可包含下列字元/ \ : \* ? " ' < > | ! @ # \$ % ^ & ( ) + = { } [ ] , °。

針對數位檔名結構，以國中圖日文舊籍數位化為例說明如下：

**單位代碼-物件類別-物件代碼-多部分之序號-使用目的-附屬檔名**

**說明：**以上五段皆為必備，各段間以“-” (dash)

作為分隔符號。

單位代碼：cca100003。（國立臺中圖書館之單位代碼）

物件類別：日文書之物件類別代碼為 thm (Taiwan Historical Materials)，以 thm 表示。

物件代碼：bjn+流水編號（5 位數字）+冊次 (Volume)；冊次以 v+冊次（2 位數字）方式呈現，若僅有單冊，則予以省略。如編號 1 之書目，因總共有 9 冊，因此第 1 冊之物件代碼為 bjn00001v1，第 2 冊為 bjn00001v2；編號 2 僅有 1 冊，物件代碼為

bjn00002。

多部分之序號：

書名頁代碼為 tp (Title Page)，以 tp 表示。

例言或序 (Preface)：pr+頁次（2 位數字），如第一頁為 pr01，第二頁為 pr02。

目次 (Table of Contents)：tc+頁次（2 位數字），如第一頁為 tc01，第二頁為 tc02。

正文以 5 碼表示頁次，如第一頁為 00001，第 500 頁為 00500。

版權頁代碼為 cp (Copyright Page)，以 cp 表示。特殊之圖或表頁（如尺寸大於原書者）：

原書已編頁碼者：命名原則依前列「正文」之規範。

原書未編頁碼者：依夾頁所在之前頁頁次+att (Attachment)+流水號（2 位數字），如 00001att01，表示正文第一頁之後為圖或表夾頁，其中第一頁的編碼。

圖書原件封面以 co 表示。

圖書原件封底以 bc 表示。

使用目的：長度 1 碼，以 t 表示預覽用、u 表示永久保存版、i 表示供線上參考/下載用。

依檔案格式給予附屬檔名，如：tif, jpg, pdf 等。

以上的檔案命名方式，在數位化作業進行前當然會盡量就可能的形式制訂，不過實際進行作業時，還是會有例外的情形。因此在數位化作業進行時，需要隨時就例外情形加以討論研訂，以利後續作業的進行。

## 七、數位版權

就行政上，當然必須先釐清數位典藏未來完成開放，有無侵犯到著作權問題。就技術面而言，較常應用的有浮水印技術。浮水印有顯性、隱性兩種方式，國中圖採用的是顯性浮水印，直接顯



示在數位典藏影像頁面上；隱性的浮水印是可以  
用軟體檢測，為避免有人改造浮水印，或註冊先  
後問題，可能必須成立認證機制及認證代碼。

在 Metadata 上有一個對權限之著錄，最好可  
以加以著錄。

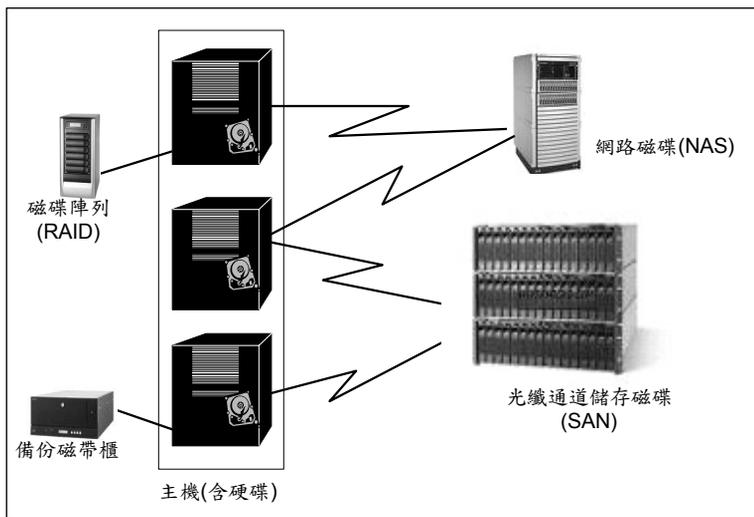
未來國中圖針對大量資訊儲存議題，也開始  
規劃如圖六的資訊生命週期儲存架構，以資訊的  
取用頻率，自動分級到不同等級、成本的儲存設  
備中，降低整體的儲存成本。因此大量極少使用

## 八、資訊儲存議題

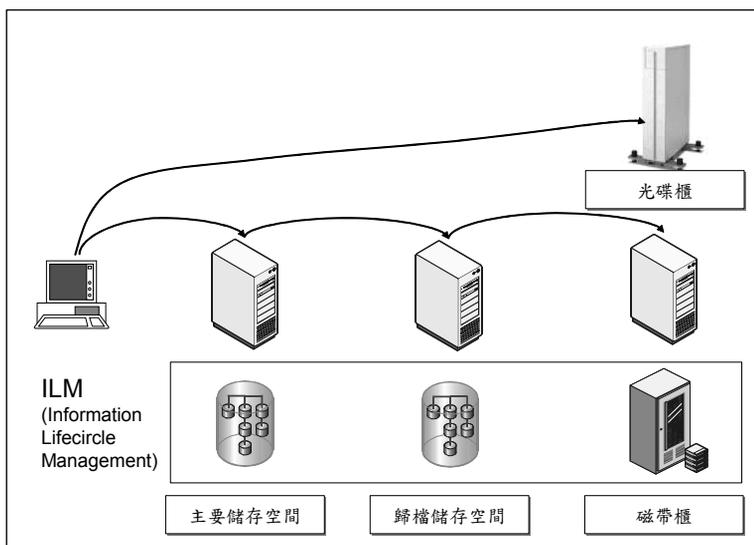
由於數位檔案大量擴  
增，必須尋求更低成本之儲  
存方案，並且也要因應高汰  
換率的資訊設備及維修問  
題。圖五所示的資訊儲存相  
關設備及架構，說明如下：

1. 主機硬碟，一般僅作為作  
業系統、數典系統運作之  
用，其擴充性有限。
2. 磁碟陣列 (RAID) 技術，  
可避免硬碟毀損所造成資  
料的損壞，可應用在主機  
硬碟、NAS、SAN 等儲存  
設備中。
3. NAS (Network Attached  
Storage)：作為網路儲存專  
用之儲存伺服器，可供多  
台主機分享。價格較便宜。
4. SAN (Storage Area Net-  
work)：存儲區域網路，主  
機透過光纖網路連結光纖  
交換器，再到儲存設備，  
傳輸速度快，但整體架設  
較昂貴。

圖五、資訊儲存設備及架構



圖六、未來儲存架構



之檔案可歸檔至低價儲存設備、具無限擴充性、可分批汰換設備

## 九、數位檔案長久保存方案

面對數位檔案的長久保存議題，我們可以先假設一個議題：50 年後，還可以看到現在建置的數位檔案嗎？目前仍是個未知數，只能就可預先努力的項目，加以準備。

1. 擬訂長久保存政策。
2. 異地備援：與其他不同地區的單位合作。
3. 儲存媒體定期轉換或重製，例如：整套 DVD 定期重製，數位影音檔定期轉換格式。

## 十一、結語

最後，我們再回想一開始所談的「為什麼要數位化？我們需不需要數位化？」數位化是一項高製作成本、高人力投入，以及高維持成本的工作，因此如果決定要進行數位化，目前網路上已經有不少資源可以取得。若有機會，也可以參觀其他機構，甚至廠商進行數位化的過程，以交換數位化作業之經驗，相信只要用心投入，數位化工作可以事半功倍，達到典藏與利用的目標。

【編者按】本文為錄音謄稿，經講者撥冗審閱刊載。

【訊息】

### 數位典藏技術規範參考資源： 數位典藏技術彙編

數位典藏技術彙編是國家型計畫的各種數位化技術規格之彙編，提供相關的經驗傳承及共享的參考，逐步建構屬於國家型計畫的相關技術規範，對於數位典藏計畫工作具推廣效益。

「數位典藏國家型科技計畫」技術彙編 [http://www.ndap.org.tw/2\\_techreport/enter.php](http://www.ndap.org.tw/2_techreport/enter.php)

