

明清內閣大庫檔案 Metadata 研擬程序

洪一梅

中央研究院歷史語言研究所研究助理

摘 要

自 2001 年起，史語所全面重新規劃數位化作業，針對史語所藏內閣大庫檔案，評估科技的進展及國際檔案管理趨勢，重新確認發展方向，主要目的在於尋求更多的人力、物力，加速檔案的整理工作、減緩檔案的毀損、促進檔案的開放，提昇檔案的利用價值，並同時盡到對國家財產保存與維護的使命與責任。

本文概括說明本計畫的目標及規劃方向與目的，主要著重在整體目標方向下之其中一個細部執行計畫「metadata 研擬」的解析，詳細描述了 metadata 研擬的程序，及過程中考量的及所受的內外部影響因素，嘗試提供予檔案典藏單位進行 metadata 分析的參考。

關鍵詞：內閣大庫、metadata、數位典藏計畫

壹、前言

史語所藏內閣大庫檔案總計約三十一萬件，包括詔令、題奏、移會、賀表、三法司案卷、實錄稿本、各種黃冊、簿冊等，為研究明清史的重要史料。史語所鑒於此，從民國八十四年始即嘗試進行數位化作業，早期因技術的限制，僅執行簡單的電腦登錄作業，檔案實體的保存與利用則有賴紙本複印及選擇性出版。

近年來，因資訊技術的精進，自 2001 年起，史語所全面重新規劃數位化作業，針對其所藏內閣大庫檔案，評估科技的進展及國際檔案管理趨勢，重新確認發展方向，主要目的在於尋求更多的人力、物力，加速檔案的整理工作、減緩檔案的毀損、促進檔案的開放，提昇檔案的利用價值，並同時盡到對國家財產保存與維護的使命與責任。

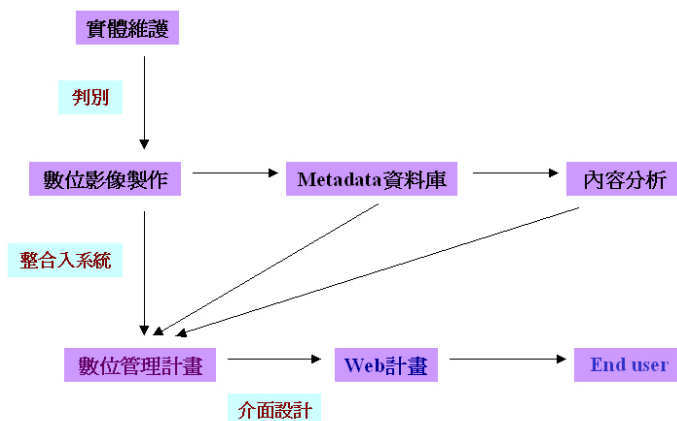
本文首先概括說明本計畫的目標及規劃方向與目的，主要著重在整體目標方向下之其中一個細部執行計畫 - metadata 研擬的解析，詳細描述了 metadata 研擬的程序，及過程中考量的及所受的內外部影響因素，嘗試提供予檔案典藏單位進行 metadata 分析的參考。

貳、史語所研擬程序

一、計畫目標

從二十世紀初以來，電腦的普及與軟體科技的成長，使得人類生活產生重大的衝擊與影響。就如同紙的發明與使用一樣，數位資料處理及數位媒體成了新一代的記錄載體，它根本地改變了傳統文件的產生、處理及應用方式。這樣劃時代的資訊革命，將檔案的維護策略從原始物件保護的觀念演變為資訊移轉的觀念。這過程的變因導引於數位媒體與資訊技術的發展，觸發於使用者與管理者需求模式的改變。檔案維護目的由「保存」到「開放利用」整合成為「生命週期的管理」，規劃方向邁向科技化、管理制度化、資訊網路化；目標為整合式的數位管理機制，兼含保存、維護與開放利用。

史語所明清內閣大庫檔案整理工作於九十一年擴大規劃，並整併該所其它數位典藏計畫成為數位典藏國家型科技計畫的一環，並確認了三大發展方向：「數位檔建置計畫」、「數位管理計畫」及「Web 計畫」。(見下圖)



圖一 史語所明清內閣大庫檔案數位計畫架構圖

(筆者撰稿，數位典藏國家型科技計畫 - 內容發展分項計畫製圖)

二、計畫目的

(一) 數位檔建置計畫

此方向主要界定在檔案數位資料的建立，包含數位影像、電子目錄及內容分析(metadata)等。而在此之前，無論是針對檔案維護保存或是檔案開放利用的目的，檔案實體維護為必要的開端與基礎，因此將數位檔建置計畫的起始程序往前回溯至檔案實體維護階段。

(二) 數位管理計畫

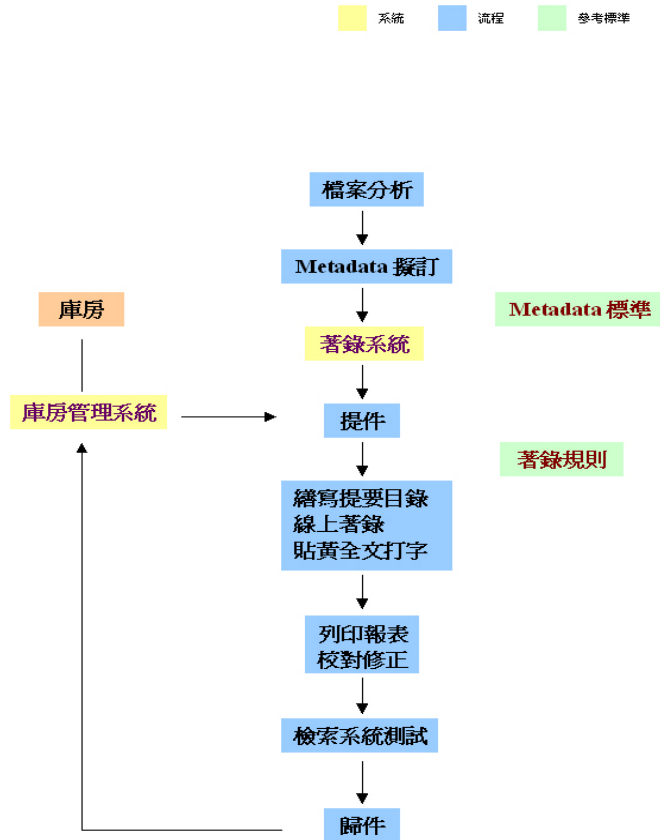
此方向主要為「整合式檔案管理自動化系統」的設計開發與應用，藉由自動化系統的管控，將檔案管理的機制與流程融入其中，使管理制度更為完善也更為嚴謹。而這包括了資料端、應用端、物件流程管控端的整合，使資訊(數位檔建置計畫之成果)得以藉由資訊流程(數位管理計畫)傳輸至使用者(web 計畫之終端)。

(三) web 計畫

此方向主要為互動式網站的建置。著重在國際化，採用中英文雙語系；並整合明清檔案之國內外相關資源，藉由互動式資料庫的呈現方式，使用者不僅可使用線上檢索系統查尋本所典藏之明清檔案，並可獲得相關參考資源電子版如六部成語等，更可藉由網站的瀏覽，了解明清檔案整理規範、細節與流程，促進檔案整理經驗的分享與整合。未來使用者更可直接透過網站回饋其本身的經驗或提供相關研究資源與研究成果的公開分享。

三、數位檔建置計畫—metadata

明清檔案分析作業因為較為困難，需由具有明清史專門知識的人員進行讀檔後，方能進行，因此通常會在數位影像製作完成後的二年左右進行。此部份首要在進行 metadata 資料庫的建置，流程中從庫房管理系統進行提件，於著錄系統中進行線上著錄，經過反復地內容校對後，於檢索系統中進行測試方算完成。其中除了在原始的 metadata 分析時有參考相關的國際標準如 MARC21、EAD 等外，在著錄過程也參考了 ISAD(G)、AACR2 等研擬了明清檔案著錄規則以為遵循，確保資料庫的品質。



圖二 史語所明清內閣大庫檔案數位檔建置計畫 - metadata 建置程序

(筆者撰稿，數位典藏國家型科技計畫 - 內容發展分項計畫製圖)

四、metadata 分析與研擬程序

(一) 界定使用者及系統目的

1. 在史語所計畫整體目標中，91-95 年為基礎建設期，主要在建立典藏資源整理的基礎架構與標準規範，因此在典藏資源分析上都期望做到最深入，如此在未來應用及推廣上將有無限可能。
2. 史語所的任务發展主要在研究工作，典藏資源的整理主要也以支援研究為目的，因此預期使用對象為研究人員，研究人員的需求在於「嚴謹」、「正確」、「快速獲得」。
3. 系統目的在計畫規劃之初，不僅只在使用端，還包括管理端，因此系統的設計開發與應用，皆朝向「整合式檔案管理自動化系統」來發展。系統中的各主要模組如下：
 - A. 安控模組(含管理者、工作人員及使用者的權限控管、報表列印與系統資料庫的備援機制)

- B. 著錄模組(含檔案目錄與權威資料庫的著錄維護介面與控管)
 - C. 線上檢索模組(主要為線上檢索工具的開發，將整合入 WEB 計畫)
 - D. 庫房管理模組(實體狀況與物件流程的控管機制)
 - E. 館內檢調模組(限於館內運作，整合線上檢索與實體調件)
 - F. 流通模組(含使用者資訊及實體調件等各項參數的控管，依據本所擬訂之「調閱規則」進行相關設定)
 - G. 數位影像管理模組(含線上閱覽控制、掃校系統、移轉系統等數位影像管理機制)
4. 綜合上述，metadata 的分析不僅止於內容資訊的分析，還包含了管理資訊的分析及使用資訊的分析。而分析的程度則是以深入分析、研究性分析為原則。

(二) 參考文獻蒐集與研究

在參考文獻的蒐集與研究上，主要以下列數種為參考指標：

1. metadata 內涵

A. 管理的 metadata(Administrative Metadata)

目的：為了管理、決策、保存記錄。

種類：

- Preservation Metadata
- Right and Access Metadata
- Meta-metadata

內涵：獲取、擁有、權限、許可、複製、合法存取條件、定址、使用、使用管理、保存等資訊。

B. 結構的 metadata(Structural Metadata)

目的：描述檔案、資料集、資訊包的組成或結構。

表現：在控制層次上，決定描述層次及資訊提供的結構。

內涵：軟硬體的描述、技術性資訊、版本控制、確認影像版本及定義、數位化資訊、關於數位影像產生的資訊(如數位檔的命名)、真實性及安全性資料、相關檢索協定、檔案控制層次描述。

C. 描述的 metadata(Descriptive Metadata)

目的：確認資訊包的特徵、分析及內容。

種類：

- Data that identifies an information package
- Intellectual organization data

- Intellectual access data

內涵：

- 題名、著者、出版日期、與數位物件有關的相關資訊。
- 權威控制、相關作品、姓名、主題配置、不同實體間的關係
- 主題標目、分類號、檢索點等[1]

D. 影響

- 決定採用包含三種資訊種類的 metadata 內涵。
- 權威控制採外部系統控制方式，另建權威資料庫。

2. EAD

A. 發展沿革

1998 年時研擬了 EAD 第一版；2002 年修訂成爲第二版。並有完整的應用指南含行政管理上的考量與做法。

B. 發展原則

- 互通性 - 可交換
- 跨語文 - 國際性
- 公開性 - 可分享可呈現的資源
- 具結構性 - 非內容標準，爲具結構性的複雜標準
- 溝通性 - 資料溝通格式。
- 描述性 - 著重結構性描述，而非呈現，呈現可以多元型態方式。
- 延展性 - 可隨著技術發展，利用現行技術進行輸出，即 DTD 的改變。
- 持續性 - 確認隨著環境改變，EAD 仍持續可被應用。[2]

C. 架構

- Header <eadheader> and Front Matter <frontmatter>：1)提供有關檢索工具本身的資訊，如檢索工具名稱及編輯日期與編輯者；2)提供有關檔案資料實體的資訊，如管理全宗、全宗或系列。
- Archival Description <archdesc>：包含了檔案資料的實體資訊，且資訊是以多元層次結構被組織，描述了每一檔案單元的各類內涵，有內容、層次及檔案資料的範疇和其它有助於檢索者使用的特徵。
- Description of Subordinate Components <dsc>：主要用來描述檔案單元的組成成份，可能包含一個或多個的下一層檔案單元。
- Adjunct Descriptive Data <add> and Other Descriptive Data <odd>：<add> 用來編碼補充的描述資訊，是可在檢索工具中發現的資料使用特徵。這包含了其它的資料開放工具，例如

索引、檔案計畫或其它檢索工具，及在檢索工具中相關的或分散的檔案資料。<odd>用來處理在目錄其它欄位無法處理的資訊。

- **Controlled Access Headings <controlaccess>**: 建立提供予使用者的查尋標目，例如人名、地名、專有名詞、主題等。
- **Pointer, Reference, and Linking Elements**: 用來指示或連結可獲得的參考資訊、數位物件等等。
- **Table Elements**: 為格式化目的重組表單欄位。

D. 編碼格式

EAD 採用二種編碼語言：SGML 及 XML，可認選其一，但因二者之間有些差異，使用者應認知清楚。

E. 著錄規則與標準

EAD 遵循 ISAD(G)以產生 Finding Aids。與 MARC 及 Dublin Core 間有進行對映。[3]

F. 影響

- 將內閣大庫視為一個全宗，以現有館藏區分為「明檔」及「清檔」，而這二個層次不在目錄中描述，放置於 finding aid 的 Header <eadheader> and Front Matter <frontmatter>。
- **Metadata** 需包含內容的及層次的描述資訊，並應提供查尋標目，且可連結相關的索引資訊或其它檢索工具，並可指引或直接對映連結數位物件。
- 針對各表單欄位需進行重組。
- 需有可對映的資料值著錄規則。
- 採用 XML 編碼語言，應先編製 DTD。
- 建立與 Dublin Core 的對映。
- 應隨時注意技術與環境的發展和變化。

3. MARC21

A. 標準對映

MARC 為長期以來圖書資訊界所使用的機讀格式，甚至在美國許多的檔案典藏單位亦使用它或使用 MARC AMC 來進行檔案資料的編碼格式，因此整合國際 MARC 的 MARC21 成為圖書館使用最多的標準。而在 EAD 中也進行了二種格式的對映，以利圖書館檔案館之間的資源交換及分享。且 EAD 格式中亦允許直接註記對映之欄位代碼。[4]

B. 權威資料

MARC21 針對權威資料有非常詳細的編碼標準。[5]

C. 影響：

- 遵循 EAD 格式進行與 MARC21 的欄位對映。
- 參考 MARC21 對人名權威資訊的編碼標準，研擬人名權威資料庫之組織架構，並建立二者間的標準對映。

4. ISAD(G)(1999 年第二版)

A. 多元層次描述

針對檔案資源的每一層次單元，皆可以第三章的描述欄位進行個別層次的描述，為 ISAD(G)的特色，稱為多元層次描述。

描述原則有：

- 從一般到特殊的描述
- 資訊關聯到層次描述
- 同層次間描述單元的連結
- 避免多元層次間相同重覆的描述資訊

B. 欄位分析

區分七種資訊領域 - (第一版為六大領域；第二版增加了 Description Control Area)

- IDENTITY STATEMENT AREA：例如參考的國際標準碼、檔案單元名稱、產生日期、描述層次、描述層次範圍等。
- CONTEXT AREA：例如產生者、行政歷史、單元聚集日期、典藏歷史、徵集來源等。
- CONTENT AND STRUCTURE AREA：範圍及內容摘要、鑑定、損壞、造冊資訊、附加物、編排系統等。
- CONDITIONS OF ACCESS AND USE AREA：法律權限、開放狀況、著作權及重製、資料語言、實體特徵、檢索工具等。
- ALLIED MATERIALS AREA：典藏位置、複本、相關、來源、出版註記等。
- NOTE AREA：無法在其它領域描述의資訊。
- Description Control Area：檔案人員的註記、原則或規範、描述日期等。[6]

C. 影響

- 考慮內閣大庫檔案的層次問題。界定內閣大庫為「管理全宗(Collection)」；明檔及清檔為「全宗(Fond)」；之下除「冊檔」有「案(File)」的層次外，其餘皆以「件(Item)」層次進行描述。
- 每一層次依其檔案單元之特性，選擇適當的描述欄位，含括七大領域的資訊。不同層次間欄位有可能重覆，但因描述層次不同，資料值應為不同；應避免欄位與資料值皆同者在不同層次

間重覆描述，除要做為層次連結的主欄位(primary element)之外。

- 每一個欄位應有定義，及其在系統中的目的與功能，以及著錄該欄位的規則。
- 應注意描述資料中所使用的國際標準代碼，如語文、國別等。

5. ISAAR(CPF)(2004 年第二版)

A. 功能目的

- 提供檔案描述單元中檢索點的控制格式。
- 更有效發揮檔案檢索工具的聚集功能，並促進檔案記錄的有效管理。
- 描述檔案的產生者在檔案管理作業中是非常重要的行為，將有助於檔案的辨識、聚集與管理。

B. 特色

- 較圖書資訊中的權威記錄更詳盡。
- 針對不同實體但同樣名稱或相似名稱的標目必須更明確地區分。
- 配合各國的特質做為發展更適合標準的基礎。
- 標準中的欄位資訊包含描述的欄位名稱、目的、原則、適用的時機與舉例。

C. 欄位分析

在檔案權威記錄的結構中包含四種資訊領域 - (第一版有三種，第二版增訂成四種)

- Identity Area：標目的唯一定義，及標目的標準化檢索點。
- Description Area：標目的領域、範圍、層次與特徵的相關資訊。
- Relationships Area：與其它標目的記錄間的關聯性。
- Control Area：權威記錄中的唯一定義與資訊及權威記錄產生的時間及責任單位。 [7]

D. 影響

- 檔案權威記錄因應檔案的特質需更具詳盡性及可區辨性。
- 檔案權威記錄所描述的標目即為檔案檢索工具中的檢索點，提供控制功能。
- 檔案權威記錄的欄位設計應包含四種資訊領域。

(三) 檔案特質分析

1. 為極重要之研究史料。

史語所前輩學者徐中舒先生指出，「清代內閣在雍乾以前為國家庶政所自出之地，在雍乾以後猶為制誥典冊之府，所存檔案都是當時構成史蹟者自身的敘述」內閣大庫的檔案對歷史研究者的重要性自是不言而喻了。

2. 為明清內閣大庫之檔案，多為殘件，並已遺失檔案原始順序。

其歷經程序如下：大庫移出(舊檔無用) - 歷史博物館(經費短絀 - 不完整 - 八千麻袋事件) - 羅振玉(史料叢刊初編 - 保留局部) - 李盛鐸 - 傅斯年(史語所) - 史語所遷台(擇要裝箱 100)。 [8]

3. 文件類別與格式複雜。

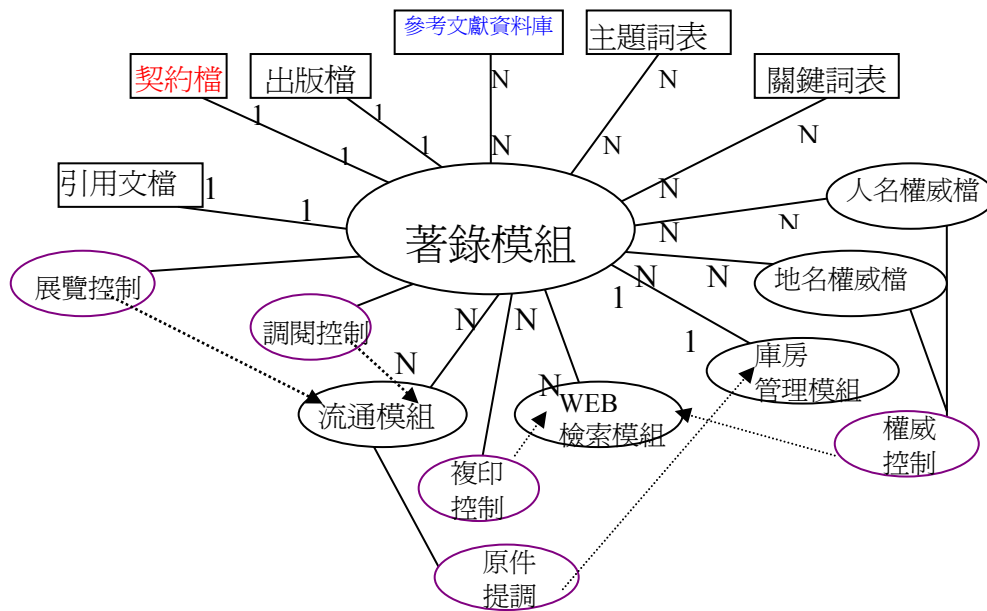
包括詔令、題奏、移會、賀表、三法司案卷、實錄稿本、各種黃冊、簿冊等，再以清代公文書體制細分文件類別，區分文件別、文本二層，形制更為複雜。

4. 整理程序複雜，耗時且費工

5. 影響

- A. 以「件」為描述層次，若為「冊檔」則增列「案」層次。力求詳細深入與完整。
- B. Tag 名稱多採史學術語(如批語、貼黃)；但因格式複雜，為求描述之定義一致，亦有部份採用通用術語(如責任者)。
- C. 以開發整合式自動化系統為目標，應考量整體工作流程，從實體維護、內容整理到開放利用應納入管理及流程的思維。

(四) 系統關聯分析(E-R model) [9]



(五) 模組欄位組織與關聯

1. 著錄模組

A. 內涵

- Administrative Metadata – Right and Access Metadata
- Structural Metadata
- Descriptive Metadata

B. 原因

著錄模組是整合式自動化系統中對檔案資訊描述的主要 schema，因此包含大部份的結構性 metadata 及描述性 metadata。另為了有效控管每一件檔案的資訊內容，將版權聲明及使用管理資訊也納入此模組的 metadata 中。由描述人員針對檔案做深入且完整的描述。

2. 庫房管理模組

A. 內涵

- Administrative Metadata – Preservation Metadata
- Descriptive Metadata – Data that identifies an information package
- Structural Metadata

B. 原因

庫房管理模組是最前端與實體管理相結合的 schema，配合工作流程

進行最初的簡易登錄，因此包含有對檔案資訊的簡易描述；並且爲了妥善地確認每一資訊包的完整性，兼含其附件及散件、殘件，會將相關資訊先行著錄，因此含有部份結構性 metadata；另因工作流程數位檔較目錄檔優先製作，因此於庫房管理登錄時即會包含數位檔相關資訊；而本模組兼含實體及數位內容的管理，因此亦有部份 Preservation Metadata。

3. 權威資料庫

A. 內涵

- Descriptive Metadata – Intellectual organization data(關聯)
- Descriptive Metadata – Intellectual access data(關聯)
- Information package about access point(檢索標目)

B. 原因

因權威控制採外部系統，所以會與描述性 metadata 中的權威控制、相關作品、姓名、主題配置、不同實體間的關係、主題標目、分類號、檢索點等具關聯，而其主要資訊內容爲針對欲控制之檢索標目進行分析與描述。

4. 檢索調閱模組(Finding Aid)

A. 內涵

- Administrative Metadata – Meta-metadata
- Administrative Metadata – Right and Access Metadata(關聯)
- Descriptive Metadata – Data that identifies an information package(目標項)
- Descriptive Metadata – Intellectual organization data(關聯)
- Descriptive Metadata – Intellectual access data(關聯)
- Information package about access point(檢索標目)(控制項)

B. 原因

檢索調閱模組與檔案的各種性質 metadata 皆有關聯，其中描述性 metadata 的資訊包資料及數位資訊爲其檢索的目標項；而描述檢索標目的權威資料庫爲其檢索點的控制項。當檢索調閱模組以 Finding Aid 的型態提供資訊服務時，Meta-metadata 即會成爲其描述與呈現的資訊內容。

(六) 人力資源配置

1. 著錄模組 vs.庫房管理模組

該部份 metadata 由史語所資深助理進行描述，每位至少皆有五年以上的明清檔案讀檔經驗，並由研究員進行審核。研擬有著錄規則，以為遵循。

2. 權威資料庫

該部份由新任助理在接受過訓練後進行描述，目前為止至少皆有二年以上史書傳記資料研讀與整理的經驗，並由研究員進行審核。研擬有著錄規則，以為遵循。

3. 關鍵詞 vs.主題詞

該部份由所有助理進行，過程中由研究員帶領，每月進行關鍵詞圈選及原則歸納。日後以所圈選之關鍵詞進行系統分析，以整合出適用之主題詞。目前仍在測試階段。

(七) metadata 研擬

1. 欄位分析[10]

A. 層次 - 案

項目名稱		模組組織	模組關聯	Metadata 性質		
層次		著錄模組	檢調模組	Structure		
登錄號	全宗號	庫房管理模組	著錄模組	Descriptive - Data that identifies an information package		
	案號		檢調模組			
題名		著錄模組	檢調模組	Descriptive - Data that identifies an information package		
案由		著錄模組	檢調模組	Descriptive - Data that identifies an information package		
責任時間	中曆	著錄模組	檢調模組	Descriptive - Data that identifies an information package		
					起	朝代
						年號
						紀年
						月
					日	
	西曆				迄	朝代
						年號
						紀年
						月
		日				
責任者	名稱	著錄模組	檢調模組 權威資料庫 - 權威控制	Descriptive - Data that identifies an information package Descriptive - Intellectual organization data		
	職銜					
	今地名					

項目名稱		模組組織	模組關聯	Metadata 性質
件數		著錄模組	檢調模組	Descriptive - Data that identifies an information package
件號		著錄模組	檢調模組	Descriptive - Data that identifies an information package
檔案內容註	項目	著錄模組	檢調模組	Descriptive - Data that identifies an information package
	說明			
檔案制度註	項目	著錄模組	檢調模組	Descriptive - Data that identifies an information package
	說明			
關聯	關係	著錄模組	檢調模組	Descriptive - Data that identifies an information package Descriptive - Intellectual organization data
	關聯號(案)			
著錄註		著錄模組	檢調模組	Descriptive - Data that identifies an information package
主題		著錄模組	檢調模組	Descriptive - Intellectual Access data
關鍵詞		著錄模組	檢調模組	Descriptive - Intellectual Access data
典藏狀況	概況	著錄模組	檢調模組	structure
	詳述			
使用限制		著錄模組	檢調模組 - 檢調控制	Administrative - Right and Access Metadata
參考文獻		著錄模組	檢調模組	Descriptive - Intellectual organization data
版權聲明		著錄模組	檢調模組	Administrative - Right and Access Metadata
編目記錄	新增者姓名	著錄模組	安控管理模組	Administrative - Preservation Metadata
	新增日期			
	更新者姓名			
	更新日期			

B. 層次 - 件

項目名稱		模組組織	模組關聯	Metadata 性質
層次		著錄模組	檢調模組	structure
登錄號	案號	庫房管理模組	檢調模組	Descriptive-Data that identifies an information package
	件號			
題名		著錄模組	檢調模組	Descriptive-Data that identifies an information package
事由		著錄模組	檢調模組	Descriptive-Data that identifies an information package
貼黃		著錄模組	檢調模組	Descriptive-Data that identifies an information package
釋文		著錄模組	檢調模組	Descriptive-Data that identifies an information package
文件別		庫房管理模組	著錄模組 檢調模組	Descriptive-Data that identifies an information package
文本		著錄模組	檢調模組	Descriptive-Data that identifies an information package
版本		著錄模組	檢調模組	Descriptive-Data that identifies an information package structure
語文	語文別	著錄模組	檢調模組	Descriptive-Data that identifies an information package
	語文代碼			

項目名稱		模組組織	模組關聯	Metadata 性質		
文向		著錄模組	檢調模組	Descriptive-Data that identifies an information package		
圖像		著錄模組	檢調模組	Descriptive-Data that identifies an information package		
質材		庫房管理模組	著錄模組 檢調模組	Administrative - Preservation Metadata		
數量		著錄模組	檢調模組	Administrative - Preservation Metadata		
裝潢		庫房管理模組	著錄模組 檢調模組	Administrative - Preservation Metadata		
高廣		著錄模組	檢調模組	Administrative - Preservation Metadata		
責任者	名稱	著錄模組	檢調模組 權威資料庫 - 權威控制	Descriptive-Data that identifies an information package		
	職銜					
責任日期	中曆	著錄模組	檢調模組	Descriptive-Data that identifies an information package		
					朝代	
					年號	
					紀年	
					月	
	日					
西曆	今地名					
批紅	描述	著錄模組	檢調模組	Descriptive-Data that identifies an information package		
					日期	中曆
						朝代
						年號
						紀年
						月
日						
西曆						
典藏沿革		著錄模組	檢調模組	Descriptive-Data that identifies an information package		
保存狀況		庫房管理模組	著錄模組 檢調模組	Administrative - Preservation Metadata		
修復狀況		庫房管理模組	著錄模組 檢調模組	Administrative - Preservation Metadata		
使用限制	展覽限制	著錄模組	檢調模組 - 檢調控制 庫房管理模組	Administrative - Right and Access Metadata		
	瀏覽限制					
	複印限制					
附件	附件	庫房管理模組	著錄模組 檢調模組	Descriptive-Data that identifies an information package		
	時間					
附註	項目	庫房管理模組	著錄模組 檢調模組	Descriptive-Data that identifies an information package		
	說明					
關聯	關係	著錄模組	檢調模組	Descriptive - Intellectual organization data		
	關聯號(件)					
主題		著錄模組	檢調模組	Descriptive - Intellectual access data		
關鍵詞		著錄模組	檢調模組	Descriptive - Intellectual access data		
地名資訊		著錄模組	檢調模組 權威資料庫 - 權威控制	Descriptive - Intellectual access data		
人名資訊		著錄模組	檢調模組	Descriptive - Intellectual access data		

項目名稱		模組組織	模組關聯	Metadata 性質
			權威資料庫 - 權威控制	
徵集方式		著錄模組	檢調模組	Administrative - Right and Access Metadata
參考文獻		著錄模組	檢調模組	Descriptive - Intellectual organization data
出版品	品名	著錄模組	檢調模組 - 互動	Descriptive - Intellectual organization data
	卷/冊			
	頁			
影像編號		庫房管理模組	檢調模組	structure
光碟註記	光碟編號	庫房管理模組	安控管理模組	structure
	櫃號			
	層號			
編目記錄	新增者姓名	著錄模組	安控管理模組	Administrative - Preservation Metadata
	新增日期			
	更新者姓名			
	更新日期			
版權		著錄模組	檢調模組	Administrative - Right and Access Metadata
暫行位置編號	架號	庫房管理模組	安控管理模組 檢調模組	Administrative - Preservation Metadata
	層號			
封存號	箱/筒號	庫房管理模組	安控管理模組 檢調模組 - 限制調閱	Administrative - Preservation Metadata
	架號			
	層號			
簡要時間		庫房管理模組	著錄模組	Descriptive - Data that identifies an information package
分類項	分類者	庫房管理模組	安控管理模組	Administrative - Preservation Metadata
	分類時間			
刷灰項	刷灰者	庫房管理模組	安控管理模組	Administrative - Preservation Metadata
	刷灰時間			
裱褙註記	狀況	庫房管理模組	安控管理模組	Administrative - Preservation Metadata
	人員			
	時間			
修補註記	狀況	庫房管理模組	安控管理模組	Administrative - Preservation Metadata
	人員			
	時間			
燻蒸註記	週期	庫房管理模組	安控管理模組	Administrative - Preservation Metadata
	時間			
定期檢查註記	週期	庫房管理模組	安控管理模組	Administrative - Preservation Metadata
	一般時間			
	詳細狀況			
	詳細時間			
庫房首登	日期	庫房管理模組	著錄模組 安控管理模組	Administrative - Preservation Metadata
	時間			
	人員			
更新記錄	日期	庫房管理模組	安控管理模組	Administrative - Preservation Metadata
	時間			
	人員			

C. Meta-metadata : Header <eadheader> and Front Matter <frontmatter>

C.1 【層次 - 管理全宗】

項目名稱	欄位值
類型	檔案
層次	Collection
登錄號	Collection No. 001
典藏單位	中央研究院歷史語言研究所
題名	內閣大庫檔案
內容描述	史語所藏三十一萬多件內閣大庫檔案的內容頗豐，包括內閣收貯的制詔誥敕、題奏表章、史書、錄書、塘報、黃冊、簿冊、朝貢國表章，以及內閣各廳房處的檔案、修書各館檔案、試題、試卷、瀋陽舊檔等，其中以題奏本章佔最大宗。這些檔案內容多涉及一般行政事務，有些並不見於會典或則例。
範圍與內容	根據徐中舒初步整理後的估計，史語所所藏內閣大庫檔案中，明代檔案數量很少，約有四千多件，多為康熙年間為修明史徵集而來，以啓禎年間的題行稿居多；清代檔案則佔大多數，內容計有：一、內閣收貯的各項檔案，如制詔誥敕、題奏表箋啓本副本帖史書錄書塘、黃冊及其他隨本進呈及繳存的文件、朝貢諸表章等；二、內閣本身的各項檔案；三、修書各館檔案；四、試題試卷及其相關的檔案；五、瀋陽舊檔。
責任單位	內閣
責任時間	[明 - 清]
傳記歷史註	<p>內閣在清初為國家庶政重心，雍正以後，其權力雖然為軍機處所分，但仍為承宣制誥典冊之府，而中央與地方官員在密奏承旨之後，國家政事仍須循正式公文管道具題，由內閣發交各部院執行。內閣大庫就是它收藏書籍、詔敕、題奏表章等檔案的場所。</p> <p>中央研究院歷史語言研究所目前所藏明清檔案是民國十八年由藏書家李盛鐸手中購入，為原內閣大庫流出檔案的一部份。這批險成還魂紙的檔案在史語所七十個年頭，這七十年來，史語所變化不可謂不大，首先所址由廣州、北平、長沙、昆明、李莊、南京而至楊梅，最後落南港，而內閣大庫檔案的收藏也從八千麻袋變成了一百個木箱。</p> <p>光緒末，因緣際會，張之洞奏請將大庫書籍歸學部，並要羅振玉至內閣相助。羅振玉因此發現奏請待焚之檔案，向張之洞力陳該批史稿的重要，並獲同意將檔案移入學部。因數量眾多，以米袋裝檔案，陸續移部，存放國子監南學敬一亭，共計八千袋。</p> <p>民國二年，教育部設歷史博物館籌備處於國子監，接管這些檔案。但因缺乏重視，而任意處置。民國十年歷史博物館因經費短絀，就將檔案中比較破碎的部份，以四千元賣給同懋增紙廠作還魂紙。由於有人將部份檔案攜出紙廠，在市面上求售，為羅振玉知悉，於民國十一年二月又以一萬二千元把檔案買了去，再度將這些檔案由毀滅邊緣搶救下來；其後並雇人檢視，把檢得部份印成《史料叢刊初編》十冊。民國十三年，羅氏留下一部份，把其餘的檔案以一萬六千元又賣給了李盛鐸。後來李氏無力維護檔案才由史語所購入。</p>

項目名稱	欄位值								
整理註	<p>內閣大庫檔案整理沿革。</p> <p>1.第一階段：初整 民國十八年至二十一年年底。初步整理，進行分類，重要檔案完成簡明日錄。</p> <p>2.第二階段：清查 民國四十年四月至七十年。重新開箱，清查件數，辦理登記，進行刊印明清史料。</p> <p>3.第三階段：保存與出版 民國七十年十月至八十三年底。鈐印編號，影印複本，出版利用，以達到保存與提供研究的目的。</p> <p>4.第四階段：數位整理 民國八十四年一月至八十九年底。數位影像，光碟儲存，檔案登錄，提要撰寫，提供電子檢索工具，以資訊技術整合人工整理，促進研究發展。</p> <p>5.第五階段：國家數位典藏規劃 民國九十年始。參考國際標準，研擬檔案整理規範與細則，開發資訊技術，結合網際網路，進行數位典藏合作、利用與規劃。</p> <p>實體整理程序。整理組：a.掃灰、除虫；b.原件綴補、順稿、分類；c.原件蓋章登錄；d.原件封存上架。摘要組：a.原件內容摘要；b.摘要內容鍵入、建檔、校對。3.檔案掃描組：a.檔案原件掃描與數位攝影；b.掃描影像檔校對。</p> <p>目錄建置程序。1.metadata 與系統規劃：a.檔案特性分析；b.metadata 發展；c.系統需求單；d.系統規格需求書。2.檔案著錄：a.編排登錄；b.著錄；c.校對修改。3.權威檔著錄：a.主款目選擇；b.考證；c.著錄；d.校對修改。</p>								
參考文獻	劉錚雲。「舊檔案、新材料－中研院史語所藏內閣大庫檔案現況」。新史學 9卷3期(1998年9月)：頁135-162。								
版權聲明	中央研究院歷史語言研究所								
編目記錄	<table border="1"> <tr> <td>新增者姓名</td> <td></td> </tr> <tr> <td>新增日期</td> <td></td> </tr> <tr> <td>更新者姓名</td> <td></td> </tr> <tr> <td>更新日期</td> <td></td> </tr> </table>	新增者姓名		新增日期		更新者姓名		更新日期	
新增者姓名									
新增日期									
更新者姓名									
更新日期									

C.2 【層次－全宗】

例一 《明檔》

項目名稱	欄位值
層次	全宗
登錄號	Collection No.
	全宗號
題名	明檔
內容描述	根據徐中舒初步整理後的估計，史語所所藏內閣大庫檔案中，明代檔案數量很少，約有四千多件，多為康熙年間為修明史徵集而來，以啓禎年間的題行稿居多。
範圍與內容	明代檔案以行稿或題行稿居多，多屬於萬曆、天啓、崇禎三朝，而在目前整理好明代檔案中，萬曆朝有五件，天啓朝五十件，崇禎朝則有三千多件。其餘較早期的檔案多為敕諭，時間最早的二件是宣德二年，稍晚的幾件是正統、成化年間所頒的敕諭。也有幾件明代的冊檔，如武職衛選

項目名稱	欄位值
	簿等。至於實錄，有明內閣進呈熹宗實錄的散頁，以及憲宗、世宗、神宗實錄抄本散頁，數量不多。
責任單位	[不詳]
責任時間	[明]
傳記歷史註	
整理註	
參考文獻	劉錚雲。「舊檔案、新材料－中研院史語所藏內閣大庫檔案現況」。新史學 9 卷 3 期(1998 年 9 月)：頁 135-162。
版權聲明	中央研究院歷史語言研究所
編目記 錄	新增者姓名
	新增日期
	更新者姓名
	更新日期

例二 《清檔》

項目名稱	欄位值
層次	全宗
登錄號	Collection No.
	全宗號
題名	清檔
內容描述	內閣收貯的各項檔案，如題奏、制誥、史書、錄書、黃冊等；內閣本身的各項檔案；修書各館檔案；試題試卷及其相關檔案；潘陽舊檔。
範圍與內容	清代檔案內容計有：一、內閣收貯的各項檔案，如題奏、制誥、史書、錄書、黃冊等。清代中央各部及地方各省官員向皇帝言事都用奏本、奏摺與題本，這些題奏本章及其抄件在史語所檔案中居最大宗。除了了揭帖、題副、奏副外，史書、錄書也是題奏本的副本。史語所現有康熙朝的史書、順治朝吏、禮、工等科的錄書，以及明崇禎朝的吏曹章冊殘檔各若干本。也有一些題奏本的抄件，有單件也有裝訂成冊的。敕諭是皇帝頒給地方重要文武官員任官的憑證，史語所藏有這兩種敕書二千多件，然以坐名敕居多，另也藏有一些詔書與誥敕。史語所藏有各衙門年終奏銷清冊，也就是黃冊，約有二千餘冊。這類冊檔在史語所所餘不多，且多屬零散破爛的本子。此外，也有一些內務府各莊頭、園頭呈報的地畝清冊。也有二份丁口清冊民數，也就是所謂民數冊。另藏有十數本各旗所造報佐領緣由冊。在檔案中也看到一種題奏小許多的本子，約十八公分長，九公分寬，上面記漢了人名，封面多有「摺本」字樣。；二、內閣本身的各項檔案，本據《光緒朝會典》，內閣下設典籍廳、滿本房、漢本房、蒙古房、滿票簽處、漢票簽處、誥敕房、稽察房、收發紅本處、飯銀庫、副本庫、批本處等房處。史語所或多或少都藏有這十二房處的檔案。三、修書各館檔案。史語所所藏的多為實錄館、起居注館、會典館、一統志館、明紀綱目館與明史館等館所修諸書。不過完整的不多，且以稿本居多。另還有為修書而向各方徵集而來的檔案。四、試題試卷及其相關檔案。史語所藏有康熙、雍正、乾隆、嘉慶、道光、咸豐、同治、光緒等朝

項目名稱		欄位值
		的文武鄉、會試題名錄、試題、試卷三百多冊。另有一些西番館、回回館、百譯館、西天館、暹邏館、俄羅斯館等譯字生的試卷，試題都是相關外文字詞的漢譯。五、潘陽舊檔。為清入關以前的文書，多以無圈點的老滿文書寫，但史語所這一方面的檔案並不多，漢文資料也不多。
責任單位		內閣
責任時間		[清]
傳記歷史註		
整理註		
參考文獻		劉錚雲。「舊檔案、新材料－中研院史語所藏內閣大庫檔案現況」。新史學 9 卷 3 期(1998 年 9 月)：頁 135-162。
版權聲明		中央研究院歷史語言研究所
編目記錄	新增者姓名	
	新增日期	
	更新者姓名	
	更新日期	

2. 欄位簡易定義

每一欄位給予要著錄的資料值定義，以區辨不同欄位所代表的資訊內容及作用。

3. 欄位組織(至系統中)

依據欄位分析表及工作流程，將欄位組織至不同的系統模組中，並進行系統及功能分析。

4. 實作

實作過程中，因工作流程的調整或是著錄之難易與問題發現，進行欄位、欄位定義或工作流程、工作程度或系統功能的調整。

5. 著錄規則擬訂

因由多人同步進行檔案描述，過程中以討論方式，將每個人對 metadata 中每一的認知進行協調與統合，並考量系統設計及檔案特性，賦予制度化的著錄規則，重新以條文式撰寫欄位定義及著錄規則，以為遵循。

6. 欄位修訂

- A. 因內閣大庫檔案不再增長，不會再有徵集及採訪的行為，因此取消「契約檔」表單。
- B. 「參考文獻」表單納入主 metadata，不另建資料庫。「引用文」因系統功能未開發及資料蒐集未竟暫不列入。
- C. 因檔案的責任地歸屬考證有疑，因此採「寧缺勿錯」原則，刪除「責

任地」。

- D. 因人名、地名問題複雜，將「人名資訊」、「地名資訊」、「關鍵詞」分別定義，分立欄位。
- E. 「高廣」欄，因尊重清代文書制度之規定，詳填至小數第一位，取消採用圖書一般著錄標準。

參、機構間整合發展—合作與分享

一、機構間整合發展目的

1. 目的：分析各計畫之典藏內容性質，整併相似性質者，共同發展檔案核心後設資料（Metadata）架構，節省發展資源，並達成相互分享及整合利用。
2. 內容：依檔案全宗原則、來源原則及原始次序原則，研擬共通可對映之層次架構及核心欄位，建立交換及整合檢索機制。

二、整合程序

1. 提供 metadata 資料表單(典藏單位)
2. 瀏覽結構整合與 metadata 分析(檔案主題小組)
3. 建立檔案核心欄位及檔案核心欄位著錄通則
4. 進行各單位 metadata 與檔案核心欄位的 mapping
5. 撰寫各單位 DTD(內容發展與檔案主題小組)
6. 撰寫 XML 轉匯程式及核心欄位對映程式(各典藏單位之系統開發)
7. 硬體與系統建構含介面設計(內容發展、檔案主題小組、中研院資訊所)
8. 匯入資料(各典藏單位之系統開發、中研院資訊所)
9. 測試檢索介面及資料更新機制(內容發展、檔案主題小組、中研院資訊所)
10. 開放檢索(含線上申請帳號機制)

三、對各計畫 metadata 的影響

1. 增加對「finding aid」的描述—header(權責描述)。
2. 突破 TAG 的字面限制，尋找定義的整合與對映。
3. 修訂著錄規則(不僅只於資料的匯入匯出，需力求資料庫品質的一致性)。

4. 外部一致性；內部彈性。讓 metadata 就整體學術領域而言更適合檔案典藏，所有的檔案典藏 metadata 更具一致性與互通性；而對單一的典藏單位而言，可在核心 metadata 的基礎上執行應用，如此應用將更具彈性。
5. 促進各典藏單位對相同領域之典藏更密切的整合，如：史語所的內閣大庫與故宮的宮中檔及軍機檔同屬明清檔案，無論在基礎 metadata 建置可採用更一致的標準，在使用端也可提供同樣的介面及整合的資料庫，使用者不需再區分二個單位二個系統進行檢索。
6. 跨典藏機構間考慮檔案層次的界定。例如：將內閣大庫由原先機構內的「管理全宗」調整為「全宗」，與軍機檔及宮中檔屬同一層次界定。各機構的檔案資源若時代、產出單位等性質不同，則各有所屬全宗，既具整合意義又尊重各計畫各全宗的檔案編排次序。

肆、結論

Metadata 的研擬是數位典藏計畫中的基礎工作，更是數位成果展現的基礎平台，並具有為使用者導航與聚集的重要功能，因此 metadata 研擬的第一階段即著重在計畫目標、目的及預期對象的計畫環境分析，以及檔案本身特質的分析；而因為國際上在 metadata 及數位工作上的發展日新月異，各領域亦有專門組織研擬有專門規範與標準，因此，文獻探討亦是此階段中不可忽略的，亦可說 metadata 研擬的第一階段，其內容重點在於整合動機、資源及參考文獻。第二階段則進入系統分析，以前階段的分析結果，確認範圍、功能及目的，做為系統發展的基礎。第三階段，則真正進行 metadata 的欄位分析，這將是真正深入檔案內容的組織工作，不僅在系統架構上，亦在檔案內容及層次進行分析與關聯組織。Metadata 研擬的妥適性需藉由實作來進行測試與調整，此時的 metadata 未達完成的階段，僅能算是草稿，需經過標準化的程序。確認為標準的 metadata，並非永久適用，因此適時適切的評估與評鑑是必要的，此階段亦是尋求合作與整合的最佳階段。歷經上述程序，方能成就一份真正合適的 metadata。(metadata 研擬程序整理為表一)

表一 Metadata 研擬程序

程序	內容重點
計畫目標、目的及預期對象	Motive
檔案特質分析	Resource
參考文獻	Reference
系統分析	Extent Function

程序	內容重點
	Purpose
欄位分析	Structure Context Content
實作與研擬	Try and Experience Principle and Standard
修訂與整合	Review Evaluation
Metadata	Product

註 解

- [1] Arlene G. Taylor, The Organization of Information. 2nd ed.(Englewood, Colo. : Libraries Unlimited, 2004),139-157.
- [2] The Library Congress, ” Design Principles for Enhancements to EAD,” 01/08/2003,<<http://www.loc.gov/ead/eaddesgn.html>>(12/01/2004).
- [3] Society of American Archivists & The Library Congress, ” EAD Tag Library for Version 1.0,” 10/12/99,< <http://www.loc.gov/ead/tglib1998/tlover.html>>(12/01/2004).
- [4] Society of American Archivists & The Library Congress, ” EAD Application Guidelines for Version 1.0 : Appendix B: EAD Crosswalks,” 11/01/00,<<http://www.loc.gov/ead/ag/agappb.html>>(12/01/2004).
- [5] The Library Congress, ” MARC21 Concise Format for Authority Data,” 05/07/2004,<<http://www.loc.gov/marc/authority/ecadhome.html> >(12/01/2004).
- [6] International Council on Archives, ” ISAD(G): General International Standard Archival Description, Second edition,” 1999,<http://www.ica.org/biblio/isad_g_2f.pdf >(12/01/2004).
- [7] International Council on Archives, ” ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons, and Families, Second edition,” 2004,< <http://www.ica.org/biblio/ISAAR2EN.pdf>> (12/01/2004).
- [8] 中研院史語所，明清檔案工作室，2003年11月11日，
<<http://archive.ihp.sinica.edu.tw/mct/>> (2004年10月29日)。
- [9] 中研院史語所，珍藏歷史文物數位典藏計畫/管理模組/分支五/後設資料需求規

格書(0.4版)，2004年11月，<<http://saturn.ihp.sinica.edu.tw/~dahcr/model5.html>>
(2004年11月5日)。

[10]同註9。

參考文獻

Taylor, Arlene G. The Organization of Information. 2nd ed. Englewood, Colo. :
Libraries Unlimited, 2004..

The Library Congress." Design Principles for Enhancements to EAD." 01/08/2003.<
<http://www.loc.gov/ead/eaddesgn.html>>(12/01/2004).

Society of American Archivists & The Library Congress." EAD Tag Library for
Version 1.0." 10/12/99.<
<http://www.loc.gov/ead/tglib1998/tlover.html>>(12/01/2004).

Society of American Archivists & The Library Congress." EAD Application
Guidelines for Version 1.0 : Appendix B: EAD Crosswalks." 11/01/00.<
<http://www.loc.gov/ead/ag/agappb.html>>(12/01/2004).

The Library Congress."MARC21 Concise Format for Authority Data." 05/07/2004.<
<http://www.loc.gov/marc/authority/ecadhome.html> >(12/01/2004).

International Council on Archives." ISAD(G): General International Standard
Archival Description, Second edition."1999.<
http://www.ica.org/biblio/isad_g_2f.pdf>(12/01/2004).

International Council on Archives." ISAAR(CPF): International Standard Archival
Authority Record for Corporate Bodies, Persons, and Families, Second
edition."2004.< <http://www.ica.org/biblio/ISAAR2EN.pdf>>(12/01/2004).

中研院史語所。明清檔案工作室。2003年11月11日。
<<http://archive.ihp.sinica.edu.tw/mct/>> (2004年10月29日)。

中研院史語所。珍藏歷史文物數位典藏計畫/管理模組/分支五/後設資料需求規格
書(0.4版)。2004年11月。<<http://saturn.ihp.sinica.edu.tw/~dahcr/model5.html>>
(2004年11月5日)。