

第二十三章

論檔案檢索工具的權威控制系統

Archival Authority Control in Finding Aids

洪一梅

I- Mei Hung

中央研究院歷史語言研究所研究助理

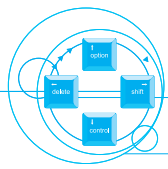
Research Assistant, Institute of History and Philology,

Academia Sinica

壹、前言

資訊檢索的目的在於利用電腦的儲存與計算能力，來幫助使用者由大量缺乏結構性的資料中，快速地擷取到所需資訊。（註1）而檔案資訊檢索是指擷取檔案資訊，是檔案利用的途徑，而最迅速的方式則是應用檔案檢索工具執行檢索，其中影響檔案利用效能最重要的關鍵在於檢索的內容。檢索的內容包含三個部份：檔案的資訊、描述檔案的資訊及檢索點。

現今，許多的資訊檢索系統申言可提供全文檢索，給予使用者錯誤的認知，誤以為只要有全文檢索，便可解決所有獲取資訊的問題，而事實上，在資訊檢索的領域中，電腦依據判斷的是字型符號，而非語意或概念的訊息。即當一個檢索概念以一連串的字彙型式輸入資訊檢索系統時，這些字彙即與存在資料庫中，代表此一符號的字彙比對。所以這是一外在型式的比對，並不是思想或概念的比對。（註2）但在人為思考的模式中，我們所需要的是一個概念模式的資料庫，因為許多原始資訊在產生時，也許因時間、空間或其它的因素，我們會以不同的字詞彙



來代表相同的涵義、事務或實體，但在日後的檢索過程中，也許我們所使用的檢索詞彙並非當初記錄時所使用的那一個，也因此，我們便無法獲得真正所需的資訊。因此，全文檢索的效能便有了疑慮，即使結合了切截、關鍵詞或布林邏輯等檢索方式，仍無法保證檢索品質的改善。

在國內外圖書資訊界多年資訊檢索的發展下，皆認同透過權威控制的一致性及參照關係，能顯著地改善檢索品質。在國際上，檔案資訊檢索無論是過去書本式目錄的查尋或是因應科技所發展的資料庫檢索工具、線上檢索工具，權威控制都是其所重視的功能，從索引典的編制到權威檔的建置，已有多數發展與應用的經驗。

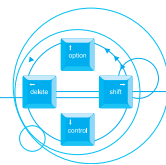
本文嘗試引介資訊檢索中權威控制的概念，並從內容與系統的建置面，及使用者的需求，來說明權威控制系統在檔案檢索上的應用。

貳、權威控制的概念

一、意義

權威控制是描述目錄控制（圖書館界稱「書目控制」）必要的工作，決定目錄檔中名稱、標題、參照款目及其它權威標目之間關係的過程。（註3）在描述系統中，用以維護標目之一致性；在線上描述系統中，藉由系統連結與指引功能，辨識匿名作者及追蹤名稱的各種更替情形；並能將相關主題、作品聚集一處，保持目錄中標目的一致性與一貫性；維護資料庫中檢索項詞彙型式一致性，進一步將資料庫中的名稱、作品與主題間之關係標示出來，其目的在促使描述記錄與權威記錄連結在一起，以輔助使用者檢索到所需的資料。（註4）

由以上可知，權威控制是藉由標目的一致性，以及建立參照關係將不用與相關的標目連結起來。獨特性、標準化及連結性是權威控制的特性。綜合中外學人對權威控制的看法，權威控制主要是在對每一個檢索款目，包括：個人名稱及團體名稱、劃一題名、標題等進行一致性及建立參照的工作，使得目錄上所有的檢索點皆具有獨特性，不會產生混淆，同時又能使得與這些標目有關的詞彙，產生關聯性，讓使用者由任何一個詞彙查尋，皆可獲得完整的資料。（註5）



二、功能

權威控制在描述標準中兼具了資料內容標準及資料價值標準的功能，在資料內容標準中是提供資訊輸入的規則；而資料價值標準則是在特定的元素中，提供可接受的術語、名稱、字母數字代碼或其它特定的資訊，以指出特定的實體，其目的除了增加資料的完整性外，更在促進查檢的精確性。（註6）

由此我們可以歸納權威控制的功能如下：

（一）查尋的功能

幫助使用者了解館藏所收錄的資料及可獲取的途徑。爲了使目錄可以達到此目的，必需提供產生者、Title、或主題作爲檢索點。如此讀者可使用簡短的詞彙進行查尋。而藉由權威控制所建立的檢索標目，便提供了使用者比對的來源，以此進一步獲得原始的資料。

（二）聚集的功能

此功能所強調的是標目的參照結構，由已知的產生者、Title、或主題來展示館藏。權威控制對所建立的標目，不僅止爲單一詞彙，更建立了一關係結構，使使用者可藉由關係結構中的任一詞彙，獲得相同、相關、相似、縮限與擴展的所有原始資料。這樣的聚集，是提供使用者在檢索時一個智慧性的選擇。

（三）連結的功能

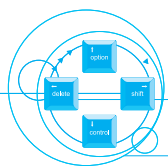
權威控制所建立的權威記錄並非完全獨立而存在，其與描述記錄之間是相依相存的。描述記錄中所完全沒有包含的資訊，權威記錄也不需包含，因爲使用者是不可能查尋或查尋得到沒有包含的資訊。反之，當描述記錄中所包含的資訊，可藉由權威控制進行關係及實體名稱界定的分析，當由同一產生者所提供的多種實體，在描述記錄中會有多筆記錄，但在權威記錄中只會有一筆產生者的記錄，包含多個產生者相關的標目，藉由此權威記錄，連結了所有此產生者所提供的實體資訊。

（四）評估的功能

資訊檢索系統的檢索效率評估在量化的計算上，主要有以下二點：

1、檢準率（precision ratio）

檢準率是指在某一查詢條件下所檢索出相關文獻數目占檢索到資料的



總數，包括相關與不相關的比率。（註7）其公式如下：

$$\frac{\text{檢索到相關資料的數目}}{\text{檢索到資料的總數}} \times 100\% = \text{檢準率}$$

（包括相關與不相關）

檢準率是評估資訊檢索系統的重要指標之一，其價值在於藉由此種分析可指出系統的缺失與改進之可能方法。

2、檢全率（recall ratio）

檢全率係指在某一查詢條件之下所檢索出之相關文獻數目占資料庫中相關文獻總數的比率，意即將系統檢索相關文獻成功的程度予以量化表示，其公式如下：

$$\frac{\text{檢索到相關資料的數目}}{\text{資料庫內所有相關資料的總數}} \times 100\% = \text{檢全率}$$

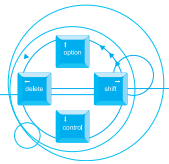
（包括檢索到與未檢索到）

檢全率是另一評估資訊檢索系統的重要指標。（註8）而權威控制便是改善檢索品質的良好方式，亦可提昇檢索效率。

三、作業程序

權威控制的作業程序，胡歐蘭教授在「圖書館學與資訊科學大辭典」中歸納了五個步驟：

- （一）建立權威記錄。
- （二）匯集權威記錄組成權威檔。
- （三）連結權威檔與描述記錄檔，組成權威控制系統，或由描述記錄檔中建立權威檔，並加以修正連結成為權威控制系統。
- （四）權威檔與權威控制系統之修定與維護。
- （五）權威檔與權威控制系統之評估。（註9）



參、權威記錄與權威檔

在權威控制的作業程序中，建立權威記錄與權威檔，是產生權威控制系統的先決條件。在權威記錄與權威檔中，其結構與內容應包含以下各部份：一、權威標目：個人、團體、地名、主題等等。二、各權威標目之元素（**element**）與元素值：即相關款目，例如參照款目、關聯款目等等。三、元素填入資訊（元素值）之規則：即款目及款目值。四、元素間之關係與結構：即指不同款目間。五、與描述記錄間之連結：即指控制執行。六、呈現與使用對應方式。

每一筆權威記錄依結構、內容著錄完成後即組成權威檔。

一、權威標目

（一）權威標目的選擇（**from** 檔案）

權威控制的主體在於權威標目。資訊檢索者的異質性與異質思考邏輯造成檢索點使用的差異，而每一個檢索點的變異也造成了權威控制的必要性，因此針對每一個檢索點進行分析以執行權威控制，成為檢索效能的重要影響因素。而在進行檢索點分析時，最優先的決定便是「何者要成為權威標目」，也就是權威標目的選擇。

考量權威標目的選擇，有三個方面：

1、使用者（檢索者）的需求

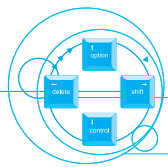
包含使用者對檢索工具及檢索內容的起始程度、使用者的資訊尋求行為及尋求能力、使用者的檢索目的等。

2、標目的種類

資訊描述摘錄自內容的資訊值，皆成為使用者的檢索點，而內容包含了人、事、時、地、物，這些不同的檢索點即是使用者在進行檢索行為時，可參考的檢索典中的標目，再依其特質予以區分種類，成為權威控制的標目，以利再分析。

3、標目的變異程度

目前最常見的標目種類有人名標目、地名標目、團體標目、主題標目…等，這些不同的標目各有其變異的情況與變異的程度，皆影響了權威控制的



方式。

以下舉例數種標目種類，並說明其變異程度：

1、人名標目：

個人名稱在檔案資訊中是非常重要的內容，因此選擇做為標目對檢索與利用及資訊加值的目的而言，

是必要的。以中文的個別人稱標目來說，可能有的變化有：筆名、改名、譯名、原名、別稱、字號、諡名、冠姓……等等。

2、地名標目：

地名標目可能有的變化，分以下三個立基點來說明：

- (1) 時間橫軸－在同一時間點時，有原地名名稱、地名簡稱、地名別稱、俗稱與習慣稱法……等。
- (2) 變異縱軸－在不同的時間點，地名名稱在原名、簡稱、別稱、俗稱與習慣稱法……上也都可能再次改變。
- (3) 空間投射－因不同的時間點，即使名稱一樣，但實際名稱所代表的領域範圍仍有可能不同，可能行政隸屬改變、可能行政所轄範圍擴大或縮小。

舉例：

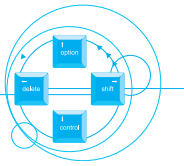
改名：山西省大同府，順治五年更府名為陽和府，順治八年復名大同府。

改隸：興縣原屬山西省太原府，雍正二年往屬山西省保德直隸州，雍正八年還屬山西太原府。

新興：雍正三年置朔平府轄左云縣、平魯縣，同時朔州、右玉縣（裁右玉衛，置）、馬邑縣（嘉慶元年六月裁入朔州）由山西省大同府改隸入。

區域改變：同上，山西省大同府範圍縮小。消失：馬邑縣，嘉慶元年六月裁，入朔州。

同名異地：清初，定海縣屬浙江省寧波府，康熙二十六年改定海縣為鎮海縣，于舟山島地方置定海縣來屬。



3、團體名稱標目：

由政府機關、民間團體、公司、企業所產出的檔案，團體名稱關係了其檔案的編排方式與結構。因此，團體名稱亦為重要的標目之一。其變化有：團體名稱、簡稱、改稱、合併、裁撤、改隸……等。

4、主題標目：

是針對檔案內容、性質、形制的再分析，運用主題標目更易於深入檔案的內容與研究。

5、其它：

官名權威標目：在制度史與官制的研究上甚為重要，其變化有異稱、簡稱、俗稱、異名同官、權限更動……等。

民族名稱權威標目：在人類學的研究上有其必要性，其變化如下：

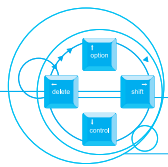
- (1) 民族分支：納西族一支為摩梭人。
- (2) 自稱或他稱，民族識別後併為一族：儼、山蘇、撲喇……等併為彝族。
- (3) 自稱或他稱及歷史上名稱的不同，如下表：

表一 自稱或他稱及歷史上名稱的不同

族名	歷史名稱	自稱	他稱
傣	【漢晉】滇越、揮、檀、越賧、僚、鳩僚 【隋唐】茫紹、金齒、面、雕題 【宋】金齒、白衣 【元、明】金齒、百夷、爨夷 【清、民國】擺夷、爨夷、歹摩	傣泐、傣那、傣繃、傣雅	傣族、水傣、漢傣、擺夷、比搓、米搓、阿搓、夏姆、勒薩姆

(二) 權威標目的管理 (from 管理機制)

權威標目的管理，是各檔案典藏單位針對其典藏檔案選擇了權威標目之後，



分析該館檔案權威標目的特性，擬訂建置方案前，所決定權威標目的管理方式，各館因性質不一，因此權威標目的管理方式亦有所不同，尤其針對整合多元資料性質的典藏館，其權威標目的管理方式有更大的彈性，例如中研院史語所，其典藏文物有一般圖書、善本圖籍（rarebook）、檔案、文物、金石拓片……等等，其權威標目的管理也許無法畢其功於一役，但肯定其若欲達最佳的管理應有最具彈性的作法，甚至在同一館內因應不同的典藏文物即選擇不同的權威標目管理方式。

- 1、使用已建立的權威檔，由其中選擇典藏範圍內的權威標目。
- 2、依據相關的標準建立新的權威標目。
- 3、在已建立的權威檔中再新增權威標目，此方式要在合作的前題下或是權威檔建置為領域內開放分享的機制。
- 4、維護權威資料。在檔案描述檔/目錄檔中，針對所選擇的標目，建立權威資料或是建立參照關係，以避免重覆。
- 5、直接在權威記錄中指定使用的權威款目做為標目，以促進使用。（註10）

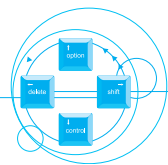
（三）權威標目的建立

國際簡例—人名或團體名稱權威標目的建立：

- 1、首要在檔案資料中找出人名或團體名稱。
- 2、檢查公用目錄。若有資料即記錄，若無資料則執行下一步驟。
- 3、檢查LCNAF。若有找到即記錄其所使用的權威標目，完成權威標目的建立；若在LCNAF 中所找到的標目格式不同於原始建立的標目，無論是在原檔中或公用目錄中，則新增參見至原檔或公用目錄中；若未找到，則執行下一步驟。
- 4、依據AACR2 建立標目，包含原始標目、描述使用的標目、參考來源的標目。在此步驟中，檢查任何可獲得的目錄資料庫或是RLIN。
- 5、若名稱有多種型態，則建立完整的款目結構，成為權威記錄或是為LCNAF 新增資料。（註11）

二、權威記錄

（一）權威記錄的元素



權威記錄包含以下各項元素，除定義各資料值的意涵外，並界定了各資料值在權威控制的功能與角色：

- 1、主款目：元素值在權威控制作業中作為控制詞彙。
- 2、參照款目（含見、參見）：元素值在權威控制作業中作為非控制詞彙。
- 3、關聯款目：元素值為主款目的擴大範圍詞彙或縮小範圍詞彙。
- 4、其它款目：元素值可能為限制不使用的詞彙。
- 5、參考資料：記錄相關詞彙的來源出處。
- 6、註記：其它不能在上述元素中闡述的資訊內容。

（二）權威記錄的著錄

為維持每一筆權威記錄的品質，需有一致的著錄原則，權威記錄的著錄原則用來規定每一筆權威記錄所包含的元素/欄位（element）其資訊（元素值）的著錄方式，以維持權威記錄的架構與一致性，一般稱為著錄規則，權威記錄著錄規則的基本項目內容應包含：

1、發展沿革：

說明著錄規則的發展源由，及其發展與認證的過程，並詳述參與發展的貢獻單位。

2、發展背景與目的：

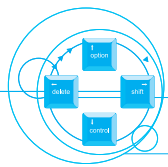
說明支援著錄規則發展的相關資源與相關標準，以及對著錄規則研擬的基本共識，並陳述發展目的、效益及相關發展原則。

3、使用方法：

包含著錄規則中的架構、組織原則，及其研擬原則、標準的建構標準與內容使用方法，以導引權威記錄的著錄。

4、內容詳列如下：

- （1）專用名詞解釋：權威控制中專用名詞的定義。
- （2）權威控制領域：建立權威標目与其它標目間的連結與架構。
- （3）資訊領域：此領域中的各個元素資訊（元素值）即是用來描述單一名稱權威標目的權威記錄。



(4) 註記領域：描述權威記錄如何被建立，包含參考資源的註記及資料維護的註記等。(註12)

三、權威記錄編碼

對權威記錄進行編碼，主要目的在於使權威標目成為已編碼的描述資訊的檢索點，針對權威標目所形成的權威記錄，其中含多元款目，再界定某些款目為權威款目，或參見至權威款目，也就是界定款目間的相互關係。以編碼導引資訊（元素值）的運作。編碼最大的優點可以於系統架構的索引典中進行交叉註記（node labels），以顯示權威、並列、相關等款目的邏輯基礎並可立即控制或呈現。這是在書面的索引典中難以設計的。(註13)

權威記錄中目前常見的編碼規則如下：

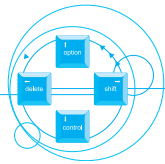
(一) MARC 21：<http://lcweb.loc.gov/marc/authority/ecadhome.html>

MARC21 針對權威記錄中的所有款目有完整的編碼規則，其做法視權威檔為一獨立的資料庫，再與主描述檔間去進行權威控制。

(二) EAD<control access>：<http://lcweb.loc.gov/ead/ag/agcre5.html#sec1>

EAD 將權威記錄視為其主描述檔中的其中一段，用以做為檢索詞彙的控制，針對不同的標目，運用不同的標記：

- Personal Name <persname>
- Corporate Name <corpname>
- Family Name <famname>
- Geographic Name <geogname>
- Name <name>
- Genre/Physical Characteristic<genreform>
- Function <function>
- Occupation <occupation>
- Subject <subject>
- Title <title>



肆、檔案的權威控制

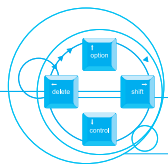
一、檔案權威控制的必要性

在檔案描述資訊中，權威控制提供了一致性的詞彙選擇，以系統的控制來彌補人為及時空變遷所可能造成的差異。在資訊檢索系統中，無論是布林邏輯、切截、全文或其它線上檢索功能都難以取代權威控制系統所提供的連結與指引功能。（註14）

檔案編排與描述所強調需遵循的來源原則及原始次序原則，使得著錄的詞彙在時間與空間差異下更顯其異稱之重要意義，例如團體名稱（或組織名稱）的變化。因此權威控制在檔案資訊檢索中將扮演極為重要的角色。不僅用於建立與維護檔案目錄資料庫，更促進檔案資訊檢索品質的提昇。

因檔案的脈絡層次特性，以下列四種名稱標目說明權威控制的必要性：

- （一）人名：檔案的產生者、保管者、捐贈者，對於檔案的存在皆有其特殊的意義，也是檢索者檢索檔案的重要檢索點之一，而這些可以為人物，因此人物名稱成為權威控制中的重要標目。
- （二）地名：檔案是歷史事實的記錄，內容牽涉了人、事、時、地、物，因此地名成為檢索內容的重要檢索點，且有時也成為標題的一部份，然而地名歷經時代的變遷，行政區的轉換，往往真正的地區未變，但名稱變異及範圍變異甚大，因此，進行地名的權威控制實為重要，對有典藏地圖檔案的檔案館尤為之最。
- （三）團體：前述檔案的產生者、保管者、捐贈者可以為人物，亦可為團體，尤其檔案的產出機關或單位，在沿革中有歷經改名、改隸、合併、裁撤、精簡等過程，檔案是重視來源原則的，因此機關團體或單位即成為重要的權威控制標目。
- （四）subject 標題：標題對檔案來說是重要且特殊的檢索點。在圖書資訊學中，標題是進行內容的分析，其架構在知識的分類體系下，但檔案的編排非以知識體系來分類，重視的是來源原則與原始次序原則，此依據是為保持檔



案於產生時即具有的意義與憑證，因此標題所進行的內容分析即不能強加以後天的分類，應以其原有脈絡體系中的主題名稱，也許是組織架構、也許是業務功能、也許是個別事件來建立適合檔案的標題表，再以時代、地理等複分方式，提供新的詞彙或不同的詞彙。檔案人員可以嘗試集結同領域的檔案或檔案全宗，建置共同使用的標題表，但不應妄想建立適合所有檔案使用的統一標題表。

二、檔案權威控制的目的

對檔案檢索的效益而言，檢索點格式與內容的控制是非常重要的，其影響資訊檢索行為與獲取的結果，同時也受資訊檢索行為的反向影響，因此是一互動性的控制。有效地呈現檔案描述的元素（element）資訊。檔案描述是對單元檔案本體的描述與分析，權威控制是藉由元素資訊以連結、排除的方式精確地獲得檔案群的橫向分析結果，即為具相同特質（元素值）檔案的聚集，而這樣的結果也代表了眾多單元檔案間的關連、層屬或交錯關係。

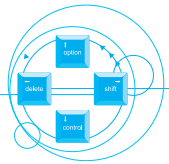
跨越檔案描述資訊本身的時空限制，也跨越管理層在檔案實體典藏及單一系統或單一資料庫對資訊值容納的時空限制。對特定元素的資訊進行一致性、專屬性的分析，以提供特定元素資訊的分享與合作，如人名元素。使特定元素中的元素值透過分析後，將縮小、擴張、同義、相反等關係清楚建立。

由上可知，檔案權威控制工作的目的有三：

- （一）跨越時間整合歷史上與層次上的差異與關連。
- （二）藉由分享，避免重覆並取得一致。
- （三）促進資訊系統的發展，並支援資訊檢索的使用。

三、檔案權威控制政策的擬訂

檔案權威控制政策是針對各項權威控制作業需求，並配合檔案館作業要求所制定的工作原則與目標。除思考檔案館內部作業要求外，應包含館藏範圍、使用者需求與資訊檢索行為、館員素質、系統功能與運作評量、名稱種類與選取來源、資料語文與格式及管理需要、工作規劃與分配、結果評鑑等各個層面的考量。（註



15)

伍、檔案檢索工具中的權威控制

一、關聯

權威控制系統的建置與「檔案描述」、「權威檔描述」及「檢索工具」三者之間實有不可分割的密切關係。

(一) 說明

依所界定的檔案層次，進行內容資訊、載體資訊、管理資訊等方面的分析與記錄，稱為「檔案描述」。檔案描述的架構、元素與元素值具體呈現檔案的價值。

使用者如何透過檢索工具，獲得檔案描述的資訊值，檢索工具分析檔案描述中的元素，決定開放提供的檢索點，透過檢索點，局部對應使用者輸入的需求與檔案描述中存在的資訊，對應成功的描述記錄筆數即代表使用者檢索的回現率。

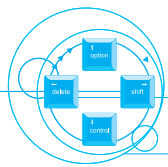
對應的基礎，在於「具有溝通能力的語言」，亦指使用者輸入需求的語言必需與檔案描述中元素值的語言彼此具溝通的能力。檔案描述者與使用者必然存在的差異，降低了檢索工具的檢全率與檢準率。權威檔描述的目的在於建立溝通的管道，如同語言編譯器，將使用者輸入需求的語言轉化成檔案描述的語言，使檢索工具得以在檔案描述中找到符合使用者需求的資訊。

「具有溝通能力的語言」包含了二個層面的格式：一為元素；二為元素值。而描述規則是用來定義元素與元素值格式的工具，亦為決定溝通方式的核心工具。

(二) 實例

1、ICA

International Council on Archives Committee On Descriptive Standards 在2002 年1 月公開發表了一份檔案檢索工具的指南：Guidelines for the Preparation and Presentation of Finding Aids。文中陳述了檢索工具建置的許多原則，並說明了檢索工具（Finding Aids）和「國際檔案描述標準（ISAD(G）」及「國際標準檔案名稱權威記錄：團體、個人、家族（ISAAR(CPF）」之間的密切關係。以檔案檢索

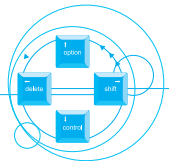


工具分析三者關係：

- (1) 思考「國際檔案描述標準」的目標，建置檔案檢索工具的原則有：
 - a、提供檔案資料的檢索利用。
 - b、提供一個準確的、一致的、自我詮釋的檢索工具。
 - c、應用多元層次描述的原則，有效呈現檔案資料的層次與內容。
- (2) 檔案檢索工具的內涵：
 - a、檔案的描述資訊。
 - b、檔案的權威檔、索引、表單及索引典等等。
- (3) 檔案檢索工具的目的：
 - a、建立一個提供檔案資訊服務的規格。
 - b、將檔案資料特色化，轉為有價值的檔案資訊。
 - c、做為檔案資料與使用者間的溝通橋樑。
- (4) 檔案檢索工具的建置：
 - a、週邊元素：手冊、自動化環境。
 - b、內容元素：提供機構智能管理的機制；提供使用者智能檢索的機制；檔案檢索工具的型態（線上或資料庫或紙本）；檔案資訊的描述方式與其呈現方式；時間的註記等。
 - c、架構元素：呈現檔案描述的多元層次；導引關聯層次的資訊獲取（可利用說明或符號註記）；定義描述的欄位；支援多語言的描述。
- (5) 檔案檢索工具的功能：
 - a、促進檔案資料內含資訊的查尋與交換。
 - b、促進權威資料的分享。
 - c、促進不同檔案典藏單位提供整合的服務。

以「國際檔案描述標準」及「國際標準檔案名稱權威記錄：團體、個人、家族」分析三者關係：

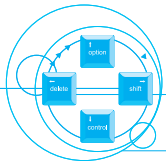
- (1) 檔案檢索工具的內含資訊應包含ISAD（G）中的26 個/組元素。並區分出必要元素及其它可選擇的元素。



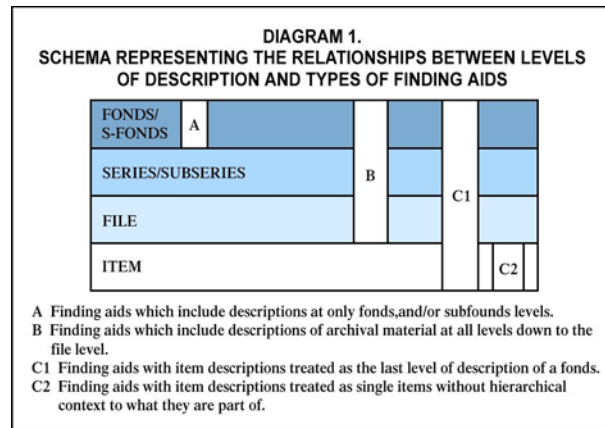
- (2) 檔案檢索工具所提供的檢索點來自於ISAD (G) 的描述元素，並依據ISAAR (CPF) 的格式使其對描述內容執行有效的檢索。也就是，檔案檢索工具的檢索點分析自ISAD (G) 的描述內容，而其格式則由ISAAR (CPF) 來定義。
- (3) ISAD (G) 與ISAAR (CPF) 對檔案檢索工具的基本影響：
- a、呈現的檔案描述層次。
 - b、每一個層次的檔案描述元素。
 - c、檢索點源自分析檔案描述元素。
- (4) 檔案檢索工具的多元層次描述架構 (Diagram 1)，依循ISAD (G) 的多元層次描述原則。
- a、CLASS A：全宗層次 (FONDS/SUBFONDS) 僅含高階層次的描述資訊，目的在提供有關產生者與檔案資料的摘要性資訊。在此層次，可整合一至多個存藏所的館藏資訊，包含檔案內容資訊及機構典藏與服務的相關機制控管。
 - b、CLASS B：全宗至案層次 (FONDS down to FILE)
檔案描述的基本單位在於系列，檢索工具可以選擇要不要包含卷冊、箱、案等層次的描述資訊。次層次的描述資訊需包含所有可提供檢索的元素值。
 - c、CLASS C：件的層次 (ITEM) 可分為二種，一為將件視為整個全宗中最基層的單位，予以描述；一為將件視為個別獨立的實體，僅提供件的描述資訊，不提供與全宗或其它層次間的關係。
- (5) 檔案檢索工具的建置需符合檔案描述的相關標準。(註16)

2、中央研究院歷史語言研究所

- (1) 明清檔案著錄規則定義了明清檔案描述的元素、元素所填錄之資訊值及填錄之格式與方法規則。在此時，即決定了檢索點、資訊值提供檢索的型態 (控制或自然語言) 及描述資訊在檢索工具中預定呈現的控管機制與方式。
- (2) 明清檔案人名權威檔著錄規則分析檔案描述所提供有關個人名稱的元素與元



素值，並考量使用者可能使用的檢索詞彙與檢索方式，去定義個人名稱的主款目、參照款目、關聯款目及其它資訊，並決定其在檔案檢索工具中執行控制的方式。

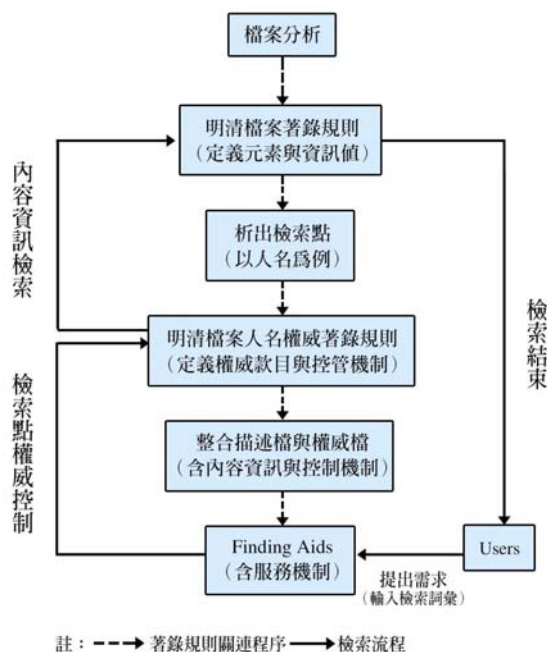


圖一 檔案檢索工具的建置需符合檔案描述的相關標準

資料來源：同註16

- (3) 明清檔案檢索系統提供的是明清檔案著錄規則所定義的描述記錄（含層次、內容元素、控管參數）；提昇檢索效率是藉由明清檔案人名權威著錄規則所定義的權威記錄，執行權威控制機制；呈現則是由描述記錄與權威記錄中的控管參數與控制機制來決定使用者的檢索結果及資訊內容。

(4) 圖說：



圖二 檔案分析

二、應用方式

檔案檢索工具權威控制的應用，主要決定於元素確認後、系統開發前的系統再分析，由系統開發人員介入，主要目的在於確認的資料庫模式下，系統內部將元素值轉換為索引詞彙表單的產出與運作方式，以及權威檔在資訊檢索系統中與描述資料檔實際進行檢核、比對、分析的方式，串連成真正的系統運作，達到對檢索者檢索行為、檢索意念的解析，輸出真正的需求資訊。

檔案檢索工具在提供檢索服務時，其權威控制的方式有下列數種：

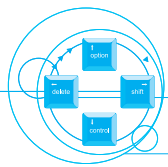
(一) 使用者選擇：線上詞彙展示

線上詞彙展示，是由使用者瀏覽系統所使用的控制詞彙，選擇適當的詞彙進行檢索。（註17）這樣的權威控制是屬於事前的控制，也就是在進行檔案描述時，針對欲提供檢索的元素（欄位）值，使用控制詞彙的方式著錄。在權威檔中，將每一檢索詞彙的控制詞、非控制詞、相關詞進行關聯，線上展示時，即為提供控制詞彙的使用指導及詞彙間關係的表達。線上詞彙展示一般有三種方式：

1、按字母順序排列/按筆畫排列

將詞彙按字母順序排列（外文）或筆畫順序排列（中文）展開。例：

<http://dsppsd.pwgsc.gc.ca/Thesaurus/index-e.html> (展開並直接串連結



果)。

2、 檢索式

提供詞彙的檢索，使用者利用檢索方式查得詞彙，再展開瀏覽相關詞彙及確認提供之檢索詞彙。例：

<http://www.getty.edu/research/tools/vocabulary/tgn/>。

3、 分類式

有些詞彙表，因循過去紙本形式，或是詞彙仍在審訂中，採用分類展示的方法。例：<http://datas.ncl.edu.tw/catweb/2-1-16.htm>（審訂中）。

（二）自動控制式：詞彙自動對應

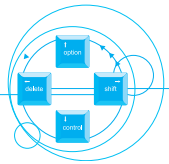
自動控制式即指詞彙自動對應，使用者不必考慮所使用的檢索詞彙是否為控制詞彙，（註18）一般自動控制有二種方式，但二者都必須仰賴一份完整的詞彙對應表來進行詞彙的自動對應。說明如下：

- 1、檔案描述中提供檢索的元素值，以控制詞彙的方式進行著錄。權威檔中主要款目皆為控制詞彙，參照款目為非控制詞彙（自然語言）及關聯詞彙。使用者輸入任何檢索詞彙，系統進行所有詞彙（含控制、非控制、關聯）的對應，找出其所使用的控制詞彙，以控制詞彙進行檔案描述資料庫的檢索，獲得結果，提供予使用者。
- 2、檔案描述中提供檢索的元素值，依照著錄原則進行著錄。權威檔中主要款目含控制詞彙與非控制詞彙，參照款目為關聯詞彙。使用者輸入任何檢索詞彙，系統對應出該檢索詞彙之控制詞彙、非控制詞彙甚至關聯詞彙，以這些詞彙進行檔案描述資料庫的檢索，獲得結果，提供予使用者。此方式可配合提供fuzzy index，以避免提供許多非使用者本意的查詢，例如：使用者的檢索詞為精確檢索（100%）；控制詞彙與其它非控制詞彙為完全檢索（80%）；關聯詞彙為擴充檢索（50%）。

（三）半自動控制式：詞彙自動對應－使用者選擇

半自動控制式主要目的在使檢索的結果更契合使用者所需，尤其當有同詞異義時。其方式分為二個階段：

1、自動對應階段



系統自動對應詞彙表，如在詞彙表中僅有一筆詞彙款目且無關聯詞彙，則系統自動檢索，將結果回饋給使用者；若在詞彙表中有多筆詞彙款目，或具有關聯詞彙，則進行第二階段。

2、使用者選擇階段

若有多筆詞彙款目或具有關聯詞彙，則系統展示詞彙款目或相關詞彙，予使用者自行選擇，點選後獲得該詞彙或該多詞彙之檢索結果。此方法對於使用個人名稱檢索尤其有助於增加其精確率，因為在同一資料庫中，同姓名卻不同人的情況是有的。

陸、檔案權威控制系統的建置

檔案權威控制系統的建置程序，從需求提出、資料庫的開發、內容著錄到系統的運作，說明如下：

一、需求分析與整合

（一）權威標目選擇

分析典藏，確認檔案檢索工具中的資訊內容。存在的資訊內容，有提供檢索的資訊內容，才有進行權威控制的必要。因此，從資訊內容中區分標目的種類、依不同種類選擇權威標目是權威控制作業的起始。

（二）控制詞彙分析

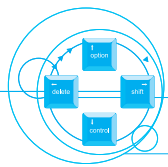
不同的資訊描述內容，在詞彙控制上有不同的要求。針對已選擇的權威標目進行詞彙分析，確定分析原則以決定控制詞彙、非控制詞彙、相關詞彙及其它資訊。

（三）使用者分析

使用者對檢索工具及檢索內容的起始程度、使用者的資訊尋求行為及尋求能力、使用者的檢索目的決定權威控制的內容與運作方式。

（四）權威資料庫分析（含權威款目、權威記錄、權威檔）

依選擇的標目及詞彙分析的結果，決定權威款目及權威記錄的欄位與定義，並依欄位彼此間的關聯、層屬、連結，組成權威檔，完成權威資料



庫的結構分析與Entities-Relationship Model 的規劃。

（五）系統需求分析

此程序是爲了界定系統架構與系統功能。分析內容包含著錄、管理、檢索、權威控制的執行及條件，並擴及與檔案檢索工具間的運作方式。

二、權威資料庫開發分爲管理端、使用端及系統運作端：

- （一）管理端：含著錄介面、著錄人員管理、人員權限控制、參數代碼維護、歷史記錄管理、統計與報表列印。
- （二）使用端：含資料庫的獨立檢索、資訊顯示及瀏覽。
- （三）系統運作端：自動索引表單。

三、內容著錄

- （一）標目來源的界定：依Collection 而不同。

針對不同的典藏品有不同的描述內容，而即使相同的標目種類，例如人名標目，因典藏領域或典藏品的差異，造成標目在不同的描述內容中有明顯不同且特殊的意義，因此界定標目來源，即成爲各領域在進行權威資料庫內容建置的首要工作。相同學術領域，亦可以時代、組織部門或專門特性去區分建置相同的權威標目種類，促進不同典藏單位的分工合作。

- （二）參考資料的選擇

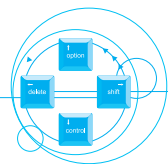
在不同學術領域中，針對權威標目有不同的著重點，在補足相關資料上，除原始主要著錄來源外，對參考資料的選擇，應著重其深度、廣度與考證性。

- （三）著錄規則的擬訂

著錄規則的擬訂，目的在有效地控制資料庫的品質與一致性，更可做爲人員訓練的教材。

- （四）著錄

在著錄過程中，除對各領域人物特性的了解外，在查找與驗證資料上



要有一定的能力。著錄需依據著錄規則來進行，著錄人員於工作過程中應常提出不同的案例進行討論獲得共識，彼此學習在著錄工作上的經驗。

（五）校對

校對是資料庫建置的必要程序。必需謹慎以對、嚴格執行，務求資料內容的正確性。

四、系統整合運作

權威控制系統中的索引表單與檔案檢索工具中作為標目來源的檢索點交互查核、連結及整合檢索的系統運作方式。

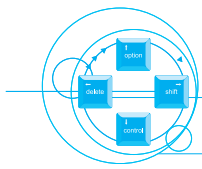
五、其它

在整個建置過程中，有相關文件規範的產出，這些文件規範與程序說明不僅可為日後相關單位進行人名權威檔建置的經驗參考外，也建立了細部作業的執行規則與工作標準，確立每一個細節成果的品質與一致性：

- （一）權威標目元素（**element**；欄位）定義。
- （二）權威資料庫系統規格書及系統分析資料。
- （三）權威資料庫建置流程說明書：從分析、系統建置至著錄、應用的流程說明。
- （四）權威標目著錄通則。

柒、結語

在圖書資訊檢索系統中，有許多的詞彙控制工具，如：分類表、標題表、索引典等，用來選取標準化詞彙做為檢索點。該詞彙控制工具即為一種權威檔，主要目的乃避免選用非權威形式的標目，及維持權威標目之一致性。在檔案內容中，最常使用的檢索詞彙類型，包含人、地、主題等，為了有效發揮資訊檢索的效能，權威標目在權威資料庫內應有一致性的著錄標準；（註19）而權威控制系統與檔案檢索工具間應具有查核、連結、整合檢索的完整功能；而檔案檢索工具對使用者應具有親善的檢索與瀏覽介面，並因權威控制系統的整合運作，有效提昇使用

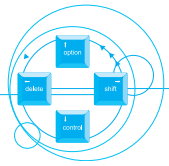


者資訊尋求的檢準率與檢全率。

【原刊載於檔案管理局出版之檔案季刊（九十二年六月）第二卷第二期】

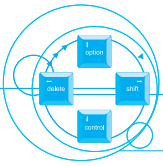
詞彙定義：

- [1] 款目（Item）：描述實體的架構元素，每一元素與其元素值在資料庫中稱為款目。
- [2] 記錄（Record）：由多個彼此間有關連，用來描述同一實體的款目所組成。
- [3] 檔（File）：由一筆以上相同結構的記錄所組成。
- [4] 資料庫（database）：由一個檔或多個檔所組成。
- [5] 標目：目錄檔中內容分析或其它分析中可用以做為檢索點的元素與元素值。元素選擇做為指定檢索點，即成為標目。
- [6] 元素（element）：指檔（file）中所包含的資訊欄位。
- [7] 元素值：元素中的資訊值。



註釋：

- 註1：卜小蝶。「資訊檢索技術的發展」。圖書資訊檢索技術，頁9。台北：文華，民國85年。
- 註2：黃嬪榕。「圖書館自動化系統權威控制功能之探討」，頁3。私立天主教輔仁大學，碩士論文，民國89年1月。
- 註3：劉春銀。「權威控制」。中國圖書館學會87年主題分析與機讀格式專題研習班講義，頁96。台北：中國圖書館學會，民國87年。
- 註4：同註2，頁5。
- 註5：同註2，頁10。
- 註6：The Society of American Archivists, Standards for Archival Description : A Handbook.” 2000.<http://www.archivists.org/catalog/stdspp/> (2000/05/12).
- 註7：蔡明月。線上資訊檢索-理論與應用，頁297。台北市：學生，民國80年。
- 註8：曾琪淑。「回現率」。圖書館學與資訊科學大辭典上冊，頁619-620。台北市：漢美，民國84年。
- 註9：胡歐蘭。「權威控制」。圖書館學與資訊科學大辭典下冊，頁2479。台北市：漢美，民國84年。
- 註10：Richard P. smiraglia Describing Archival Materials : The Use of the MARC AMC Format (New York: the HaworthPress,1990),pp.93-94.
- 註11：同註10，pp.95-97.
- 註12：ICA/CDS, ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families : pp.1-24 (Ottawa : ICA,1996).
- 註13：The Library of Congress. “ MARC 21 Concise Authority: Introduction” . 2001.<http://www.loc.gov/marc/authority/ecadintr.html>(2001/11/20)
- 註14：劉春銀，陳亞寧。「書目資料庫之權威控制系統規劃」。計算中心通訊11卷6期：1，民國84年3月。
- 註15：陳亞寧。「泛談權威控制政策的制定」。圖書與資訊學刊10期：28-29，民國83年8月。



註16：ICA. "Repot of the Sub-committee on Finding Aids Guidelines for the Preparation and Presentation of Finding Aids". 2002.

<http://www.hmc.gov.uk/icacds/eng/findingaids.htm> (2002/05/27)

註17：陳光華。「數位圖書館中權威控制系統的設計」。圖書與資訊學刊34期：55-61，民國89 年8 月。

註18：同註17。

註19：同註2，頁56。