



# 中文期刊報紙典藏數位化作業概述

林淑芬 國家圖書館閱覽組編輯

## 一、計畫簡介

期刊報紙是一般訊息傳遞、研究發展之成果發表最快速的園地，也是記錄當代人文、社會、科技等活動最重要的媒體，讀者使用量極高，但保存不易，使用不便，因此極需數位化予以改善。

國家圖書館（以下簡稱本館）因根據圖書館法，具有接受國內出版品的法定職責，館藏報刊資料極為豐富。民國 70 年起，本館就已著手規劃中文期刊資訊系統之建置，而「中華民國期刊論文索引及影像系統」自從紙本升級到光碟，以至線上資料庫以來，一直都深受國人的歡迎。

民國 91 年本館以期刊及報紙館藏參加由國科會主持的國家典藏數位化計畫，預定在 5 年內進行館藏臺灣地區發行期刊約 1,000 種、報紙約 30 種的數位化工作。主要目的在以數位方式保存報刊資料，充實國家圖書館期刊報紙影像資料庫，便利民眾取閱，建立學術文化資產的永久典藏，並期能藉助資訊網路科技的進步，結合數位版權管理以及電子商務機制，提供使用者方便快捷的文獻傳遞服務，達成無牆圖書館的理想。

該計畫 93 年度的工作目標是加強報刊掃描的工作數量，預計進行館藏 200 種期刊與 6 種報紙（含微縮片轉製）掃描，共約 140 萬頁期刊及報紙的全文影像數位化。主要工作項目包括：期刊典藏影像數位化、報紙典藏數位化、期刊篇目詮釋資料分析建檔等三項作業。

## 二、數位化工作流程說明

該計畫的數位化作業，大致依照下列六項步驟進行，依序分別為：（一）數位典藏前置作業；（二）詮釋資料分析與著錄；（三）報刊資料數位化，（四）數位化資料的儲存與管理；（五）智慧財產權管理；（六）數位化資料的利用。茲分別介

紹如次。

### （一）數位典藏前置作業

本館依據圖書館法規定，典藏國內之出版報刊資料，目前報刊種數已逾一萬餘種，因此館藏藏品在進行數位化作業之前，首先需對預定數位化的藏品內容、範圍、類別、數量、優先順序等作一整體的規劃方案。

有了整體規劃方案之後，再依據當年的經費、人力配備以及儲存設備等條件進行年度計畫的草擬，年度計畫完成後，即可著手進行數位化的前置準備作業，以利工作的順利開展。這些前置作業包括：訂定標準與規範、選擇數位化物件、確定數位化格式、建立工作程序、確立施作方式等。

#### 1. 規劃年度計畫

本計畫的數位物件為本館館藏的期刊及報紙，因此在年度計畫規劃中要考慮的因素，除了當年的經費預算及人力配備外，還要對館藏藏品的狀況詳加考量，例如：期刊報紙的館藏數量、保存狀況、破損狀況、紙質、印刷品質、原件尺寸、複本數量、缺期缺頁狀況以及裝訂方式等；另外，部分藏品除原件外，另有微捲形式的藏品，以原件直接進行數位化，或是以微捲進行數位影像轉製，也是必須考量的項目。由於本館藏品平日仍需提供到館讀者的調閱，因此，如何將藏品分批進行數位化，以減少讀者無法調閱館藏原件的不便，也是考慮要素之一。

根據上列因素詳加研擬之後的年度計畫，應明確提列年度擬數位化的藏品、範圍、項目、方式以及預定數量，如此計畫方能順利執行。

#### 2. 訂定標準與規範

年度計畫確定之後，即可展開各項前置作業。首先，必須訂定嚴謹明確的標準與規範，



以便作業過程中有所依據，這是絕不可省略的工作。

本館為進行報刊數位化的相關計畫，期刊部門特邀請館外圖書資訊界專家，協同館內各組同人（特藏組、資訊組、參考組、政府出版品部門），成立「文獻分析機讀格式計畫小組」，共同修訂期刊文獻資源建檔之詮釋資料格式；另外並參與館內之「自動化及網路諮詢委員會」、「中國機讀編目格式計畫小組會議」、「詮釋資料 (Metadata)」格式標準研訂小組會議、「資料數位化標準研訂小組會議」、「分散檢索標準研訂小組會議」等多項會議，共同訂定數位化作業的相關標準與規範。

至於有關期刊篇目的主題類別與主題詞、關鍵詞等的選用與著錄，則直接採用圖書資訊界向來沿用之「中國圖書分類表」及「中文圖書標題表」等規範工具。

另外，由於報刊係屬連續發行的出版品型式，進行數位化時，出版日期、卷期以及頁碼的編碼方式，也必須明確訂規則，以免日後連結詮釋資料時產生問題。因此，本館同人特針對期刊以及報紙的卷期特性，訂定編碼原則各一份，作為數位影像檔案編碼的依據。

本館訂定的標準與規範是本計畫數位化作業人員在執行時所遵循的準繩，可以幫助各項作業有所依據，各項規範包括：

- (1) 詮釋資料 (Metadata)：文獻分析機讀格式。  
(臺北市：國家圖書館，民 92 年)
- (2) 產出規格
  - A. 掃描影像檔：TIFF 檔。
  - B. 影像製作規格：300dpi、G4 壓縮、黑白。  
在原稿清楚的前提下，影像掃描品質力求清晰，內容完整。部分藝術類雜誌掃描成全彩影像，早期報紙掃描成灰階或彩色影像。
  - C. 微縮捲片：35mm。
- (3) 作業規範
  - A. 資料數位化標準-檔案數位化與命名原則。
  - B. 國家圖書館期刊影像編碼原則。
  - C. 國家圖書館報紙影像編碼原則。
  - D. 分類規範工具。

(a) 中國圖書分類法 (試用本) (臺北市：國立中央圖書館，民 68)

(b) 中國圖書分類法 (增訂八版) (臺北市：文華圖書館管理資訊公司，民 92，第八版)

#### E. 主題詞規範工具。

(a) 中文圖書標題表 (臺北市：國立中央圖書館，民 84)

(b) 中文主題詞表·人文社會科學類 (臺北市：國家圖書館，民 92)

### 3. 選擇數位化物件

確立檔案格式之後，即可進行選擇數位化物件的作業，由於館藏期刊甚多，數位化之優先次序為：(1) 收錄於「中華民國期刊論文索引及影像系統」之學術性刊物；(2) 民國 50 年以前發行之臺灣地區出版品；(3) 館藏無復本者；(4) 使用量大且破損嚴重者；(5) 絕版而難以購得者。

至於報紙方面，則因使用量極大，且紙質脆弱，應即時進行數位化，是以自民國 90 年起，決定將館藏臺灣地區發行中英文報紙約 30 種，全部皆進行數位化，惟近來已有多家國內報社皆已開始發售其報紙光碟，因此，本館數位化的報紙種數，則以不與業界商品重複為原則，凡能價購者，皆不進行數位掃描，以節省資源。

### 4. 確立檔案格式

為能提供數位物件的檢索，本館決定採用「文獻分析機讀格式」作為期刊篇目文獻的詮釋資料建檔依據，至於報紙部分，因市面上已有報紙標題資料庫，因此本館決定不進行報紙的詮釋資料建檔。

另外，依據媒體資料的不同，所制定的檔案格式亦不相同，本計畫的數位物件既是期刊及報紙，屬於文獻類藏品，首要要求為文獻內容的清晰可讀以及永久典藏，因此，本館依據「資料數位化與命名原則」(臺北市：國家圖書館，民 92 年) 之建議規格，決定採用文字檔及影像檔資料永久保存格式進行數位化，至於網路下載以及預覽影像規格，則因限於經費，暫不進行。(參閱表一：數位化檔案格式)



表一 數位化檔案格式

檔案格式	建議規格	說明
文字檔		
資料永久保存格式	檔案格式：TIFF 壓縮：不壓縮 色調深度：黑白；灰階—每像素 8-bits；彩色—每像素 24-bits 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）	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
影像檔		
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。

5. 建立工作程序

本館期刊之數位化採「期刊篇目文獻詮釋資料之分析與著錄」與「報刊資料的數位化掃描作業」兩部分分別進行，其相關工作流程請參考圖一。

6. 決定施作方式

在確定當年的數位化計畫後，緊接著則可依據工作內容項目，逐一檢視每一項目的施作方式，確定是自行製作或委外作業。以近幾年狀況而言，若經費足夠，通常會採委外作業招商。當然，若決定採委外招商，則需著手訂定報刊數位化影像掃描作業招標規範。

在委外作業的招標規範中，除了需明確提列數位化物件項目、數量、工作地點、工作期間之外，亦需提供明確的數位化作業規範、作業程序、品質檢驗規範等，務使承包商確實據以遵行。

(二) 詮釋資料分析與著錄

1. 準備資料庫建檔系統

準備一套可以處理期刊文獻詮釋資料的資料庫系統是一件重要的工作，本館早年即已著手進行期刊論文索引資料的建檔工作，因此特委託業者開發建置一套詮釋資料建檔系統，本計畫的期刊篇目文獻詮釋資料亦在同一資料庫

內持續進行。

2. 確定資料庫收錄範圍

(1) 確認收錄對象及範疇

本館期刊館藏收藏甚豐，無法負荷將全數刊物篇目予以分析建檔，因此，特從收錄的刊物中，挑選臺灣地區出版發行、內容具學術性質者，進行分析著錄，收錄期刊中不包括一般消息報導及活動性質者。中華民國期刊論文索引資料庫內總收錄約有 3,000 餘種刊物，目前仍持續出版發行者約有 1,000 餘種。新創刊物之收錄係由本館期刊部門論文索引刊物選編小組同人負責。

(2) 挑選收錄之篇目

收錄之新刊物經確定後，為求時效，通常於刊物發行並寄達本館，且經登錄處理後，交由期刊論文索引主編同人進行各卷期收錄篇目之挑選。經收錄後，則需在資料庫系統建立書刊的基本資料，註記新到卷期登錄，若有拖期嚴重者，亦需進行書刊催缺。另外，對於設有網站之刊物，本館亦予以註記，並作連結，以提供使用者服務。

(3) 詮釋資料分析、建檔及審核

之後接著進行詮釋資料建立的步驟。凡經決定收錄的期刊篇章，都依據「文獻分析



機讀格式」、「中國圖書分類法」及「中文圖書標題表」等規範工具，針對各篇章進行分析與著錄，著錄項目包含：篇名、著者、刊名、卷期、出版年月、起迄頁次、類別、關鍵詞等。

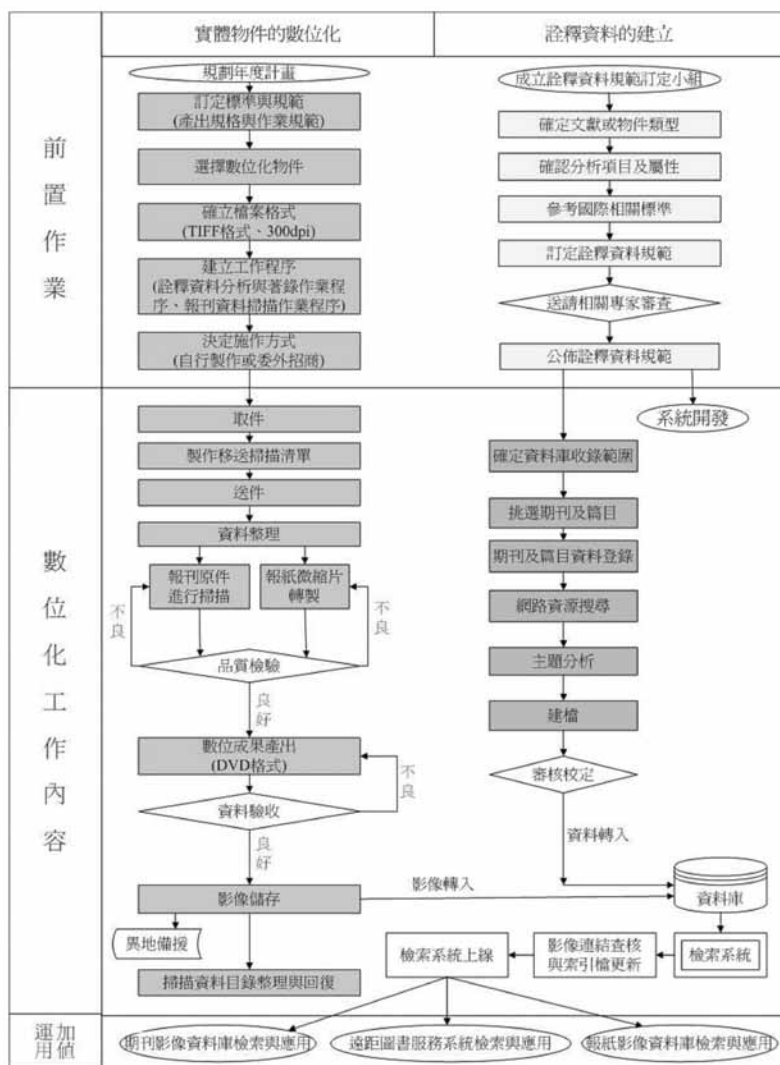
建檔工作完成後，需交付同人進行審核及校正，經總校完畢的詮釋資料，即可收錄於「中華民國期刊論文索引及影像系統」，並於隔日系統更新索引後，隨即提供使用者線上查詢。

本項工作限於人力，部分作業（如：詮釋資料之欄位分析、分類、建檔等項目）係採委外招商駐館服務處理，本館同人則負責審核及總校正之品質管控工作。

(三) 報刊資料數位化作業

依照本館年度數位掃描規劃，分期展開館藏期刊及報紙的數位化作業。本項工作，本館係採委外招商進行。承包廠商需依照本館制定之招標規範，逐一完成委託項目。

本項作業依照工作流程順序，包括以下幾個步



圖一 報刊資料數位化影像掃描流程圖

【館務報導】





驟：取件、製作掃描清單、送件、資料整理、進行掃描、品質檢驗、數位成果產出、資料驗收、影像儲存以及索引更新等。茲簡要說明如下：

### 1. 取件

#### (1) 架上調書

本館工作人員依據年度計畫中擬定的工作項目，擬製「待掃描報刊清單」，並從期刊書庫書架上直接調出欲進行掃描數位化作業的期刊、報紙或微縮片。

#### (2) 複本徵集與補缺

取件的同時，應對館藏狀況進行註記，基本上數位化的刊期，以不影響到館讀者調閱為原則，因此，數位化物件盡量以有複本者優先。若發現期刊缺期或沒有複本者，應隨即註記，以便通知館內採訪同人再作徵補或催缺；若該刊物卷期已絕版或無法取得，則應與他館進行館際合作，盡可能在數位化執行期間取得該件期刊或報紙。

#### (3) 核對「已移送／已掃描目錄」

自書架上取件後，接著應核對「已移送／已掃描目錄」，以確認該卷期是否已經數位化，若已經移送或數位化者則隨即歸架，以避免重複；若確認尚未移送者，即可準備移送數位化，並在書架中加註「XX 刊物 X 卷 X 期移送掃描中」之標記。

### 2. 製作移送掃描清單

準備移送數位化的物件，應建立移送掃描清單，清單中應載明以下各項目：系統識別號、刊名、卷期、出版日期等。為避免錯誤，應再核對「中華民國期刊論文索引資料庫」中之卷期、出版日期、系統識別號等項目。務求紀錄與系統一致，避免日後與詮釋資料連結時的錯誤。

為不影響到館讀者之借閱，數位化掃描期刊一般分成：批次輪轉式送掃（可送館外處理）、每週平台送掃、每日平台送掃等多種不同週期類別。除批次輪轉式外，平台掃描大都須於館內進行。以平台處理者多為館藏無複本或藏品較為珍貴，不宜批次移送館外處理的刊物。

清單項目經核對無誤後，即可依據不同週

期類別製作各種「移送掃描清單」。

### 3. 送件

根據之前完成正確無誤之「移送掃描清單」，由期刊或報紙負責同人將該清單及物件移送委外廠商，準備進行影像掃描數位化作業。

### 4. 資料整理

委外廠商於接獲「移送掃描清單」及數位化物件後，應派員赴本館先進行移送數位掃描報刊資料整理的工作，工作內容包括：核對刊名、卷期、頁數、逐頁檢查有否缺頁、髒污或頁碼異常等，嚴重異常狀況（如嚴重缺頁）應即時通報本館同人處理。

報刊資料檢查完畢後，即可打包裝箱，準備進行館內或移送館外的數位掃描工作。

另外，若有書冊過厚的報刊，本館同人會先行註記「可裁切」或「可拆裝訂」字樣，承包廠商工作人員可以在館內進行書冊裁切或裝訂拆解，以便進行掃描。

### 5. 進行掃描

此項作業由委外廠商工作人員依照招標規範與契約書之工作項目、作業規範與作業程序在館內或館外進行數位掃描。

### 6. 報紙微縮片轉製

本館早期報紙因都已拍攝微捲，因此部分報紙的數位化作業，係以微縮捲片直接進行數位影像轉製。此項作業亦採委外招商處理，承包商應依照招標規範與契約書之工作項目、作業規範與作業程序進行掃描數位化。

微縮片轉製數位影像的工作流程，與一般原件直接掃描相同。

### 7. 品質檢驗

本館依據招標規範與契約書要求，進行品質檢驗。委外廠商於影像數位化後，應隨即進行影像檔的品質校正，逐頁檢查掃描後的文件是否有太濃、太淡、黑邊污點、歪斜或是影像內容不清楚等情形發生。若檢驗出品質不佳的數位影像檔，則必須進行修改至合乎品質規格要求。

### 8. 數位成果產出

委外廠商將經過品質檢驗的數位化影像檔案，以數位光碟片（DVD）至少二套，連同各式



「掃描清單」，交付本館申請辦理資料驗收。

#### 9. 資料驗收

本項工作由本館工作人員按月以整批檢驗或抽驗等方式，依據「期刊影像掃描檔案編碼原則」、「報紙影像編碼原則」、「期刊影像掃描暨報紙微縮資料轉製影像規格」、「掃描期刊回復裝訂規格」等，併同廠商交付之各式交付清單及掃描期刊問題整理紀錄單，配合合約之規格要求以及項目，對委外廠商製作完成之數位成品，進行影像品質之驗收工作。所製作之光碟片與微縮捲片若有品質不佳或編碼錯誤等情形發生，皆依招標規範及契約書處理。

至於抽驗的樣本數值檢驗抽樣程序及抽樣則是依據中國國家標準（CNS2779 Z4006，計數值抽樣程序及抽樣表，頁9-10）的表 I-樣本大小代字，以及表 II-A--正常檢驗單次抽樣計畫（主抽樣表）為準則；另外，在品質方面的檢驗，國圖並編製「影像品質驗收基準範例」作為參考依據。

#### 10. 影像儲存

本項作業係由本館同人於館內進行。將通過驗收之光碟片等數位成果，逐片轉入影像資料庫。待所有的成果皆完成驗收點交後，光碟片可以準備移交資訊組予以典藏。目前典藏設備為光碟櫃，為便於提供服務，已另行購置影像伺服器儲存數位化成果，並複製一套光碟片儲存於和平東路科技大樓資訊圖書館，以達異地備援之目的。

#### 11. 影像連結查核與索引檔更新

影像資料轉入系統後，經由系統定期自動更新影像索引檔，並與詮釋資料進行連結註記該篇文獻「☑」（已掃描），以上步驟執行完畢後，即可提供使用者線上利用。

#### （四）智慧財產權管理

本項作業係將已完成的數位化成品，進行使用權限的管理，以便提供各種不同的應用。依據著作權法的規定，國家圖書因是典藏單位，並非期刊文獻的著作權利人，因此，數位化成果，僅能提供典藏之用，若欲提供網路服務的加值與利用，則須獲得著作權利人的授權同意才能進行。

目前該計畫之期刊篇目文獻，凡已獲得授權同

意者，將註記於館內的數位版權管理系統，授權許可分有償及無償授權兩種，分別以「○」（無償）及「☑」（有償）兩種代號作為提示，凡已取得授權同意者，使用者可以在網路上瀏覽或列印該篇文獻。若尚未取得授權者，則仍無法提供線上服務。目前國圖正積極推動爭取著作權利人授權的相關計畫。

另有關數位成品的加密保護處理，本館雖尚未進行，但在提供線上服務時，係以特別設計的瀏覽軟體進行相關保護及管制，目前成效良好。

#### （五）數位化資料的利用

該計畫之數位成果已全部轉入本館相關系統中提供服務，其中「中華民國期刊論文索引及影像系統」之營運，係由本館以委外經營方式，併同館內其他幾項重要資料庫，推出「遠距圖書服務系統」供國人使用。該系統結合電子商務，提供期刊論文索引資料庫以及期刊文獻全文影像，可提供未能到館讀者線上檢索以及文獻傳遞申請服務，甚獲使用者好評。

目前該計畫數位成果分別呈現於以下三種資料庫，歡迎各界多多利用：

1. 期刊影像資料庫檢索與應用：  
<http://readopac.ncl.edu.tw/periodical/hypage.cgi>
2. 報紙影像資料庫檢索與應用：  
[http://192.192.58.101/cgi/ncl9/m\\_ncl9\\_news](http://192.192.58.101/cgi/ncl9/m_ncl9_news)
3. 遠距圖書服務系統檢索與應用：<http://readopac.ncl.edu.tw/>

後記：本文原係「數位典藏國家型科技計畫內容發展分項計畫」製作之系列報導之一，圖片拍攝及圖文編輯承蒙新聞主題小組助理汪怡慧、陳美智、曾欣怡、呂俊毅等多位助理協助，謹致謝忱。