

從九二一震災談臺灣文獻史料數位化之重要性

楊正寬

壹、前言

兩件不相干的事，因為臺灣文獻史料的典藏，便變成了密不可分，也成為探討本文的動機。

第一件事是大家餘悸猶存，驚慌甫定，發生在民國八十八年九月二十一日一時四十七分，芮氏七·四級的二二大地震。剎那間天搖地動，東倒西斜，形同廢墟。災區如此，位在災區的臺灣省文獻委員會所典藏的臺灣文獻史料亦復如此。災後日本關心文獻會典藏日治時期臺灣總督府檔案的檜山幸夫等二、三十位學者，不但捐款襄助本會從事檔案復舊，而且還蒞會關切。其中一位來自東京的國文學研究資料館的青木睦表示，這些珍貴文獻史料，很僥倖沒有像日本阪神大地震時很多珍藏史料遭受災後，因為失火又加上消防水漬的多重打擊。事實上，在斷層密布的南投文獻會三棟建築物仍能屹立，只是檔案史料與玻璃櫥櫃一片狼藉，只要費神復舊，臺灣的文獻史料還是可以任重道遠的傳承香火。

第二件事是數位化時代的來臨。根據「數位時代」雜誌的調查發現，臺灣十五歲以上的「數位人」約為四百二十萬人，約佔總人口數的四分之一。如果將去年數位指數訂為一、〇〇〇，今年則為二、五〇〇，比去年增加一·五倍（註一）。因此，總體而言，臺灣正朝向數位社會邁進。

英國中小企業暨電子商務部部長派翠西亞休特（Patricia Hewitt）在二〇〇〇年世界資訊科技大會強調「電子化政府」是英國最有野心的計畫之一，並預計在二〇〇二年，英國將變成全球電子商務領先的國家。而世界資訊科技與服務業聯盟（WITSA）董事長哈里斯·米勒（Harris Miller）在該大會閉幕後確認「全球數位革命中心在臺北」（註二）。

乍看下數位化時代的來臨與文獻史料似不相關，但是仔細研究瞭解，數位資訊影響所及，不僅提供了新的寫作方式，新的多媒體表現方式，也改變了基本的數據，資料和知識表達與呈現，進而影響到溝通、學習、研究和所有做事的態度，方法和效果。臺灣既為全球數位革命的中心，文獻史料又攸關提供人類基本的數據、資料知識表達和呈現。因此讓文獻、史料或檔案配合數位化時代的需求，使人類文明得以忠實記錄和傳承，也就是將文獻史料典藏數位化，已是我們努力的不二方向。本文即是植基於如何善盡臺灣文獻史料典藏責任，從九二一大地震的毀壞及數位化的典藏，趨吉避兇將所有典藏史料掃描變成電腦顯像的數位（Digital）處理，不但可以安全保存，且可以藉網路廣為開放利用。

貳、臺灣文獻史料數位化可行性評估

檔案史料典藏的目的，在於開放利用與保存永久，世代傳承。但如開放原始史料直接利用，不免有受污損、毀壞的風險；不提供利用，則有失典藏的意義。為解決此一問題，世界各國之檔案或史料館，均投注莫大的心力於相關的研究上，典藏數位化正好提供這種功能，可以減少原始史料直接調借利用之次數，使檔案在「保存」與「利用」之間取得平衡發展的滿足點。

典藏數位化，對管理典藏文物資料制度的建立及運作規範方面，皆產生不小的衝擊。此一新興科技迅速的成長發展，正快速地改變博物館、美術館、圖書館或其他性質相近的文教機關（構）的經營管理模式。同時，由於此一科技具有廣泛運用的特性，影響及大、中、小型機構，使公私文教團體無法自外於這股趨勢。國內較大型而具有計畫性型態之專業工作，由國科會於民國八十六年展開籌備，八十七年度起推動「數位化博物館專案」。

「數位化博物館」專案之目的在於網際網路上建立具有大眾科學或文化教育價值之多媒體資料呈現機制，包括網站及其後續技術發展所需之各項模組。「國家典藏數位化計畫」目的，則側重在典藏文物資料的保存與利用。所謂典藏的「利用」，是將數位化之後的各種產出，作為相關產業的原料或素材，透過相關產業的加值，產生各式各樣的產品，以補助學術研究、各種教育、文化事業所需的多元化利用。八十八年七月行政院「電子、通訊、資訊策劃會議」正式通過「國家典藏數位化計畫」，指定七個單位參與。臺灣省文獻委員會為七個單位之一，故立即根據典藏特性，提出以日據時期至光復初期檔案數位化為主題的計畫——典藏日據與光

復初期史料數位化，參與「國家典藏數位化計畫」。預計自九十年元月起，以五年為期，完成包括檔案文書原始手稿的分類編目、影像掃描、索引建檔、檢索系統的開發與建置等任務（註三）。

史料的保存與運用為典藏管理首要的工作。「保存」又是「運用」的基礎。而原始檔案十之八九為手稿式史料，內容珍貴但形式脆弱，其特性即其「唯一性」。且若干檔案文書已超過一百年以上，紙張自然老化的現象無法避免；但史料以能提供研究、法律信證或歷史事件稽核為其保存下來的積極目的，故應設法善用此一資源。透過數位化保存，則可達到永久傳承史料文物之目的，且就積極面而言，已數位化之檔案史料，將可充分提供各種運用，發揮典藏管理的目標。

本會在自據時期檔案方面，目標在製作管理大尺寸文件的數位化，並配合與中研院雙方合作進行中的「臺灣總督府檔案清查整理與製作光碟計畫」，使檔案文件不拘其尺寸大小，共同設計於同一個應用系統，提供網際網路線上檢索及調閱服務。其次，在光復後檔案方面，刺配合本會策劃執行「臺灣省各機關（構）、學校檔案移轉管理作業」進行後續整理典藏，加強數位化製作之目標，預計針對「臺灣省行政長官公署時期檔案」試行整合分類、編目，使同一時期檔案史料不論原來分屬臺灣省政府任何廳、處、會管理，未來都在同一應用體系之下，一目瞭然，開放各界調閱利用。

參、臺灣文獻史料典藏內涵

臺灣省文獻委員會目前典藏的臺灣文獻史料包括日據時期及光復後兩大部分，臚陳如下：

一、日據時期文獻史料

包括三部檔案，第一部是臺灣總督府內各部、局、課、係保存之公文文書。檔案涉及至今五十餘年至一百餘年的歷史現象，包括政治、社會、文教、經濟……等各個層面施政的情況，稱為「臺灣總督府公文類纂」檔案總數一萬五千零六十二卷。另一部是獨立於臺灣總督府之外的專賣局業務檔案，稱為「臺灣總督府專賣局公文類纂」，約九十餘卷，記載明治三十（西元一八九七）年至日據末期臺灣總督府實施專賣制度的歷史檔案，目前仍在積極編目整理中，內容包括早期樟腦、食鹽、煙草、鴉片、糖務、火柴、酒精、酒、石油……等專賣事業之經濟類文獻。兩者都是研究日據時期臺灣史不可或缺的原始史料，亦是國家所珍貴的文化資產。第三部是臺灣總督府主導成立的「臺灣拓殖株式會社」文書，計二千八百五十三冊，是研究日治時期，由總督府主導，在昭和十一（西元一九三六）年底設立的國策公司，日本政府以臺灣為根據地，發展對南進政策的重要參考史料（註四）。

二、光復後文獻史料

包括「臺灣省行政長官公署」時期檔案，約四千餘卷，以及本會於八十八年底前，利用精省機會，將原省屬各機關檔案移至本會珍藏的省屬九十九個機關學校，高達九萬四千零四十二卷三十年以上的公文檔案（註五）。

肆、計畫辦理數位化情形

一、作法

本會在建構所存檔案數位化計畫時，從橫向分工而言，初步擬定下列四項子計畫（註六）：

（一）檔案史料整理子計畫：

在各種型態之資料數位化之前，應先探討其資料構造特性、內在邏輯秩序，擬定整理規則，前後一貫。此為前置性作業。因此整理的工作為前置性作業，包括下列項目：

1. 清查、盤點
2. 登錄、編碼
3. 分類
4. 排序整架

此部份工作需具備史學專業知識者參與或居間指導協調。

（二）目錄及索引建檔子計畫：

包括下列項目：

1. 目錄格式的研究及定義
2. 史料性質的分析、鑑別、選擇
3. 目錄編寫
4. 文字的輸入建檔
5. 校對

此部分工作銜接史料整理的前置作業而加以發展，需具備史學與相關人文社會科學領域之專家參與協調，

並有資訊專業背景者參與。

(三) 檔案史料影像掃描子計畫：

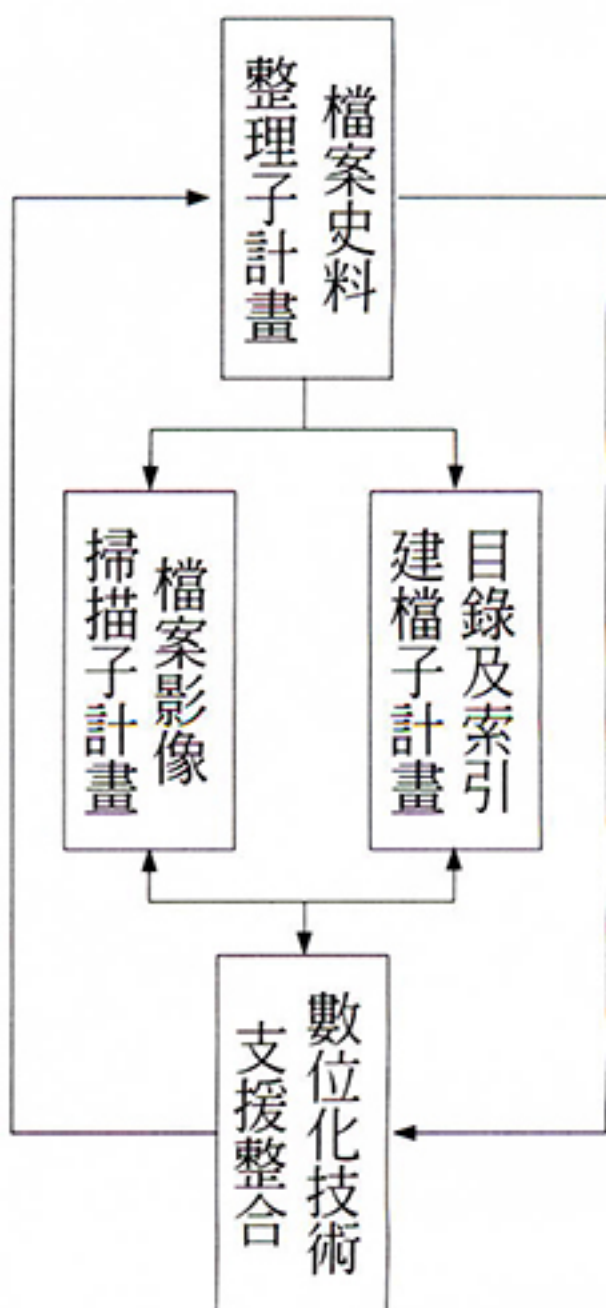
此部分為操作性工作，在作業規範、檔案規格討論確定之後，可採取公開徵選廠商，駐點委辦方式實行。所需參與人才應介於同時具備資訊科技與歷史學知識者擔任。

(四) 數位化技術支援子計畫：

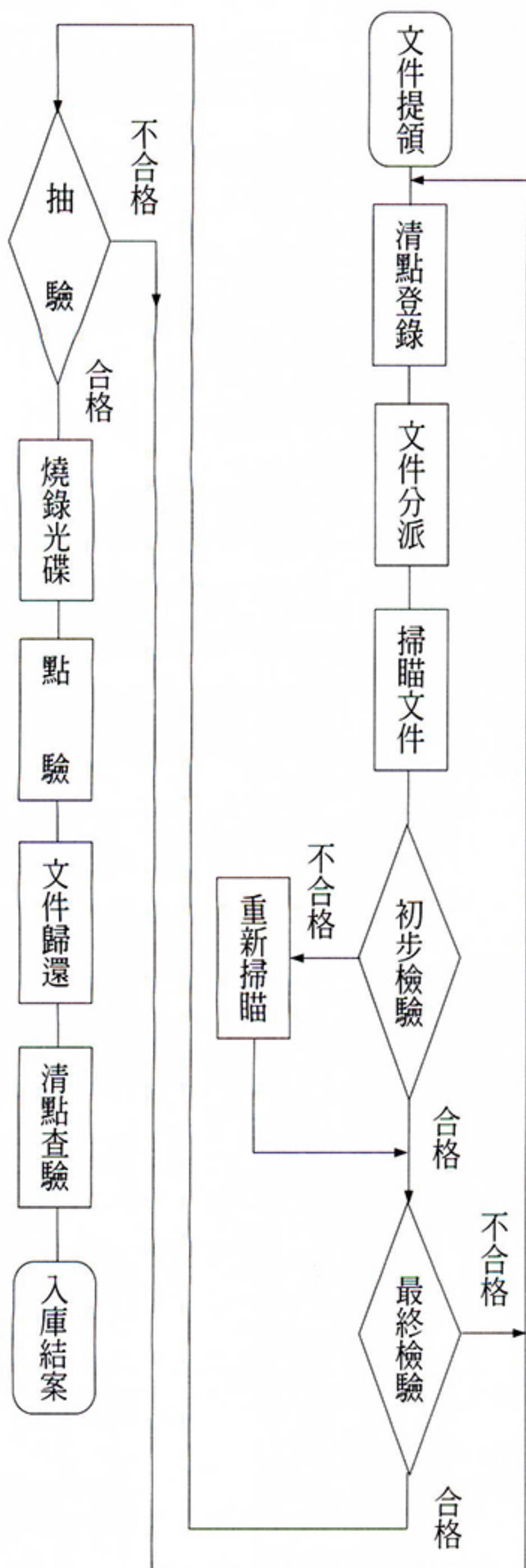
包括以下工作項目：

1. 資料庫研發
2. 網路應用環境規劃設計
3. 軟體設備規劃、建置
4. 系統整合運用

此部分需資訊科技背景人才參與協助，特別是有關大型資料庫管理與運用、網路規劃設計人才。
茲將以上四者的關係圖示如下：



各項工作應有具體流程與規範，以檔案影像掃描為例，其流程表示如下：



(取材自：典藏日據與光復初期史料數位化中程綱要計畫書，臺灣省文獻委員會，民國八十九年元月，頁十三。)

二、數位化執行情形

本會珍藏之日據時期臺灣總督府檔案，自民國六十八年起至七十五年止，即分八年編列預算製作微捲一千六百八十六捲，開放調閱。惟早期微縮影片欠缺索引工具，不易普遍推廣運用，且影像處理技術限於媒材，使得微縮影像與原始手稿有相當大的差距。本會乃於民國八十二年另訂「日據時期檔案製作光碟計畫」，擬以十五年時間完成檔案數位化工作。迄至八十六年，完成四千餘冊檔案掃描建檔。

在民國八十二年影像處理技術仍未純熟、網際網路亦未普遍使用之前，檔案製成光碟，有許多保存及使用上的限制。例如，影像掃描無法採用全採模式，或全採模式成本昂貴，時間效益不彰……等。在資料庫系統開發方面，亦格於作業系統激烈轉換之際，前一年建置在DOS下之系列軟體，使用一年便被迫而改採Windows介面。操作系統的多變，正是過去六、七年來科技技術快速變遷絕佳的寫照。此階段所製作之光碟或數位化製品，嚴格而言，皆具有若干實驗性質。這種實驗性的計畫使資料處理的特性及限制更為彰顯。歸納而言，民國八十二至八十六年間檔案數位化為階段性產品，其情形類同於六、七〇年代盛行微捲製作一樣，在當時的科技條件之下，從事典藏管理與運用的普遍性選擇。

「日據時期檔案製作光碟計畫」在規劃之初，由於面臨科技進步速度難以預測，乃採較保守的估計，以十五年為期，列為業務發展長程計畫。十五年為期包括將「臺灣總督府檔案」一萬五千餘卷（冊）全部數位化，並建立索引，開發資料庫系統。然而，國內外學術界人士對此一重要文獻開

放利用之需求日益迫切，頻頻促進儘早完成數位化計畫。因此本會乃在民國八十六年檢討修訂計畫，縮短數位化計畫為五年，並增加「專賣局檔案」，將兩類檔案整合為同一應用系統下的影像資料庫。民國八十六年十一月十日，本會與中央研究院計算中心、臺灣史研究所籌備處、近代史研究所、中山人文社會科學研究所、經濟學研究所等共同簽訂學術合作協議，成立專案，預計以五年時間，每年由合作雙方共同籌措資金、投入人力，聯合整理臺灣總督府與專賣局檔案數位化計畫。此一專案至今已邁入第三年，初步完成檔案文件影像掃描約二百萬頁，建檔文字約五千萬字。八十八年十一月十一日曾於中央研究院近史所舉行「初步成果展示會」，將影像資料試著整合於網際網路瀏覽器之上運用。雖然在分類、編目上尚未臻理想，但在網路上的運作已經實證為可行。

三、過去二年重要執行成果

(一)自民國八十七年七月起至民國八十八年十二月止，累積檔案數位化成效共計完成約八千冊、二百餘萬頁文件，索引建檔約完成十餘萬筆，五千餘萬字。

(二)處理古老文件影像掃描的產能逐漸從最初的每日三千多頁，提昇至今九千多頁。每年處理影像掃描頁次的能力可達二百萬頁次以上。

由於處理檔案數位化的流程及工作規範日趨嫻熟，各項技術規格亦已有跡可尋，將來處理大型數位化影像資料庫，似已奠立寶貴的實務經驗。

伍、結論

臺灣文獻史料的典藏，除了注意防火防水、恆溫恆濕外，並且要將破舊文獻史料趕緊裱褙維護。然而此次地震災害後，本會亦特別於八十九年六月舉行為期三天的臺灣文獻史料學術研討會，其中日本國文學研究資料館研究員青木睦，即針對阪神大地震經驗，提出：「被災史料の救助と修復復元」^(註七)，寶貴的經驗仍然無法根本達到保存典藏與開放運用的目標，因此評估結果，面對數位化時代與網路世界的來臨，特別與國科會自九〇年度起為期五年，繼與中研院合作結束後，進行國家典藏數位化計畫^(註八)。茲將預期效益分析及其影響總結如下：

一、效益分析

- (一)挽救已瀕臨風化、腐損而不堪使用之珍貴文獻史料檔案。此檔案為國家珍貴文化資產，應予永久保留典藏。
- (二)配合國家檔案法的公佈實施，永久保存檔案經一段時間之後，應開放調閱、研究。數位化後的檔案文件資料庫搜尋效率將大為提昇，皆因科學技術進步，提供之應用方式將可更多元化，效率更佳。服務的方式或範圍並可達無遠弗屆之理想目標。

二、影響評估

- (一)過去在日據時期檔案方面，以製作大尺寸圖表為目標。大尺寸文件，大於A三以上，最大達全開，在製

作技術方面，技術規範及能力難度頗為嚴格。而建立大尺寸文件製作之範例，可提供典藏數位化另一種特殊面向，建立有價值的參考模式。

- (二)大宗文件的數位化案件，以影像掃描製作為例，數位影像預估將達到數百萬件；因而建立之資料庫系統是否得以順利而有效率運作，考驗主事者的能力與經驗。反之，倘能順利推動，則在大型典藏資料數位化開發與應用方面，將極具參考價值。以影像檔為例，預估將完成數位化之影像儲存量高達三至五TB以上。

- (三)因應「檔案法」通過實施^(註九)，將來許多機關（構）甚至各級學校保存管理之檔案文件，亦將陸續產生數位化需求。故以檔案管理結合現代科技完成數位化典藏，應有建立指標性的作用。

- (四)如在數位化的過程中實施順利成功，本會預定除光復初期臺灣行政長官公署檔案外，方可與時俱進，依序推動四十、五十，甚至六十年代等所有記載臺灣經驗之官署檔案數位化。

【註釋】

註一：「臺灣數位人較去年增加一·五倍」，中央日報，民國八十九年六月二十六日，第五版。

註二：「全球數位革命中心在臺北」，聯合報，民國八十九年六月一日，第六版。

註三：臺灣省文獻委員會：典藏日據與光復初期史料數位化中程綱要計畫書，民國八十九年一月，頁一。

註四：臺灣省文獻委員會：日據時期檔案簡介，民國八十八年六

月，頁一至三十六。

註五：同註三。

註六：楊正寬：臺灣精省後原省屬各機關檔案移轉政策分析。臺灣文獻，第五一卷第一期，臺灣省文獻委員會，民國八十九年三月出版，頁十至二十。

註七：青木睦：被災史料の救助と修復復元。臺灣省文獻委員會，民國八十九年六月七一九日，臺灣文獻史料學術研究研討會發表。

註八：行政院國家科學委員會民國八十九年二月二十四日（八九）臺會企字第〇〇七〇六一號函。

註九：檔案法，總統民國八十八年十二月十五日華總一義字第八八〇〇二九七四八〇號令公布。

