



黨史檔案數位典藏作業及其數位人文 工具應用經驗

林巧敏*  王立勛** 

【摘要】

中國國民黨度藏之檔案資料為海內外研究中國近代史重要資料來源，2018 年政治大學與中國國民黨簽署「黨史資料委託管理合作協議書」，將此批史料移至政治大學社會資料中心典藏。政治大學社資中心接受黨史檔案委託代管後，著手對於此批檔案之整理與數位化工作。本文介紹黨史檔案掃描與建檔作業原則，並提供運用數位人文工具協助自動分類之經驗，期許可提供其他典藏機構進行檔案數位化及內容分析工作之參考。

關鍵詞

檔案數位化 數位人文 數位典藏 檔案應用 自動分類 社團檔案

壹、前言

中國國民黨自孫中山先生於 1894 年創立興中會迄今，已歷經超過 120

* 國立政治大學圖書資訊與檔案學研究所教授兼社會科學資料中心校史與檔案組組長

ORCID 0000-0002-9309-9884

通訊作者 E-mail: cmlin@nccu.edu.tw

** 國立政治大學社會科學資料中心校史與檔案組組員

ORCID 0000-0002-6042-9218

E-mail: matt@nccu.edu.tw

年的歷史，為近代中國歷史最久、執政時間最長的一個政黨。記錄中國國民黨黨務發展過程之檔案資料，自然成為研究中華民國史或中國近代歷史重要之史料，學者有稱「如果沒有國民黨的研究作基礎，整個民國史的研究就不可能全面、系統、深化」（劉健清、王家典、徐梁伯，1992）。史學研究有賴於充分的史料，中國國民黨黨史館典藏之檔案資料為海內外研究中國近代史重要的參考資料來源（劉維開，1996）。

國立政治大學社會科學資料中心（簡稱社資中心）成立於 1961 年 10 月，為國立政治大學（簡稱政大）校內一級單位，在多年經營下，已成為典藏近現代史資料之重鎮，2012 年中國國民黨面臨空間不足問題，優先與政大社資中心合作，簽訂「孫中山紀念圖書館委託管理合作協議書」。2017 年 10 月中國國民黨對於黨史館經營政策有所轉向，因有孫中山紀念圖書館託管前例，在既有經驗下，黨史檔案合作協議醞釀成形，社資中心啟動相關前置作業，陸續盤點校內可使用空間，評估珍貴黨史資料託管可行性。

2018 年 3 月政大與中國國民黨正式簽署「黨史資料委託管理合作協議書」，2019 年 2 月 25 日社資中心資料組改組為校史與檔案組，掌理校史資料、孫中山紀念圖書館與黨史檔案之經營，組織任務從原本偏重於出版資料的匯聚，逐步著眼至原始檔案的管理與應用，亦成為我國大學中少數同時掌理圖書、校史、檔案之單位，兼具大學檔案館及歷史檔案館的角色與功能。

本文介紹此批黨史檔案委託代管背景以及政大社資中心對於黨史檔案整編、數位化與數位人文應用之作業經驗，期許可提供其他典藏機構進行實務工作之參考。

貳、黨史檔案代管作業規劃

有鑑於中國國民黨黨史檔案數量浩繁，檔案實體脆弱需妥善典藏，事先規劃典藏環境與搬運整理工作，相當重要。除了必須盤點既有資源與經費、擬妥各類作業程序、人員專業培訓外，同時與託管單位充分溝通，建立互信機制，安排適當交接程序，以滿足託管單位委託目的。

社資中心結合政大專業師資發揮高等教育機構學術優勢，並爭取校方計畫型人力協助檔案整理與數位化作業，讓此批檔案在原有黨史館建立的

基礎上發揮加乘效果，使檔案管理作業步上軌道。

一、社資中心典藏檔案資源條件

綜觀國內研究機構，由社資中心進行黨史檔案之典藏與管理，具有下列條件：

(一) 學術機構專業人才

政大擁有圖書資訊與檔案學研究領域學者專家，亦培育出眾多檔案學專業人力參與實務工作，校內亦有史學研究者共同關注，整體行政團隊為提供完善檔案品質而努力。

(二) 學術社群研究需求

黨史檔案資料內容主要使用族群為近代史研究學者，政大在近代史研究領域卓有聲譽，亦為政大核心發展學科，此批黨史檔案資源存於政大，學術資源與學者亦匯聚於此，不僅可提供研究社群便於利用，降低研究奔波所耗損之時間成本，未來亦可發展各種學術研究合作之可行性，達成託管單位（中國國民黨）、受託管方（社資中心）、研究社群（政大各教學研究單位）三贏局面，並增進史料資源之交流，刺激研究動能。

(三) 建置典藏空間設備

社資中心為因應黨史檔案入藏，整繕空間改造為擁有溫濕度控制、防火時效 1 小時以上之標準庫房，配置門禁保全系統、錄影監視、通訊設備等，購置鋼製檔案架，妥適安置格局，發揮空間最大效能，同時規劃設置檔案調閱使用空間，提供檔案應用之條件。

數位化工作室則購置各類掃描設備，視處理檔案之性質與需求，採選適當設備（表 1），以提升檔案數位化工作品質。檔案工作室備妥各類修護器具，如除塵台、無塵室級吸塵器、整燙器、托裱設備、無酸耗材等，在數位化同時亦一併進行檔案原件之整理維護。

表 1

購置檔案掃描設備說明表

掃描設備名稱	數量	特色	適用時機
Bookeye V1	2 台	1.平台非接觸式。 2.歪斜校正。 3.可去除指印。 4.V 型書稿台。	1.珍貴、紙質脆弱檔案。 2.裝訂成冊式檔案。
Microtek LS-3700	1 台	1.掃描範圍可達 A1 尺寸。 2.亦適用於畫作、木雕板等各類實體物件。	用於掃描大型物件，如檔案內含之工程圖面。
Avision AV5400	1 台	1.自動饋紙器。 2.自動雙面掃描。	1.紙質較堅韌，可承受饋紙滾輪。 2.雙面內容檔案。

二、檔案內容數位化構想

數位化是指經過「輸入」、「處理」、「儲存」與「輸出」四大步驟所形成，是影像光線的類比信號經由感光電子元件取樣後，所記錄的數位影像信號，此「輸入」過程被稱為類比／數位（analog to digital）轉換，是轉換類比信號成為數位信號的過程，最基本觀念就是取樣後形成的「畫素」資料，也稱為點陣圖數位影像；而「處理」的過程中，須依不同使用目的進行轉檔，亦包含後設資料（metadata）之建置，以清楚敘述所有檔案內容的來源；「儲存」則需具備永久保存的策略，並有遠見的考量未來使用之趨勢；「輸出」在螢幕的播放亦為一種呈現方式，而印表機則是常用的方法，輸出時應透過校正程序與色彩管理系統，使輸出之成品與輸入之原始物件達到相同之呈現（徐明景，2011）。

資料經數位化後產出數位內容的目的，其一為文化保存，因相較傳統原件典藏，透過數位化方式，可增添不同儲存媒介的保存方式，分散文化保存風險。其二為學術研究，經數位化後之研究資源可提供網路檢索，使資源更易獲取與傳播，改善學術研究工具。其三是為教育學習，教師可應用於教學實務上，更可對不限定對象之多數使用者開放，可達到提升自主

學習、拓展社會教育等目的。其四可便於加值應用，經數位化後可賦予資源創新內涵，並加值應用於各領域，如多媒體內容、文創產品、藝文活動、展覽推廣等。

檔案管理機構進行檔案數位化作業，會關注數位化之後的應用，數位化目的之一在於避免調閱原件的損耗，加速內容的開放應用。因此，評估數位化檔案內容的優先順序時，除考慮檔案價值、保存風險以及使用者需求等因素外，也必須考量數位化後能否開放應用問題。對於此批黨史檔案，委託協議已授權政大師生學術研究使用，加上檔案年代久遠、原件脆弱且屬唯一性資料，其內容深具研究價值，實有優先進行數位化工作之必要性。

近 20 年來，數以萬計的研究者享受數位化成果之便利性，大幅改變資料獲取、研究方式與產出途徑。針對此批黨史資料除了進行數位典藏（digital archives）作業之外，也針對開發檔案內容價值，進行檔案檢索系統更新以及資料內容探勘分析，希望能藉助資訊科技協助，突破既有研究框架，運用數位人文（digital humanities）工具進行資訊內容之開發，以促進使用者對於檔案內容資訊之加值運用。

參、黨史檔案數位典藏作業

黨史檔案陸續入藏社資中心後，數位典藏作業旋即開展，社資中心依循管理計畫之原則，按部就班的執行各項作業，過程中運用各類數位化設備與工具，在組織內設計各種檢核機制，不斷檢討、優化作業流程，分述各項工作重點與成果如下：

一、託管移交前期準備工作

由於黨史檔案數量龐雜，交由社資中心託管執行前，尚有大批重要檔案未經妥善整理維護，檔案內容缺乏完整編排描述，將造成檔案查找及調閱過程的困難。社資中心接受託管後首要執行目標在於全面性檢視檔案，分析託管方提供之數據，瞭解過往類似檔案管理作業之經驗，以決定優先數位化之檔案系列。在正式數位化前，前期準備工作如下：

（一）建置檔案典藏環境

檔案庫房須符合恆溫恆濕要求，具防火、防水、防蟲、門禁等設施，使檔案實體安全典藏其中。

（二）制定相關作業規範

為使黨史檔案數位典藏順利進行，社資中心已建置一套完整的工作規範與標準流程，迄今已制訂完成 5 種作業規範，包括「檔案業務專任助理工作注意事項」、「檔案庫房管控規範」、「檔案數位典藏前置作業規範」、「檔案數位化格式與作業規範」、「檔案描述規範」，以及 3 種作業表單，包括「入庫提調單」、「溫溼度紀錄表」、「檔案處理作業註記單」，均列為社資中心 ISO 程序文件，使檔案作業流程一致化，並維持作業品質（詳如附錄一至三）。

此外亦訂定附屬文件，如「新進檔案業務專任助理訓練計畫」、「檔案描述用字用語參考表」以因應實務工作所需。研訂中之規範，尚有檔案閱覽應用、檢索系統使用、業務統計等方面之作業要求，以求完善整體檔案管理作業制度。

（三）滾動式研討與標竿學習

託管黨史檔案為社資中心創新業務，作業過程中遇有瓶頸時即不斷研究相關資料、請益校內外專家學者，並召開工作小組討論，同時向外尋求標竿機構之典範參考，如參訪國史館在檔案數位化流程的著力，與歷史檔案內容解析應用之豐厚經驗，也參訪學習檔案管理局已建置之檔案實務作業指引，及其工作現場與庫房之嚴謹管理制度，汲取效法各單位之長，力求完備社資中心之檔案管理經驗與能量。

（四）規劃數位化作業流程

建立檔案數位化處理作業流程，依序規劃為前置作業、數位掃描、影像檢核、後設資料建檔等，其中各項流程又可細分為多個步驟，數位化工作要求作業人員需遵照嚴謹的作業規範，施行數位掃描工作，社資中心數

位典藏作業各步驟圖示說明如下 (圖 1)：

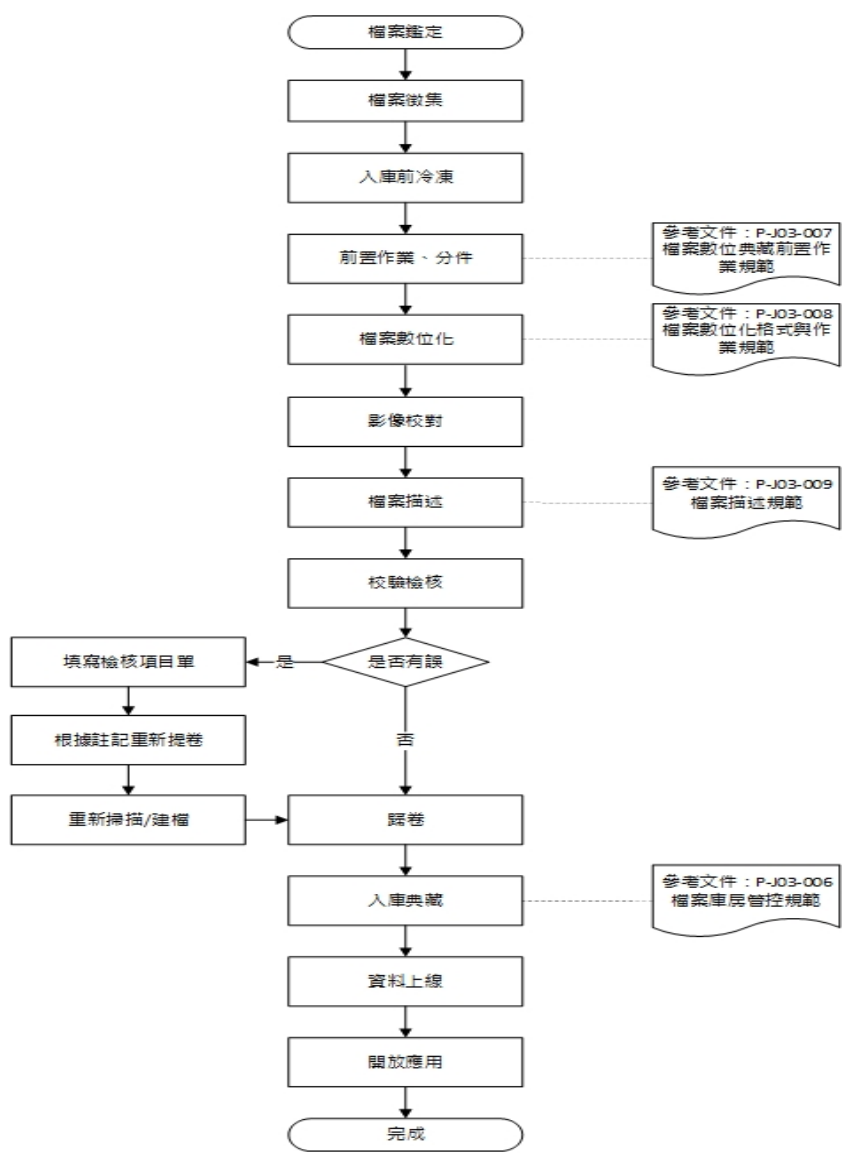


圖 1 檔案處理標準作業流程圖

二、數位化前置處理

前置作業規劃重點在於檔案徵集後，開始數位典藏之前，為了數位化作業而進行的所有必要措施與準備，以便後續正式進行數位化流程時，得以順暢無礙，提升作業品質與效率。針對此批黨史檔案進行之前置作業，包括：

（一）預除蟲菌

由於託管檔案入館前之環境不一，社資中心難以掌握，為避免檔案若有蟲菌，恐污染傳播至其他未汙損之檔案，入庫前進行除蟲滅菌可有效阻絕滋生，並維持新設檔案庫房之清潔。

社資中心目前採低溫法，分批將檔案裝入聚乙烯塑膠袋密封，再放入攝氏零下 20 度至 30 度之冷凍櫃達 7 日以上，待回復至室溫後取出檔案，去除塑膠袋，將檔案逐頁檢視，將乾化、粉粒化之雜質或蟲屍，以軟毛刷輕拭清除後，或上架典藏、或是置於工作區等待後續處理。

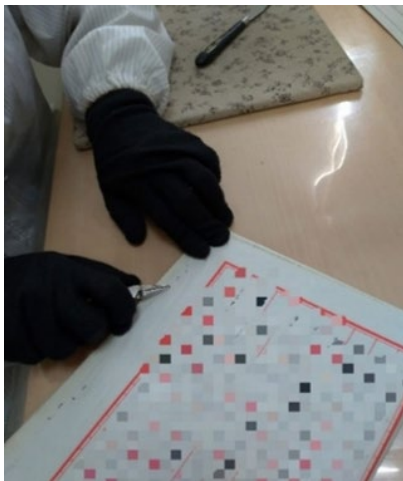
（二）拆、整卷

在不傷害檔案內容前提下，運用各類作業工具，如鉗子、鑷子將原始檔案裝訂之釘針與其他附加物去除，並將每頁紙張攤平，邊緣折角理平，如皺摺嚴重致使掃描影像之摺痕有遮蓋文字之虞者，需使用電熨斗襯紙低溫熨平，再依序整齊疊放。

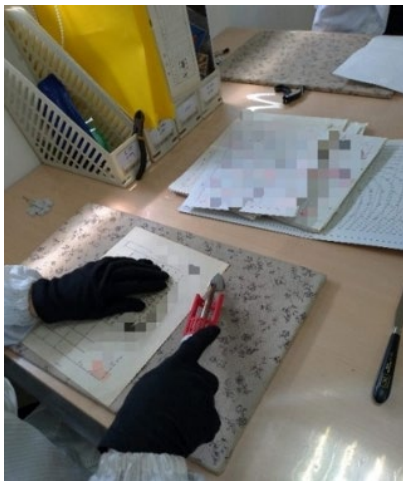
（三）分件編頁

由於拆卷後檔案恢復單頁散裝態樣，為維護檔案原始順序避免錯亂，應即逐頁編訂頁碼，編頁時檔案方向定位以檔案內容行文或閱讀方向為準，定位後於檔案正下方置中處，編寫頁碼，若該處已有內容可調整挪移位置。

若遇有浮簽、夾簽者須獨立編頁；對於非破壞式裝訂之書冊，如線、膠裝者，暫不拆卷，以原始形制編訂頁碼即可。



檔案拆卷編頁



檔案折邊熨平

圖 2 檔案拆卷整平作業

(四) 複本處理

社資中心將複本定義為內容、形制均一致，且重複超過三頁者。為使後續掃描階段降低重複作業成本，在前置階段需先行判斷需掃描之複本範圍，並於實體夾註標籤，以便作業，判斷標準如下：

1. 完整資訊原則

- (1) 檔案內容相同、原件差異為銜章之有無：掃描有銜章之原件。
- (2) 檔案內容相同、有不同批註，則掃描有批註之整份原件。
- (3) 內容互有不同、互有資訊增減者，則皆須掃描。
- (4) 不同檔案中的相同附件（例如表格），為保留檔案脈絡完整，附件皆須掃描。

2. 從優原則：若有多件複本，原則上掃描第一件，但可依實體狀態挑選較好的單件掃描。

（五）檔案保存狀況檢視

依據《國家檔案管理作業手冊》，將紙質類檔案受損程度區分等級如下（國家發展委員會檔案管理局，2020）：

1. 第一級：嚴重蟲蛀、腐朽、沾黏、脆化、粉化、紙材或媒材變質致內容字跡無法判讀，檔案無法翻閱或已形成檔案磚，屬嚴重受損之檔案。
2. 第二級：局部蟲蛀、破損、脆化、撕裂痕、褐斑或發霉，紙材或媒材變質致影響內容字跡判讀，檔案難以翻閱，屬中等受損之檔案。
3. 第三級：局部輕微破損、撕裂痕、泛黃、摺痕，檔案可翻閱，屬輕微受損之檔案。
4. 良好：檔案紙材及媒材狀況穩定，無明顯變色、撕裂痕、變形皺褶之檔案。

依據上述標準，逐卷檢視檔案保存狀況，並隨卷記載「檔案處理作業註記單」，註記檔案受損等級與範圍，若有危害長久典藏情況、檔案脆裂、散落等，需掇拾拼齊，並放入無酸膠套特別保護。

三、數位掃描作業

掃描機的成像原理為光線照射至原物件，透過感光元件將接收到的光訊號轉變為電訊號，經過類比／數位轉換器轉換成數位訊號資料（徐明景，2011），這樣的特性使得運用數位掃描得以快速產生數位化成品，但必須依據實體檔案的性質，妥適選用合宜的設備，避免檔案因掃描過程造成損害。數位掃描之操作方式，及其成品之影像規格、產出電子檔的命名方式、後設資料建檔等，均需有完善規範與流程控管，俾於維持作業一致性，以及後續開放加值等運用。

（一）掃描影像規格

影像規格隨著硬體技術發展，支援的標準不斷提升，除了考量硬體規格是否可支援外，訂定影像規格時亦須考量成本問題，例如儲存媒體的空間成本。可知影像規格之訂定，必須衡量品質、需求與成本，且考量未來

使用的前瞻性。因一旦訂定後，若有超出規格之需求時，即難以變更，不得以只好再次進行掃描作業，不但是耗費資源，亦對檔案實體造成傷害。

因此，社資中心參考近年主流規格（Besser, 2003／林彥宏譯，2009），編訂典藏級與閱覽級影像規格如下，目前閱覽級格式係因配合系統瀏覽界面設計採用 PDF 檔，未來如有其他使用需求，亦可根據典藏級檔案轉換為其他提供運用之電子檔格式。

表 2

社資中心處理黨史檔案影像規格表

規格項目	典藏級	閱覽級
檔案格式	TIFF	PDF
色調深度	全彩（RGB），24-bits	加入浮水印保護
壓縮比	不壓縮	適當壓縮
解析度	300DPI	檔案格式不分解析度
最小單位	一頁一檔	一件一檔

（二）數位檔命名方式

數位掃描後的影像檔需賦予唯一的命名編碼作為識別並避免重複，亦能提供影像檔與後設資料於檢索系統中相互對照，因此，需要訂定符合邏輯與彈性的命名標準，又因原有之系列號均以中文表示，如「省」、「環」、「會」等，因檢索系統儲存中文有內碼差異問題，電子檔命名原則宜採英文及數字為主，故將系列號以英文代碼代替原有中文系列號。根據完成處理之省黨部檔案為例，說明掃描數位檔命名方式如下：

1. 典藏級數位檔命名方式，依序為系列檔號代碼、卷號、副卷號、頁次號，因所有檔案為國民黨全宗，故不在檔名重複加上全宗號，影像以頁為儲存單元，故採頁次號。表列範例如下：

表 3

典藏級數位檔命名方式

以省黨部省 1/1 第 1 頁為例				
結構	系列檔號代碼	卷號、副卷號（預留百位）	頁次號（預留千位）	
範例	tw	省 1/1	0001	
tw_001_001_0001				
若該頁原件過大需分區掃描，切割之影幅命名方式為頁次號加上英文				
範例	tw_001_001_0001a、tw_001_001_0001b			

2. 閱覽級數位檔命名方式，依序為系列檔號代碼、卷號、副卷號、件次號，範例如下：

表 4

閱覽級數位檔命名方式

以省黨部省 1/1 第 1 件為例				
結構	系列檔號代碼	卷號、副卷號（預留百位）	件次號（預留百位）	
範例	tw	省 1/1	001	
tw001_001_001.pdf				

（三）掃描作業須知

掃描過程應由訓練有素之工作人員妥善進行，除對原件情況有一定認知外，亦須於掃描過程中注意以下事項：

1. 放置原稿前需仔細檢查掃描器承物玻璃，應保持潔淨而無灰塵和污跡。
2. 為保存原件完整樣貌，掃描之影像必須涵蓋原件周邊至少 0.2 公分以上，以利確認影像檔之完整性。
3. 如果原件包含浮貼，掃描時需分別掃描浮貼蓋著、浮貼掀開之影像，如果同一處有數張重疊之浮貼，需逐張掀開掃描。
4. 若檔案原件過薄導致透字，需裁切與原始檔案大小一致之襯紙加襯底後掃描，且影像不得外露襯紙。

5. 檔案襯貼或裱補之紙材、黏疊文件、書冊之四周陰影，在不影響影像完整性的原則下，可適當切除，以維閱讀明確度。
6. 檔案內容模糊時，以提高對比、解析度及銳利度方式為原則，將影像調整清晰。
7. 原件規格過大時（超過 A3），原則以大圖掃描器一次掃描完畢；無法一次掃描完畢時，始得分區掃描，分區掃描採由上至下，由左至右方式，以實際閱讀方向決定，且重疊二公分接合區域，以顯示其關聯。
8. 掃描過程應同時確認前置作業是否有缺漏之處，記載於檔案處理作業註記單，並退回修正。

（四）影像檢核修正

雖有標準化之作業流程，惟影像成果仍難免因作業過程疏漏產生瑕疵，在提供給使用者應用前，必須嚴格把關影像品質，避免應用端難以使用，以及衍生日後需要重新提調、掃描之作業成本。檢核過程採建檔人員建檔完成後交互逐筆檢驗，再由檢核人員整合抽驗，過程中若發現有以下瑕疵，必須重新修正：

1. 影像檔案之數量與原始文件頁數不相符。
2. 影像跳、漏頁，次序混亂等。
3. 影像歪斜超過 5 度，原稿即已歪斜者除外。
4. 影像不完整或留邊過大。
5. 折痕過於明顯，影響內容判讀者。
6. 透頁或影像模糊，影響內容判讀者。
7. 色階、亮度不足，影響內容判讀者。
8. 以上影響內容判讀者，若係原稿即有此現象則除外。

四、後設資料建檔

來自黨史館移交之檔案目錄已有基本之後設資料建置，惟其後設資料架構與名稱，與檔案描述特性略有出入，例如原目錄欄位「出版者」，應是指「產生者」，原目錄欄位「題名」所指應是「案件名」。為能提供後續檔

案檢索系統顯示欄位明確而易於理解，修正並重整後設資料欄位，分述各項著錄要點如下：

（一）案件名：依據分件後各案件脈絡與牽涉之單位，以及原件所示主旨，以照錄為原則，若原件主旨過於冗長、發散，可在不影響主旨前提下稍事精簡。

（二）產生者：紀錄檔案產生單位或個人。

（三）資料類型：因黨史檔案實體類型單純，多為紙質類文件，故將資料類型整合區分為公文書、簡報、日記、圖書、期刊、照片、電報、合約、其他。

（四）案卷名：案名之賦予，可參考照錄該案卷已有之名稱，其用語宜簡要明確，避免冗長累贅，並具涵蓋性。無案卷名稱，或原案卷名稱未臻完善者，依該案卷內各件內容所具共同性質、產生年代等，以揭示案情，並在最大涵蓋程度之範圍，賦予妥適名稱。

（五）出版日期：著錄檔案內產生案件之年、月、日，此欄為保留原國民黨著錄欄位值，有民國、年號或西元等不一致情形，保留原值備供查考。

（六）出版日期西元：新增著錄檔案案件之西元年、月、日，共計八碼，年月日間使用斜線「／」區隔，如為案卷層級，則著錄起迄。

（七）內容描述：針對案卷內容加強著錄，揭露該案卷之內容資訊。

（八）產生地：著錄起始案件之產生地，行政區層級以直轄市、省縣市為原則。

（九）相關人名資訊：本項目著錄檔案內容涉及之重要個人名稱，並以其於檔案中之名稱為依據。

（十）相關地名資訊：此產生地非屬於行政區劃之名稱，是指著錄檔案內容涉及之重要地理名稱，且可擴及至地標、建築物等，並以其於檔案中之名稱為依據。

（十一）數量：著錄檔案數量與計量單位，計量單位著錄以頁為原則。

（十二）外形描述：針對物件實體之形制、特徵、標示等進行描述。

（十三）語文：依據檔案主要行文語言著錄。

（十四）檔號：組成結構為系列號簡稱、案卷號、件次號，均預留百位數。

- (十五) 典藏單位：依據目前檔案典藏單位著錄。
- (十六) 製作單位：依據檔案數位化建置單位著錄。
- (十七) 備註：著錄有描述價值、便於管理檔案之事項。
- (十八) 拍攝備註：著錄該案件於案卷中之頁碼起訖。

前述「檔號」之組成，因過往黨史館檔案描述層級均以「案件」為基本描述單位，且原有檢索系統未有層級描述之架構設計，導致社資中心在受託管初期，後設資料建檔作業時，受限於既有框架沿用舊制，仍以「案件」層級建檔。但為解決此一問題，社資中心研究黨史館過往描述邏輯，對照實體案卷排架順序，推敲出檔案層級描述方案，於後設資料保留原檔號著錄，但進行檔號結構梳理對照，提供檢索系統批次轉檔時判別層級關係，提供瀏覽功能使用，以便社資中心新建置之檔案檢索系統可呈現反映檔案實體結構之目錄層級瀏覽功能。

經清點此批檔案各文件系列之檔號編碼方式相當不一致（如表 5），未來將配合檢索系統改版重新梳理檔號，將以往黨史檔案中各系列描述標準差異之情況進行整合修訂，將現行檔號多樣態之命名方式統一，並依據檔案描述編碼格式（Encoded Archival Description, EAD），研發適合館藏內涵特性之後設資料欄位，作為社資中心描述黨史檔案的後設資料標準。配合檢索系統改版擴充，檔案將不限以「案件」為基本描述層級，將建立描述政策指引，視不同檔案系列性質與研究檢索需求，可彈性選擇以系列、案卷或案件之層級描述，藉此於檢索系統中實踐階層瀏覽功能，呈現檔案資料之關聯性。

表 5

黨史檔案各系列檔號命名方式

檔案系列	原檔號命名方式	系列號 英文代碼	現行檔案編號範例
流水號無意義			
環龍路檔案	流水號	MN	環 00069
漢口檔案		HK	漢 0014
五部檔案		FD	部 0132

（續下表）

(接上表)

檔案系列	原檔號命名方式	系列號 英文代碼	現行檔案編號範例
監察檔案		CS	監 0003
狄膺史料		DY	膺 101
吳稚暉檔案		ZH	稚 01182
流水號有意義			
敵方廣播新聞紀要	檔號為出版卷次號	DG	敵廣 0051
有分卷，斜線前後均為案卷號			
特種檔案	案卷號、件次號	SP	特 1/2.1
中央秘書處檔案		CC	中 1/10.12
台灣省黨部		TW	省 1/1
有分卷，僅斜線前為案卷號			
政治檔案	案卷號／件次號	ZG	政 6/12
國防檔案		DC	防 005/0042
一般檔案		GA	一般 523/1
有年代號綜合型			
蔣中正總裁批簽檔案	年代號／件次號	ZP	總裁批簽 57/0002
副總裁批簽檔案		FZ	副總裁批簽 48/0001
大溪檔案黨務類		DD	大黨 049/009
特殊類型			
會議紀錄	屆次、會議別／流水號	CR	會 9.1/85
一般檔案－期刊	圖書分類號	GJ	一般期刊 559.405 6418

肆、黨史檔案之數位人文發展經驗

中國國民黨的黨史資料，雖是中國國民黨的重要文化資產，但早期對於檔案整理缺乏建制，加上幾經遷移，部分文件系列之案卷編排缺乏有序整理，僅依時間將文件排放置入卷夾，根據檔案產生時間上架典藏。通常檔案編排或依主題、或依分類，可協助將相同主題檔案加以類聚，有助於檔案典藏管理及檢索使用。

進行「文件主題分類」或簡稱「文件分類」是指依文件「內容主旨」給予「類別」(class or category)，文件分類的目的，在對於文件進行分門別類的加值處理，使得文件易於管理和利用(曾元顯，2002)。缺乏分類編排的檔案資料，不僅無法有序的組織檔案資訊，提供檢索查詢，也無法揭示內容發揮檔案價值。進行檔案分類編排如同圖書資訊的主題分類，是檔案技術服務工作的核心，能將關聯案情之檔案資料加以類聚，可提供後續檔案價值鑑定判斷、檔案檢索以及檔案編輯選件等內容應用工作之依據。但面對數量浩繁且未經整理的檔案，如果藉由人力重新編排入卷，不僅曠日廢時，且有違檔案編排的原始順序理論，致使多數未經有序整卷之檔案，僅能進行選擇性的作為，致使多數未經整理的檔案，尚無法提供檢索應用。

資訊技術的進步，或許能為檔案管理問題帶來改善的契機。數位工具的長項是能協助處理大量資料，降低人力負荷，尤其針對缺乏內容分類的檔案全宗(註 1)，既要符合全宗及原始順序原則，維持原有實體案卷編排順序，但又需要提供主題瀏覽查詢。在不改變實體原件編排順序下，可運用文本探勘技術，分析大量文本內容，建構文本內容主題分類架構，提供線上檢索系統查詢使用，並無須重新整編實體案卷，也不用抵觸多年來在檔案學理論是否堅守檔案編排原始順序之爭，有助於提升文件整編與檔案開放應用的效益。

文本探勘(text mining)是利用資訊技術，從數位資料挖掘有意義的資訊，運用的是分類(classification)與分群(clustering)、推估與預測(estimation and prediction)等演算理論與模型技術(項潔，2017)，可協助檔案管理領域以往單憑人工方式，建立主題類表和沉浸檔案文件間尋找脈絡的方法，可改以電腦機器學習方式，透過文本資料斷詞和探勘分析，找出文本內容主題概念並加以自動分群，能有效簡化檔案人員以往人工繁瑣又重複的工作。

在數位人文研究方興未艾之際，文本探勘工具的開發已有相當成果，可提供人文社會科學領域學者，不必鑽研技術學理即可快速運用開發之數位工具，完成資料統計與分析(曾元顯，2011)。人文學者也開始具備資訊素養能力，有資料處理的經驗或者習得相關程式語言知識，能夠知道數位工具的操作方式，並擁有運用數位工具融入自身研究的能力，能發展成一套步驟程序，提供本身專業領域解決問題之用(Liu, 2017)。過往之文本探

勘工具輔助主題分析的研究中，自動化主題分類尚無法堪稱完美，但能節省大量人工分類的時間，如能在機器出現錯誤時予以修正，仍然具有可行性（曾元顯，2014）。再者，進行人工智慧的自動分類判斷是目標，在此之前需要先找出檔案分類和內容判斷的決策因素，因此，需要嘗試建立檔案自動分類與價值判斷的分析過程，提供進階到自動化處理技術發展的依據，即使初時尚無法完全取代人工分類，但降低人力負擔還是大有可為。

有鑑於檔案主題分類之重要性，社資中心嘗試運用文本探勘工具，以總裁批簽檔案系列為對象，進行檔案自動分類之實證調查與分析過程，利用文本內容特徵詞彙為「總裁批簽檔案」系列建立自動主題歸類，以機器學習方式計算詞頻、關聯性，找出文本中隱含的資訊，建構此批檔案內容的知識架構，最後輔以專家訪談，修正檔案主題分類架構。此一運用數位工具導入檔案實務作業，期許可協助改善檔案編排無序而缺乏主題分類之問題。

由於資訊技術處理所關注的面向，往往與人文研究目標不完全相同，如果能導入人文學者研究中的專業詮釋知識，能讓資料探勘分析的結果，可以獲得更有理論背景的結論（劉昭麟，2018）。因此，自動分類結果加入人文學者協助判定，可回饋修正文本探勘的技術處理過程。實施程序如下：

一、自動化主題探勘

將文本經過斷詞處理，利用名為 CATAR（Content Analysis Toolkit for Academic Research）的內容探勘工具，對於檔案標題、內容摘要以及內容涉及人、事、物等含有主題概念欄位中詞彙，結合「相關係數」及「詞頻」排序詞彙的分析，自動擷取類別特徵詞，作為類別標題以及進行檔案文件的分群，最後再根據分群之特徵詞，初擬訂定適當之主題標目，以整理出此批檔案系列之主題架構。

二、專家意見調查

完成初步文本資料探勘後，邀請民國史學專家對於前述檔案文本探勘及自動主題歸類架構結果進行「訪談」，提供相關主題修正意見建議，以完成整體主題架構。最後依據訪談意見，並以民國史專家提供之重要概念詞

彙，輔以民國大事紀斷詞呈現的重要人物和事件概念，與文本探勘結果進行比對，以分析此批檔案與歷史價值之關聯，從而判定此批檔案的史料研究價值。

本測試研究工具採用曾元顯 (2011) 開發之 CATAR，是採概觀分析的自動主題分類方式，歸類結果會以「主題樹」(topic tree) 及「主題地圖」(topic map) 視覺化的方式呈現。主題樹是根據前述凝聚歸類與多階段歸類結果，配合類別描述詞自動擷取以及多面向資料交叉分析，將所有結果以網頁形式呈現，以便於檢視各個主題類群之關鍵詞、歸類層級、文章數量等資訊。CATAR 自動歸類依其門檻值設定不同，會得到不同歸類群組數，門檻值基本上有 0.0、0.05、0.07 及 0.1 供參考，本研究將 2,427 件檔案以不同門檻值進行檔案後設資料內容自動分群 (含案件名、案卷名、內容描述以及內容涉及人名、地名等有主題概念欄位)，接續輔以人工初步檢視主題歸類結果，為避免主題分類架構過於細瑣，反而不具參考與分析的意義，經與歷史學者討論後，選擇門檻值 0.05 之 13 個主題歸類結果較為貼近學科專家認知，類目架構如表 6 所示：

表 6

總裁批簽檔案系列自動分類之類目名稱與內容主題關鍵字

序號	類名	主題關鍵字
1	改造委員會	中改會、會議、次會議、通過、委員會
2	中央常務委員會	登記、中常、中常會、辦法、黨員
3	黨務工作檢討	檢討、報告、工作、意見、黨部 革命、研究院、實踐、同志、釣座
4	黨部工作單位考核	各單位、考核、成績、考評、設計
5	政策推動組織	秘書長、委員會、政策委員會、組主任、黨部 總支部、總裁、懇請、本黨、大會
6	全國代表大會	代表大會、黨部、教授、知識青年黨部、全國
7	大陸工作	大陸、張君勱、香港、中共、中國 港澳、加強、工作、擴大、中共
8	社會活動	政策、會議、次會議、鑒核 草案、復興、實施、標語、文化

(續下表)

(接上表)

序號	類名	主題關鍵字
9	海外活動	來台、活動、同志、本黨、工作 華僑、婦女、越南、服務、方案 澳門、出國、夫人、戰俘、緬甸 亞盟、代表團、擔任、召集人、谷正綱
10	人員遴選與指派	主委、黨部、遺缺、現任、擬請 召見、人選、名單、總裁、敬請
11	民意機構聯繫	立法院、候補委員、立法委員、立法委員黨部、典禮 行政院、院長、監察院、中華民國、副院長 國民大會、專案、專案小組、幹事、反共建國聯盟 臨時、國大、總統、國大黨、總統候選人
12	新聞宣傳	題詞、特刊、出版、僑報、計劃 美金、會談、主管、撥付、宣傳會談 日報、社論、時報、香港時報、台港
13	中央評議委員	中央委員、評議、評委、中評委、規定

接續進行專家意見調查，整理各小組主題專家瀏覽檔案內容相關性判斷意見，據以修正自動分類主題架構，修正完成之分類架構維持前述 13 類架構，並增補主題關鍵字，預計後續將根據此一分類架構作為分類號基礎，而分析產出之主題關鍵字將加入後設資料增加檢索點。本測試研究完成總裁批簽系列檔案所有檔案之內容分析，藉由斷詞分類所類聚之主題，呈現該系列檔案內容的核心概念；並藉由史料分析與專家訪談，完成檔案分類主題架構之調整。完成自動分類結果可提供兩項應用：

(一) 可將分類產生之主類及次類名稱，批次鍵入後設資料欄位，增加檔案描述欄位之主題檢索詞彙，提高檢索用語可查獲的結果。

(二) 經自動分類及人工修正之分類架構，未來可提供檔案檢索系統使用者介面以主題瀏覽方式，提供使用者利用此一虛擬主題樹，直接點選瀏覽相關之檔案文件。

伍、結語

檔案數位化目的在於珍視檔案原件保存，並可加速檔案內容傳播與應用。此批深具史料研究價值之黨史檔案，經委託程序入藏政大之後，政大以大學之學術能量進行資料專業整理與研究使用，充分發揮資料內容價值。在數位化工作推展之際，也利用自動分類之數位工具，進行檔案內容探勘分群處理，協助建構可供檔案檢索主題瀏覽之分類表，取代人力逐一瀏覽案件後，重新建構分類表的冗長作業時程，並可經由內容斷詞後產出有意義之主題詞彙，增加後設資料內容主題著錄來源，提高使用者檢索詞彙豐富度。此一實驗性研究也證實數位人文工具，可運用於提升檔案管理工作效率，增加檔案內容主題分析之成效。

本次檔案自動分類的測試研究目的，在於協助檔案管理以科學化之作業方式，深入探究檔案內容與價值，進而提供應用數位工具促進檔案實務工作之創新思維，與檔案學理捍衛之原始順序原則並不悖逆，因實體案卷編排如舊，僅是運用數位工具梳理檔案主題，進而建構主題詞表提供檢索瀏覽輔助並增加主題檢索詞彙，期許本實測研究經驗，可提供檔案典藏機構未來對於運用數位工具解決檔案主題判別、提升檔案分類作業效率之實務工作參考。

(接受日期：2021 年 3 月 25 日)

註釋

註 1：檔案全宗 (fonds) 是指一個機關、企業、事業、團體、組織或個人在社會活動中形成，而彼此具有歷史聯繫或邏輯聯繫的檔案文件整體。在中國歷史上會將相關事件的一組文件稱為卷宗，故源自法文 fonds 一詞，中譯成為「全宗」，其字面意義就是指全部的卷宗，是來自同一產生者社會活動記錄的關聯文件。全宗是檔案館對於檔案進行分類整理和所有管理活動的基本單位。

參考文獻

- Besser, H. (2009)。影像製作入門 (*Introduction to imaging*) (林彥宏譯)。臺北市：拓展臺灣數位典藏計畫。(原作 2003 年出版)。
- 徐明景 (2011)。數位化工作流程指南：影像資料。臺北市：拓展臺灣數位典藏計畫。
- 國家發展委員會檔案管理局 (2020)。國家檔案管理作業手冊。檢自 <https://www.archives.gov.tw/Publish.aspx?cnid=1644&p=3723>
- 曾元顯 (2002)。文件主題自動分類成效因素探討。《中國圖書館學會會報》，68，62-83。
- 曾元顯 (2011)。文獻內容探勘工具－CATAR－之發展和應用。《圖書館學與資訊科學》，37 (1)，31-49。
- 曾元顯 (2014)。自動化資訊組織與主題分析近二十年來的研究與發展。《教育資料與圖書館學》，51 (特刊)，3-26。
doi:10.6120/JoEMLS.2014.51S/0652.RV.AM
- 項潔 (2017)。檔案與數位人文：數位人文視野下的檔案系統。《國史研究通訊》，13，152-162。
- 劉昭麟 (2018)。從《數位人文研究》展望資料開放、工具協同與議題網絡。《數位典藏與數位人文》，1，11-50。doi:10.6853/DADH.201804_1.0002
- 劉健清、王家典、徐梁伯主編 (1992)。《中國國民黨史》。江蘇：江蘇古籍出版社。
- 劉維開 (1996)。中國國民黨中央委員會黨史委員會藏會議史料的内容與運用。在《海峽兩岸檔案暨微縮學術交流會論文集》(頁 95-112)。臺北市：中華檔案暨資訊微縮管理學會。
- Liu, C. L. (2017). Flexible computing services for comparisons and analyses of classical Chinese poetry. In *Proceedings of the 2017 International Conference on Digital Humanities (DH 2017)*. Symposium conducted at McGill University & Université de Montréal. Montréal, Canada.

附錄一

入庫提調單

[illegible]

附錄三

檔案處理作業註記單

檔案處理作業註記單						館藏號：		
第一聯：前置作業			處理人員：			日期： 年 月 日		
達第 1 級受損類型		頁碼		備註				
紙質破裂								
嚴重脆化								
嚴重蟲蛀								
嚴重霉斑								
紙面沾黏物				張數		頁數		
其他：								
分件欄								
件次	頁碼範圍	建檔	件次	頁碼範圍	建檔	件次	頁碼範圍	建檔
1			29			57		
2			30			58		
3			31			59		
4			32			60		
5			33			61		
6			34			62		
7			35			63		
8			36			64		
9			37			65		
10			38			66		
11			39			67		
12			40			68		
13			41			69		
14			42			70		
15			43			71		
16			44			72		

黨史檔案數位典藏作業及其數位人文工具應用經驗

17			45			73		
18			46			74		
19			47			75		
20			48			76		
21			49			77		
22			50			78		
23			51			79		
24			52			80		
25			53			81		
26			54			82		
27			55			83		
28			56			84		
第二聯：檔案數位化 處理人員： 日期： 年 月 日								
處理情況			頁碼			備註		
編頁漏頁								
編頁重頁								
重複代頁								
破損代頁						張數	頁數	
第三聯：校驗檢核 處理人員： 日期： 年 月 日								
檢核項目		頁數/數位檔名/欄位					備註	
前置	編頁漏頁							
	編頁重頁							
數位化	其他：							
	影像歪斜、不全							
	影像重複、漏頁							
	折痕、遮字							
	透頁、模糊							
	色差、亮度調整							
	其他：							

目錄	加強描述		
	用字錯誤		
	其他：		
第四聯：修正		處理人員：	日期： 年 月 日
備註			
第五聯：描述建檔		處理人員：	日期： 年 月 日
備註			



Digitalization of Party History Archives and Application Experience of Digital Humanities Tools

Chiao-Min Lin^{*} Li-Hsun Wang^{**}

【 Abstract 】

The archives of the Kuomintang (KMT) are an important source of research on modern Chinese history at domestic and international. In 2018, The National Chengchi University and the Kuomintang signed the “Party History Archives Entrusted Management Cooperation Agreement” and moved this batch of historical archives to the Social Sciences Information Center of Chengchi University. After accepting the entrusted management of the party history archives, the Social Sciences Information Center began to organize and digitize this batch of archives. This article introduces the principles of party history archives scanning and data filing, and provides empirical case in using digital humanities tools to assist in

* Professor, Graduate Institute of Library, Information and Archival Studies, National Chengchi University

Section Chief, University History and Archives Section, Social Sciences Information Center, National Chengchi University

ORCID 0000-0002-9309-9884

Principal author for all correspondence E-mail: cmlin@nccu.edu.tw

** Officer, University History and Archives Section, Social Sciences Information Center, National Chengchi University

ORCID 0000-0002-6042-9218

E-mail: matt@nccu.edu.tw

automatic classification. It is expected to provide experience in archive digitization and content analysis as a reference for other collection agencies.

Keywords

Archive digitization, Digital humanity, Digital archive, Archive access, Automatic categorization

【 Summary 】

Kuomintang archives are important historical data of modern Chinese history. National Chengchi University (NCCU) and Kuomintang signed the “Party History Archives Entrusted Management Cooperation Agreement” in 2018 to entrust Kuomintang archives to NCCU for the management. This study aims to introduce the background of entrusted management of the batch of party history archives and the experience of the Social Sciences Information Center of Chengchi University (the Social Sciences Information Center) in the reorganization, digitalization, and digital humanities application of the batch of party history archives.

Since the user groups of the party history archives are modern history researchers and scholars and NCCU presents excellent reputation on the research on modern history, storing the batch of party history archives in NCCU results in the gathering of academic resources and scholars. It does not simply provide the utilization of research communities, but could also develop the practicability of distinct academic research cooperation.

Key Emphasis on Archives Organization and Digitalization

Building Archives and Application Space

To cope with the collection of the party history archives, the Social Sciences Information Center repairs the space and remodels the standard record

storage with temperature and humidity control and fire resistance period for more than 1 hour, equips access security system, surveillance cameras, and communication equipment, purchase steel archives racks for proper arrangement and setup, as well as plans the archive retrieval use space for archives application.

Developing Work Specifications and Record Chart

Aiming at archives organization, the Social Sciences Information Center develops job specifications, including “Archives Storage Management and Control Specifications”, “Code of Practice for Digital Archives Predecessor Activity”, “Code of Practice for Digital Archives Format”, and “Archives Description Specifications” as well as record charts of “Warehousing and Transfer Order”, “Temperature and Humidity Records”, and “Archives Processing Notes” to uniformize the archive workflow and maintain the operation quality.

Planning Digital Workflow

The archive digitalization processing workflow is established, including pre-processing, digital scanning, image recognition, and metadata filing, where the processes could be divided into several steps. The operating personnel are requested to follow the strict job specifications for the scanning.

Setting Image Scanning Specifications

The setting of image specifications should consider costs and the prospective for the future use. By referring to mainstream specifications, archive level and reading level image specifications are set. TIFF file are applied to archive level, with 300DPI, and the conversion to PDF with watermark is used for reading level.

Setting the Naming of Digital Files

Image files, after digital scanning, need to be give the sole naming code for identification and avoiding repetition. The electronic file naming principles apply English and number that English codes are used for replacing the original Chinese serial numbers. As the example of naming archive level digital files, the file name covers serial number code, volume number, secondary volume number, and page number °

Experience of Party History Archives Digital Humanities Development

Kuomintang's party history data are important cultural assets of Kuomintang. However, the early archive organization was not established. Besides, the file folder arrangement of some file series was not in order, after several times of move. They were simply placed in folders chronologically and shelved according to the archive generation time. Archive data without classification and arrangement could not be orderly organized the information for retrieval and inquiry or reveal the content to develop the archive value. If such huge quantity and unorganized archives are re-arranged by manpower, it will waste a lot of time and violate the primitive sequence theory for archival arrangement.

Without changing the arrangement order of original files, the Social Sciences Information Center therefore applies text mining to analyze text content of archive folders lack of subject classification, construct subject classification structure of text content, and provide online retrieval system for inquiry and use, without reorganizing the physical file folders and contradicting the primitive sequence principle for archival arrangement insisted in archival theory. It could help enhance the efficiency of file reorganization and archive retrieval.

A content mining tool, named CATAR (Content Analysis Toolkit for Academic Research), is utilized for the automatic classification (Tseng Yuan-

hsien, 2011). Archive title, abstract, and content involving in people, events, and objects, which cover vocabulary in subject concept field are ordered and analyzed according to “correlation coefficient” and “term frequency” to automatically capture category feature words as the category titles and cluster archive files. Finally, according to the clustered feature words, 13 categories of topic structure are generated. To confirm the automatic classification result, expert interview is preceded for adjusting the classification structure. The completed automatic classification result provides two applications.

- (1) To key the names of main category and sub-category generated by classification into the metadata field to increase the subject access vocabulary of the archive description field.
- (2) The classification structure corrected by automatic classification and manpower could provide archive retrieval system user interface for directly clicking on archive file through subject browse.

The batch of party history archives present historical research value. After being entrusted to NCCU, NCCU, with the academic energy, precedes professional data organization and research and fully develops the data content value. Moreover, digital tools for automatic classification are also utilized for archive content mining and clustering and constructing classification table for archive subject browse. The experimental research proves that digital humanities tools could be applied to enhance archive management efficiency and promote archive content and topic analysis effectiveness.

Romanized & Translated Reference for Original Text

Besser, H. (2009)。影像製作入門 (*Introduction to imaging*) (林彥宏譯)。臺北市：拓展臺灣數位典藏計畫。(原作 2003 年出版)。【Besser, H. (2009). *Introduction to imaging*. (Lin, Yan-Hong trans.) Taipei: Taiwan Digital Archives Expansion Project. (in Chinese)】

- 徐明景 (2011)。數位化工作流程指南：影像資料。臺北市：拓展臺灣數位典藏計畫。【Xu, Ming-Jing (2011). *Digitization procedures guidelines: Image data*. Taipei: Taiwan Digital Archives Expansion Project. (in Chinese)】
- 國家發展委員會檔案管理局 (2020)。國家檔案管理作業手冊。檢自 <https://www.archives.gov.tw/Publish.aspx?cnid=1644&p=3723> 【National Archives Administration (2020). *Handbook on national archives management*. Retrieved from <https://www.archives.gov.tw/Publish.aspx?cnid=1644&p=3723> (in Chinese)】
- 曾元顯 (2002)。文件主題自動分類成效因素探討。《中國圖書館學會會報》，68，62-83。【Tseng, Yuen-Hsien (2002). Effectiveness issues in automatic text categorization. *Bulletin of the Library Association of China*, 68, 62-83. (in Chinese)】
- 曾元顯 (2011)。文獻內容探勘工具－CATAR－之發展和應用。《圖書館學與資訊科學》，37 (1)，31-49。【Tseng, Yuen-Hsien (2011). Development and application of a content analysis toolkit-CATAR. *Journal of Library and Information Science*, 37(1), 31-49. (in Chinese)】
- 曾元顯 (2014)。自動化資訊組織與主題分析近二十年來的研究與發展。《教育資料與圖書館學》，51 (特刊)，3 -26。【Tseng, Yuen-Hsien (2014). Research and development on automatic information organization and subject analysis in recent decades. *Journal of Educational Media & Library Sciences*, 51(special issue), 3-26. (in Chinese)】 doi:10.6120/JoEMLS.2014.51S/0652.RV.AM
- 項潔 (2017)。檔案與數位人文：數位人文視野下的檔案系統。《國史研究通訊》，13，152-162。【Hsiang, Jieh (2017). Dang an yu shu wei ren wen: shu wei ren wen shi ye xia de dang an xi tong. *Academia Historica Newsletter*, 13, 152-162. (in Chinese)】
- 劉昭麟 (2018)。從《數位人文研究》展望資料開放、工具協同與議題網絡。《數位典藏與數位人文》，1，11-50。【Liu, Chao-Lin (2018). Toward linking open data, synchronizing digital tools, and mapping relevant topics: A survey of the series on digital humanities. *Journal of Digital Archives and Digital Humanities*, 1, 11-50. (in Chinese)】 doi:10.6853/DADH.201804_1.0002
- 劉健清、王家典、徐梁伯主編 (1992)。《中國國民黨黨史》。江蘇：江蘇古籍出版社。【Liu, Jian-Qing, Wang, Jia-Dian, & Xu, Liang-Bo (Eds.) (1992). *Zhong*

guo guo min dang shi. Jiangsu: Jiangsu Gu Ji Chu Ban She. (in Chinese)】

劉維開 (1996)。中國國民黨中央委員會黨史委員會藏會議史料的内容與運用。

在海峽兩岸檔案暨微縮學術交流會論文集 (頁 95-112)。臺北市：中華檔

案暨資訊微縮管理學會。【Liu, Wei-Kai (1996). Zhong guo guo min dang

zhong yang wei yuan hui dang shi wei yuan hui cang hui yi shi liao de nei rong

yu yun yong. In *Hai Xia Liang An Dang An Ji Wei Suo Xue Shu Jiao Liu Hui*

Lun Wen Ji (pp.95-112). Taipei: Zhong Hua Dang An Ji Zi Xun Wei Suo Guan

Li Xue Hui. (in Chinese)】

Liu, C. L. (2017). Flexible computing services for comparisons and analyses of classical Chinese poetry. In *Proceedings of the 2017 International Conference on Digital Humanities (DH 2017)*. Symposium conducted at McGill University & Université de Montréal. Montréal, Canada.