

第四章 詮釋資料格式對映與轉換實作

本章分為七個小節，第一節瞭解檔案全宗與全集之關係。第二節針對「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)兩種不同描述層次(level)詮釋資料格式進行結構與欄位分析。第三節瞭解國民黨歷史檔案內涵，接著進行兩詮釋資料格式欄位對映(mapping)與說明兩詮釋資料格式內容轉換(conversion)。第四節透過可延伸樣式表語言轉換(XSLT)技術，進行詮釋資料格式結構轉換實作，由平面式單件層次詮釋資料編目記錄轉換成階層式全集層次詮釋資料編目記錄。第五節探討自動全集轉換實作相關問題。第六節探討「檔案描述編碼格式」(EAD)與「RSLP 全集描述綱要架構」之對映關係。第七節探討檔案全集描述呈現介面。

第一節 檔案全宗與全集

不同類型的典藏機構環境，由於典藏資源的特性與管理文化上的差異，除了對全集之概念有不同的認定與定義之外，也產生不同的全集型態與全集描述方法。本節深入瞭解檔案的特質、全集的特性以及全集的需求面。

一、檔案的特質

檔案的特質可歸納為以下六點：(薛理桂，民 87)

1. 唯一性：檔案管理的前身是文書管理，在文書管理的階段，大約有 95% 的文書將被銷毀，只餘留下約 5% 的文書具備持續性價值，進而成為檔案。因而，所遺留下的檔案大都具備唯一性的特質。
2. 多元性：檔案大多係政府單位處理公務的結晶，政府單位包羅萬象，因而檔案的內涵也無所不包。檔案所使用的媒體也具多元性從紙本式手稿、公文到錄音帶、照片、地圖等無所不包。
3. 公平性：檔案係政府單位處理公務的結果，大都屬於真實性的資料，除非少數有造假的情況。因而，在法律的證據與歷史的研究，檔案具有「第一手證據」(First hand evidence) 的價值。可依據檔案解釋事實的真實情況。

4. 可靠性：檔案的產生係由於政府單位公務人員處理公務的需要，並非特意要產生的文書。經由產生後的公文書在公務單位若干年的保存後，經鑑定有保存的價值、再移轉到檔案館典藏，成為正式的檔案。由於檔案具有官方的特性，其處理過程是歷經法定的程序，因而具有可靠性。
5. 自然性：檔案的產生是自然的彙集、漸進的與連續性的成果。
6. 關連性：公務人員在處理公務的過程中，並非單只處理本身的業務，常會有與其他單位有關連的情況，可能在一件檔案中與數個單位有密切關係。除了本機構內與其他單位有關連，還可能與機構外的其他單位有關連。

二、檔案的編排與描述

瞭解檔案的特質之後，進一步瞭解檔案在編排與描述上之特點。

檔案學上，對於檔案的典藏或管理上十分重視「整體編排」與「描述的完整性」，以下說明檔案的編排與檔案的描述：

（一） 檔案的編排（arrangement）

由於檔案是隨著組織（機關）或個人之業務與活動，有機的成長累積，因此檔案處理過程中尊重三個重要的原則，分別為「全宗原則」、「來源原則」以及「原始順序原則」：

1. 「全宗原則」：「全宗原則」為檔案學的核心，指的是同一組織（機關）或個人所產生之檔案，應視為一不可分割之整體予以處理，以確保原始檔案文件之間關聯性及整體性。
2. 「來源原則」：依檔案來源及產生的單位，進行檔案分類，以利檔案整理與編目。
3. 「原始順序原則」：維持原始檔案單位之檔案案卷排列順序，以反映出原始單位在檔案產生時之業務需求及檔案使用情況。

（二） 檔案的描述（description）

在檔案處理過程中，「全宗原則」、「來源原則」與「原始順序原則」等三個重要原則之具體應用，形成了檔案處理的全宗（fonds）、系列（Series）、案卷（file）以及單件（item）四個主要層級。

由於檔案處理的過程中強調層級的處理與層級之間的關係，以維護檔案的原始歷史價值，便於檔案的編排與描述以及協助檔案管理者處理檔案，因此，檔案在描述上屬於「多層級描述」。

在「多層級描述」中，全宗層級為檔案描述的最高（大）層級，提供檔案記錄之背景描述資訊，例如檔案記錄的範圍與內容、摘要、相關全宗人物的傳記與歷史等資訊。在全宗層級進行描述，即以檔案資源的集合體作為描述的單位進行描述，也可稱為檔案資源的全集描述。

在全宗層級進行描述之優點可從使用者、典藏機構管理者以及資料交換等三個面向進行說明：

1. 使用者面向：全宗的描述資訊，可幫助使用者在查檢全宗檔案之初，就可對此全宗檔案建立起一個概略性的了解。
2. 典藏機構管理者面向：檔案典藏具有保存歷史的重要意義，全宗描述有助於釐清典藏主題的歷史傳記、重要事件及典藏價值，由管理角度來看，全宗描述有助於了解典藏品的相關典藏狀況等資訊。
3. 資料交換面向：全宗層級是檔案描述時基本的描述層級，因此在全宗層級進行描述將有助於與國內外相關計劃進行資料的交換。

三、 檔案全宗與全集

經由上述第一節檔案的特質與第二節檔案的編排與描述，瞭解檔案界習慣上使用「全宗」（fonds）來稱呼其典型描述資料的集合體，而不用「全集」（collection）來指稱其典型描述資料的集合體；雖然「檔案的全宗」與「人為的全集」都可稱之為「全集」，但兩者實際上仍有許多不同的特質：

1. 「檔案的全宗」：可以瞭解單件檔案的起源（起源原則），且它們的排列反映出一個組織或個人記錄它們最初的工作順序（原始順序原則）。

2. 「人為的全集」：「人為的全集」當中的單件是關連的，但是缺乏全宗的凝聚一致性（coherence）。

以下進一步整理「檔案的全宗」與「人為的全集」兩種「全集」之異同，請參考表 4-1：

表 4-1 「檔案的全宗」與「人為的全集」之異同

異同 \ 類型	檔案的全宗	人為的全集
聚成全集準則	「全宗原則」、「來源原則」以及「原始順序原則」	主題、類型、格式、地點
聚集方式	非特意聚集，屬於自然的彙集、漸進的與連續性	人為特意聚集
描述單位	以一個機構或個人所有的檔案文件集合體進行描述	以單一資源之集合體進行描述
單件與單件之間	單件與單件之間有關連性	單件與單件之間除了擁有共同的某一特質之外，缺乏缺乏全宗的凝聚一致性
組成全集之單件內涵	多元性	
全集類型	兩者皆可稱為全集	

四、本研究單件轉全集之必要性

本研究以國民黨歷史檔案為全集描述實作對象，目前「國民黨黨史館」在檔案描述編目上是以單筆檔案作為描述的單位，以單筆檔案所建立起的檔案單件層次描述對使用者而言無法瞭解檔案整體脈絡之外，從管理角度來看，僅建立檔案單件層次描述將無助於了解典藏品的相關典藏狀況；因此有必要將目前的檔案單件層次描述轉換成檔案全宗（全集）層次描述，轉換後可便於瞭解檔案之整體脈絡與檔案典藏品相關典藏狀況。檔案單件層次描述與檔案全宗層次描述之優缺點比較請參考表 4-2。

表 4-2 檔案單件層次描述與檔案全宗層次描述之優缺點

描述類型 優缺點	檔案單件層次描述	檔案全宗層次描述
詮釋資料格式	「國民黨歷史檔案詮釋資料」	「檔案描述編碼格式」(EAD)
描述單位	單筆檔案	檔案紀錄整體—全宗
優點	不需透過檔案專家就可進行單筆檔案編目。 單筆檔案特質詳細描述。	幫助使用者瞭解檔案整體脈絡架構。 幫助管理者掌握典藏品相關典藏狀況資訊。 資料交換。
缺點	檔案單件層次描述無法瞭解整體檔案架構脈絡。	需經由對檔案整體內涵有全盤瞭解之檔案專家才能進行檔案全宗層次的描述。

第二節 詮釋資料格式結構與欄位分析

瞭解檔案之特質、全集的特性、全集的需求面以及本研究單件轉全集必要性之後，進入本研究自動全集轉換實作部分。

本研究在自動全集轉換實作部分以描述單筆檔案之「國民黨歷史檔案詮釋資料」以及描述檔案全宗之「檔案描述編碼格式」(EAD)兩種詮釋資料格式為例，探討由單件層次描述編目記錄轉換為全集層次描述編目記錄之自動全集轉換相關問題。

由「國民黨歷史檔案詮釋資料」(以單件為主要描述單位之單件層次詮釋資料)對映轉換(mapping/crosswalk)成「檔案描述編碼格式」(EAD)檔案檢索工具(archival finding aids)(以資源集合體為主要描述單位之全集層次詮釋資料)，首要工作便是分析與列出「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)描述項目欄位(fields或稱metadata element sets)之間的對映。

本節首先分析要進行對映轉換的詮釋資料格式之結構與欄位，即「國民黨

歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)之結構為何?欄位有多少,定義為何?

本研究進行兩詮釋資料格式架構對映與轉換,「國民黨歷史檔案詮釋資料」是以前資料欄位,而「檔案描述編碼格式」(EAD)則是以檔案單元主要結構之檔案描述作為格式對映之項目。

以下分別從「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)之背景、結構、組成的欄位(或稱元素)以及相關屬性進行分析。

一、「國民黨歷史檔案詮釋資料」結構與欄位分析

(一) 背景

中國國民黨歷史檔案,根據其所編纂的史料目錄,係分為「總類」、「屬於人的」、「屬於事的」、「專檔」四類;但是這份目錄不同於實際上所使用之檔案卡片目錄。

檔案卡片目錄分為「一般」(從興中會時期至1969年的檔案)、「漢」檔(即「漢口檔案」,包括國民黨武漢時期第一屆及第二屆中央執委會檔案)、「會」檔(各種會議記錄檔案)、「中」檔(中央委員會相關檔案)、「部」檔(聯俄容共時期士農工商內政五部的檔案)、「監」檔(國民政府時期中監會)、「政」檔(有關內政及各項國政業務的檔案)、「特」檔(各種特種檔案)、「環」檔(環龍路檔案)、「稚」檔(吳稚暉專檔)與「鷹」檔(迪鷹專檔)等十一類。

對於一個身處數量龐大、內容包羅萬象歷史檔案中的使用者來說,僅僅透過目前的檔案卡片目錄進行檔案的檢索將十分困難,且這些歷史檔案檔案也並非全部對外開放,造成實際使用上相當不便。

(二) 「國民黨歷史檔案詮釋資料」結構與欄位分析

「國民黨黨史館」提供的檔案卡片目錄是以單筆檔案作為描述單位,而「國民黨歷史檔案詮釋資料」是根據卡片目錄記載的內容進行欄位之設計,屬於平面式單筆檔案描述結構,其結構之示意圖請參考圖4-1:

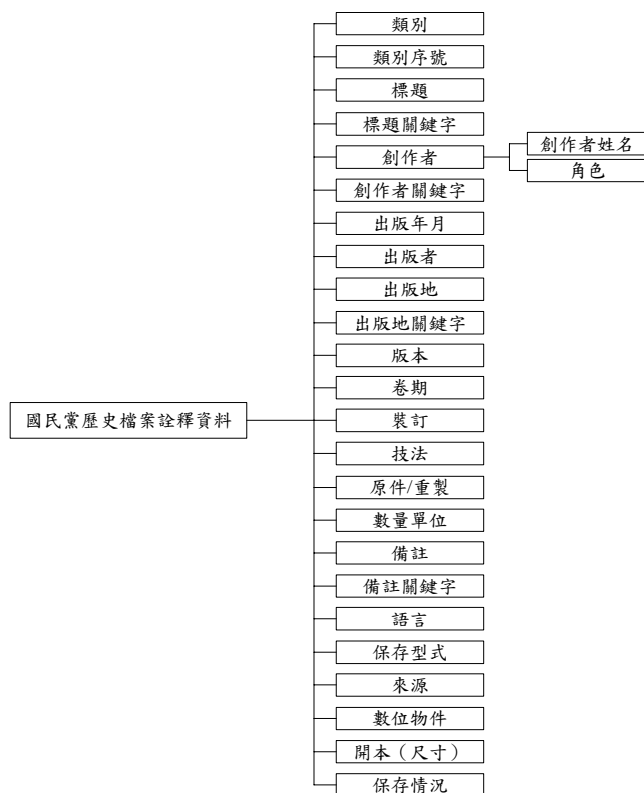


圖 4-1 「國民黨歷史檔案詮釋資料」平面式單筆檔案描述結構

資料來源：本研究

「國民黨歷史檔案詮釋資料」其描述項目欄位包含一類別、類別序號、標題、標題關鍵字、創作者、創作者姓名、角色、創作者關鍵字、出版年月、出版者、出版地、出版地關鍵字、版本、卷期、裝訂、技法、原件/重製、數量單位、備註、備註關鍵字、語言、保存型式、來源、開本（尺寸）、數位物件、保存情況等欄位，以下透過表 4-3 分析「國民黨歷史檔案詮釋資料」結構之欄位與相關說明：

表 4-3 「國民黨歷史檔案詮釋資料」欄位及說明

欄位名稱 / <標籤>	說明
類別 / <classification>	國民黨歷史檔案卡片目錄分類。必備且不可重複。著錄範例：稚、環、001、6.41。
類別序號 / <classificationNumber>	國民黨歷史檔案類別後的一組純數字編號。必備且不可重複。著錄範例：00001、001、1、2.1。
標題 / <title>	給予檔案紀錄之名稱或題名。必備且不可重

		複。著錄範例：吳稚暉致戴傳賢函。
標題關鍵字 / <titleKeyword>		標題欄位解析出的關鍵字，表達檔案紀錄主題之片語或關鍵字。非必備且可重複。
創作者 / <creator>	創作者姓名 / <creatorName>	對檔案紀錄內容主要負責之人或機關團體名。非必備且不可重複。著錄範例：吳稚暉。
	角色 / <character>	對檔案紀錄內容主要負責之人。非必備且可重複。 檔案紀錄內容負責之人或機關團體所扮演的角色。非必備且不可重複。著錄範例：撰、擬、主編。
創作者關鍵字 / <creatorKeyword>		創作者欄位解析出的人名或機關團體名之關鍵字。非必備且可重複。
出版年月 / <issuedDate>		檔案紀錄以目前形式呈現之日期。非必備且不可重複。著錄範例：1927/03/29。
出版者 / <publisher>		將檔案紀錄以目前之形式呈現出版之負責人或機關單位。非必備且不可重複。著錄範例：大陸雜誌社。
出版地 / <issuedPlace>		檔案記錄、出版品或刊物出版之地理位置。非必備且不可重複。著錄範例：南京、重慶。
出版地關鍵字 / <issuedPlaceKeyword>		出版地欄位解析出的地名關鍵字。非必備且可重複。
版本 / <edition>		檔案記錄、出版品或刊物之版本類別。非必備且不可重複。著錄範例：1版、再版、中英文對照版。
卷期 / <volume>		檔案記錄、出版品或刊物之卷期。非必備且不可重複。著錄範例：第一期、第一卷第四期。
裝訂 / <binding>		檔案記錄、出版品或刊物之裝訂方式。非必備且不可重複。著錄範例：平裝、裱裝、線裝。

技法 / <technique>	呈現檔案紀錄所使用之技術或方法。非必備且可重複。著錄範例：鋼筆、毛筆、鉛印、剪報。
原件/重製 / <duplication>	檔案記錄目前呈現之類型。非必備且不可重複。著錄範例：原件、抄件、稿件。
數量單位 / <quantity>	檔案記錄之數量。非必備且可重複。著錄範例：1 張、全 1 冊。
備註 / <note>	著錄檔案紀錄相關之註解。非必備且可重複。
備註關鍵字 / <noteKeyword>	備註欄位解析出表達檔案紀錄資源主題之關鍵字或片語。非必備且可重複。
語言 / <language>	檔案紀錄內容所使用之語言。非必備且不可重複。著錄範例：日文、法文、英文。
保存型式 / <pattern>	檔案記錄呈現保存之形體類型。非必備且不可重複。著錄範例：名片、明信片、報刊、小冊子、袖珍本。
來源 / <source>	敘述檔案紀錄的來源，讓使用者得知該檔案紀錄之來源。非必備且不可重複。著錄範例：中央秘書處計劃檢查士編印、中委會第一組編印。
開本（尺寸） / <format>	檔案紀錄之實體尺寸描述。非必備且不可重複。著錄範例：16 開、10 吋、長條。
數位物件 / <digitalObject>	檔案紀錄數位化影像檔。非必備且不可重複。
保存情況 / <condition>	檔案紀錄目前的保存狀態。非必備且不可重複。著錄範例：不全。

二、「檔案描述編碼格式」(EAD) 結構與欄位分析

(一) 背景

「檔案描述編碼格式」(EAD) 是網際網路環境下用於檢索工具編碼的標

準，也是專門用於描述檔案手稿紀錄所設計之詮釋資料格式，其階層式的檔案資源描述架構，可完整表達檔案編排與描述的「全宗原則」、「來源原則」、「原始順序原則」以及「多層次描述原則」，將相關的檔案紀錄聚合在一起，成為一個檔案全宗，反映出檔案資料實體的原始排列順序與結構，有助於使用者瞭解和檢索被描述的檔案紀錄資料。

「檔案描述編碼格式」(EAD)除了能提供全宗(全集)整體概括性之描述，幫助使用者迅速判斷組成該全宗之系列或單件是否符合所需之外，也能進一步描述組成該全宗之系列或單件相關資訊，提供使用者進一步查詢使用，即透過全宗層次的描述讓使用者在資訊查詢時可從大處著眼(全宗的描述)，小處著手(系列或單件的描述)。

「檔案描述編碼格式」(EAD)之結構由「EAD 標目」(EAD Header)、「序言」(Front Matter)與「檔案描述」(Archival Description)三大部分所組成，以下針對這三個部分所包含之元素(elements)、子元素(subelements)以及相關屬性(attributes)進行整理與說明。

(二) 「EAD 標目」

「檔案描述編碼格式」(EAD)結構的第一個組成部分為「EAD 標目」(EAD Header)。「EAD 標目」主要用於描述檢索工具本身相關的資訊，例如檢索工具的產生、出版、修訂或發行等相關訊息。

在「EAD 標目」元素之下有四個子元素，依序為「EAD 識別」(EAD Identifier)，屬於必備元素；「文件描述」(File Description)，屬於必備元素；「背景描述」(Profile Description)，屬於選擇性元素以及「修訂描述」(Revision Description)，屬於選擇性元素。「EAD 標目」元素結構如圖 4-3 所示：

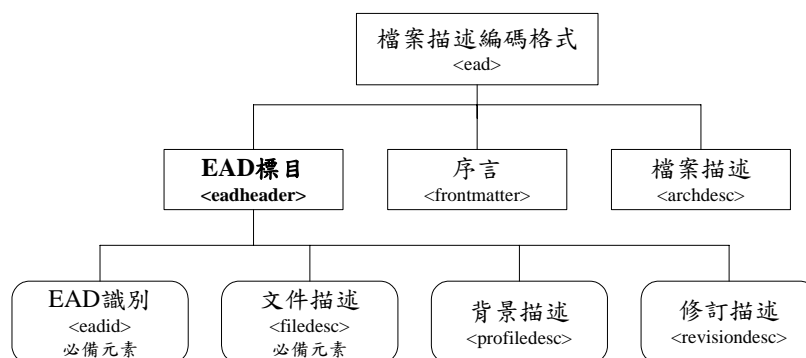


圖 4-3 「EAD 標目」元素結構

資料來源：本研究

詳細的「EAD 標目」元素說明階層式架構表格請參考附錄一。

(三) 「序言」

「檔案描述編碼格式」(EAD) 結構的第二個組成部分為「序言」(Front Matter)。「序言」元素將焦點集中在檢索工具的建立、出版或運用上而非有關檔案被描述的資訊。「序言」元素結構如圖 4-4 所示：

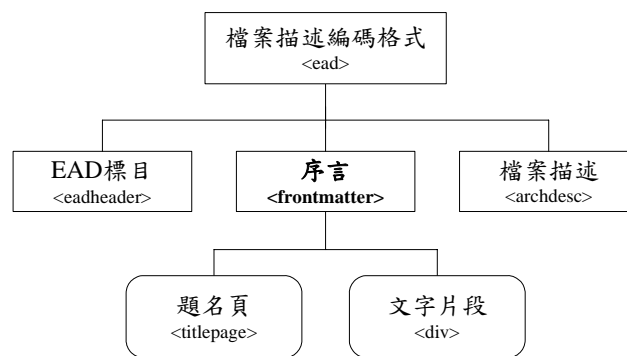


圖 4-4 「序言」元素結構

資料來源：本研究

詳細的「序言」元素及說明之階層式架構表格請參考附錄二。

(四) 「檔案描述」

「檔案描述編碼格式」(EAD) 結構的第三個組成部分為「檔案描述」(Archival Description)，該元素及其子元素主要用於描述檔案記錄本身以及檔案記錄管理上的相關資訊。

「檔案描述」元素能表達多層次之描述架構，彙編包括「視檔案資料整體為一個單位」以及「組成該檔案資料整體之部件」的相關描述資訊，體現檔案檢索工具表達階層結構之能力，讓不能親臨檔案實體典藏館之使用者只要透過網際網路查詢，就可瞭解一個全宗檔案組成之系列、卷、件的完整資訊。

「檔案描述」元素結構如圖 4-5 所示：

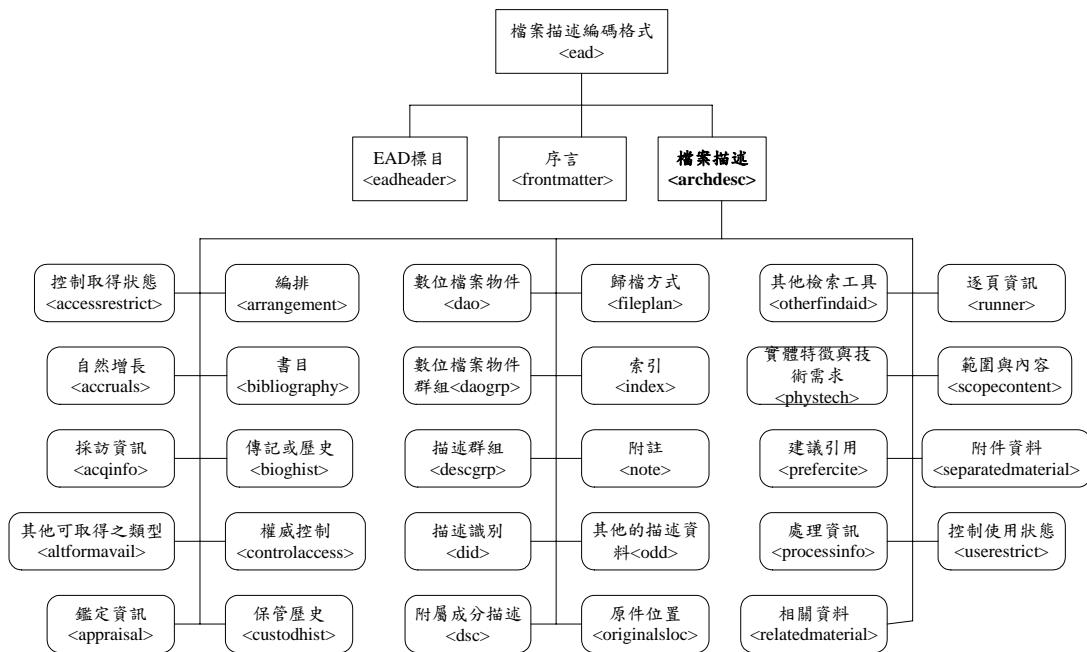


圖 4-5 「檔案描述」元素結構

資料來源：本研究

詳細的「檔案描述」元素說明表格請參考附錄三。

典型的檢索工具對一個全宗（全集）有兩個或三個觀點，每個觀點描述相同資料主體，但是在詳細程度上呈現變化：(Frusciano, 2000)

- 多層次描述之「第一層」
 - 用一個非常一般方式描述全集整體。
 - 提供資料類型之概觀。
 - 指出有意義的重要人物以及主題代表。
 - 提供檔案起源（provenance）與存取（access）等相關資訊。
 - 全面地描述全集，可能包含一個書目性梗概或代理者歷史以及範圍與內容註。
- 多層次描述之「第二層」
 - 焦點集中在全集當中的資料群集（groupings）。
 - 比在第一層更加詳細描述每一個群集。

- 強調更多具體的資料類型、附加的個體與主題代表。
 - 包含在整體中的系列或副系列之敘述性描述，可能被描繪在檢索工具中。
 - 依全集的複雜性與典藏機構之慣例而定，該中間層次描述（mid-level description）或許是不需要的。
- 多層次描述之「第三層」
- 可能描述每卷（file）或每個單件（item）。
 - 常採取檔案卷夾（container）或文件夾（folder）清單目錄形式。
 - 這些清單目錄明確地呈現出檔案資料智力活動階層（intellectual hierarchy），研究者運用這些清單目錄來取得資料。

「檔案描述編碼格式」（EAD）之「檔案描述」元素中透過「附屬成分描述」元素表達多層次描述之架構，如圖 4-6 所示，依實際檔案資料層級或典藏機構描述慣例而呈現出不同的檔案描述層次：

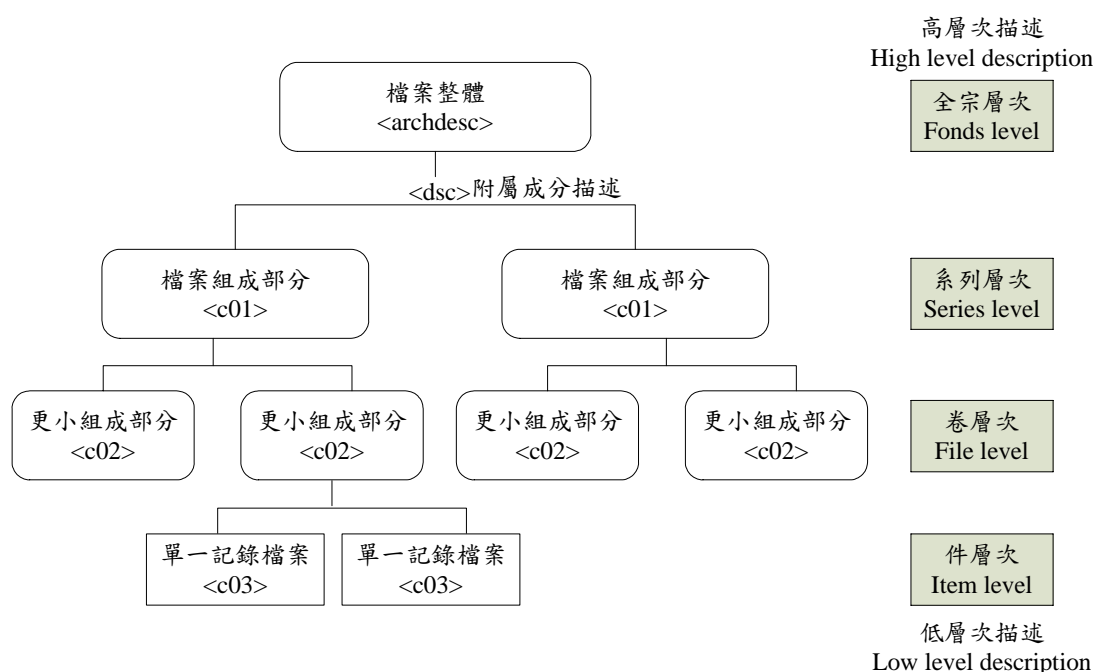


圖 4-6 「附屬成分描述」元素之階層式描述架構

第三節 詮釋資料格式欄位語意對映與欄位內容轉換

目前「國民黨黨史館」中典藏的歷史檔案在資訊組織方面主要是以單件檔案作為描述的單位以及資源發現的最前線；若要從目前底層詳細的單件層次檔案描述建立起上層的全集層次檔案描述，需先瞭解國民黨歷史檔案整體之內涵，再根據這些內涵設計符合本研究之「檔案描述編碼格式」(EAD)，進行全宗描述，分析與設計步驟如下：

1. 國民黨歷史檔案內涵之分析

瞭解全集描述實作對象檔案整體之內涵，例如國民黨歷史檔案在編排上所包含的層級，以確定全集描述之主體描述單位以及構成該主體描述單位之群組。

2. 分析全集層次詮釋資料「檔案描述編碼格式」(EAD)

根據國民黨歷史檔案內涵，分析「檔案描述編碼格式」(EAD)中需包含的檔案層級以及採用的元素、屬性與屬性值，即分析「檔案描述編碼格式」(EAD)中哪些元素用於全集資源的揭示與描述，哪些元素用來與「國民黨歷史檔案詮釋資料」欄位進行對映。

3. 進行兩詮釋資料格式欄位對映。

一、國民黨歷史檔案內涵分析

以檔案編排之「全宗原則」、「起源原則」與「原始順序原則」三大原則為基礎以及參考「國民黨黨史館」之檔案卡片目錄，將國民黨歷史檔案分為全宗、系列與件三個層級，以下針對這三個層級之間的關係進行說明：

1. 全宗：在本研究全宗是「中國國民黨」此一組織內產生的所有檔案紀錄文件之彙集，屬於「組織全宗」。
2. 系列：在全宗層級的下一個層次為系列，用來定義國民黨歷史檔案的第二層結構，本研究依照「國民黨黨史館」檔案卡片目錄進行劃分，系列層級包含「國防最高委員會會議記錄」、「國防最高委員會檔案」、「上海環龍路檔案」、「中央改造委員會檔案」、「中央改造委員會紀錄」、「中央改造委員會各種工作簡報」、「第七屆中央常務委員會議記

錄」以及「吳稚暉先生檔案」等八個系列。

3. 單件：在系列層級之下，用來定義國民黨歷史檔案的第三層結構，是本研究中檔案的最小描述單位。

二、本研究「檔案描述編碼格式」(EAD)分析

以國民黨歷史檔案內涵分析為基礎，分析「檔案描述編碼格式」(EAD)中透過哪些層級與元素用於全宗的描述，哪些元素用來與「國民黨歷史檔案詮釋資料」欄位進行對映，分為四個部分進行分析，請參考圖 4-7：



圖 4-7 本研究「檔案描述編碼格式」(EAD)分析

資料來源：本研究

以下分析這四個部分所採用的元素。

(一) 「EAD 標目」元素與屬性分析

「EAD 標目」(EAD Header) 元素主要是用來描述國民黨歷史檔案檢索工具的相關資訊，並非用來描述檔案主體。在此部分構成的元素包含：

- EAD 識別 (EAD Identifier) 元素
- 文件描述 (File Description) 元素
 - 標題敘述 (Title Statement) 元素
 - ◆ 檢索工具正題名 (Title Proper of the Finding Aid) 元素
- 背景描述 (Profile Description) 元素
 - 建立 (Creation) 元素
 - ◆ 日期 (Date) 元素
 - 使用語言 (Language Usage) 元素
 - ◆ 語言 (Language) 元素

「EAD 標目」元素、屬性與值之階層式架構請參考圖 4-8，詳細的階層式分析表格請參考附錄四。

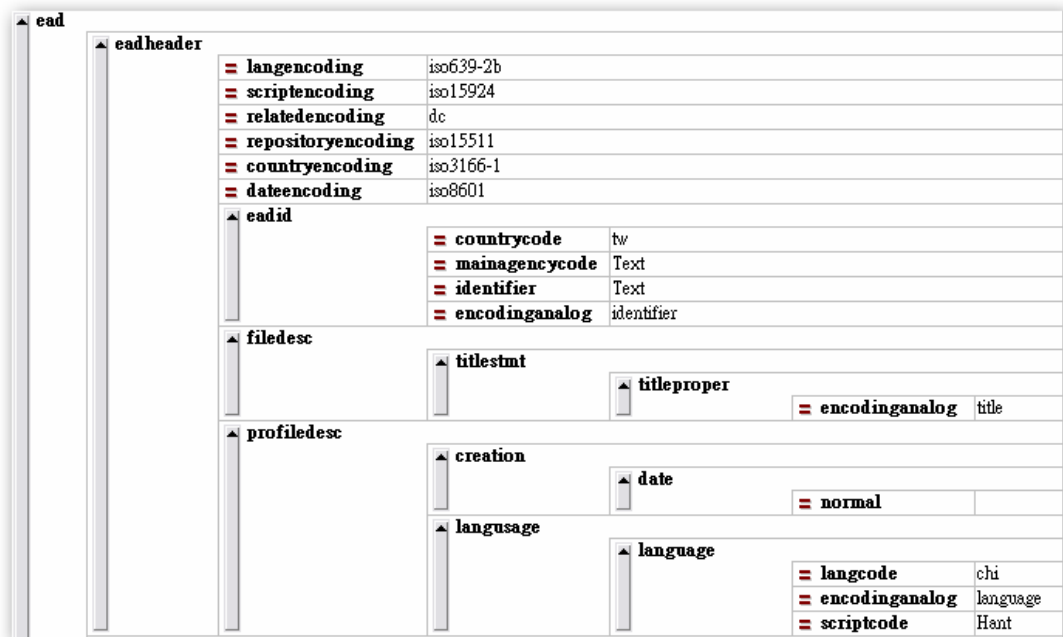


圖 4-8 「EAD 標目」元素、屬性與值

資料來源：本研究

(二) 「檔案描述」元素與屬性分析

「檔案描述」(Archival Description) 元素第一層以全宗作為描述的單位，進行檔案資料整體之內容、背景和範圍的描述，除此之外，也提供有助於檔案資料使用上與管理上的相關資訊，透過這些元素(欄位)所建立起的描述資料可稱為檔案的全宗層次描述或稱之為檔案資源的全集層次描述，提供檔案資源整體概觀的描述，讓使用者可根據這些描述資料，瞭解全宗及其所包括檔案之情況。

本研究中「檔案描述」第一層用來描述檔案全宗整體之內容、背景、範圍與數量等相關元素與子元素是參考英國「檔案中心」EAD 線上樣版(EAD 2002 Online Template)中所採用的元素與子元素。

用來描述全宗的元素與子元素包含：

- 描述識別 (Descriptive Identification) 元素
 - 單元識別 (ID of the Unit) 元素
 - 單元題名 (Title of the Unit) 元素
 - 單元日期 (Date of the Unit) 元素
 - 典藏機構 (Repository) 元素
 - 實體描述 (Physical Description) 元素
 - ◆ 範圍 (Extent) 元素
 - 起源 (Origination) 元素
 - 資料之語言 (Language of the Material) 元素
 - ◆ 語言 (Language) 元素
- 傳記或歷史 (Biography or History) 元素
- 範圍與內容 (Scope and Content) 元素
- 控制取得狀態 (Conditions Governing Access) 元素
- 其他檢索工具 (Other Finding Aid) 元素

- 相關資料 (Related Material) 元素
- 書目 (Bibliography) 元素
- 權威控制 (Controlled Access Headings) 元素

「檔案描述」元素、屬性與值階層式架構請參考圖 4-9，詳細的階層式分析表格請參考附錄五。

archdesc	
level	collection
type	inventory
relatedencoding	marc21
did	
unitid	countrycode=tw repositorycode=Text type=collection label=Accession number
unittitle	encodinganalog=245\$a
unitdate	normal=Text type=inclusive encodinganalog=245\$f
repository	
physdesc	
origination	
langmaterial	
bioghist	encodinganalog=5451_
scopecontent	encodinganalog=5202_
accessrestrict	encodinganalog=506
otherfindaid	
relatedmaterial	
bibliography	
controlaccess	
dsc	type=combined

圖 4-9 「檔案描述」元素、屬性與值

資料來源：本研究

(三) 「附屬成分描述」元素與屬性分析

「附屬成分描述」(Description of Subordinate Components) 元素主要用來描述全宗所包含的「系列」以及「系列」當中所包含的「件」。

用來描述「系列」的元素與子元素包含：

- 附屬成分描述 (Description of Subordinate Components) 元素
 - 組成部件 (Component) 元素 (系列)
 - ◆ 描述識別 (Descriptive Identification) 元素
 - 單元識別 (ID of the Unit) 元素

- 單元題名 (Title of the Unit) 元素
- 單元日期 (Date of the Unit) 元素
- 資料之語言 (Language of the Material) 元素
- ◆ 範圍與內容 (Scope and Content) 元素
- ◆ 編排 (Arrangement) 元素
- ◆ 控制取得狀態 (Conditions Governing Access)
- ◆ 控制使用狀態 (Conditions Governing Use)

「附屬成分描述」元素、屬性與值階層式架構請參考圖 4-10，詳細的階層式分析表格請參考附錄六。

▲ dsc	= type	combined
	▲ c01	
	= level	series
	▲ did	
		▼ unitid encodinganalog=099
		▼ unittitle encodinganalog=245\$a
		▼ unitdate type=inclusive encodinganalog=245\$f
		▼ langmaterial
		▼ scopecontent encodinganalog=5202_
		▼ arrangement encodinganalog=351
		▼ accessrestrict encodinganalog=506
		▼ userrestrict

圖 4-10 「附屬成分描述」元素、屬性與值

資料來源：本研究

(四) 單件層次描述元素與屬性分析

單件層次描述元素與屬性主要用來描述單件檔案紀錄，為了進行「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD) 兩種不同描述粒度層次之詮釋資料格式對映轉換，本研究將「檔案描述編碼格式」(EAD) 之單件層次描述所使用之元素、屬性與值的分析獨立成單個表格。

描述單件檔案的元素與子元素包含：

- 組成部件 (Component) 元素 (件)
 - 描述識別 (Descriptive Identification) 元素
 - ◆ 單元識別 (ID of the Unit) 元素
 - ◆ 單元題名 (Title of the Unit) 元素
 - ◆ 起源 (Origination) 元素
 - ◆ 單元日期 (Date of the Unit) 元素
 - ◆ 資料之語言 (Language of the Material) 元素
 - ◆ 實體描述 (Physical Description) 元素
 - ◆ 數位檔案物件 (Digital Archival Object) 元素
 - ◆ 附註 (Note) 元素
 - 權威控制 (Controlled Access Headings) 元素

單件層次描述元素與屬性分析階層式架構請參考圖 4-11，詳細的階層式分析表格請參考附錄七。

level	item
▲ did	<ul style="list-style-type: none"> ▼ unitid (2) ▼ unittitle encodinganalog=245\$a ▼ origination label=Creator: ▼ unitdate type=inclusive encodinganalog=245\$f ▼ langmaterial ▼ physdesc ▼ dao linktype=simple href=Text ▼ note (7)
	▼ controlaccess

圖 4-11 單件層次描述元素、屬性與值

資料來源：本研究

三、 「國民黨歷史檔案詮釋資料」欄位對映 EAD 欄位

(一) 詮釋資料格式欄位對映與轉換相關問題

當兩種詮釋資料格式進行對映轉換時，需考慮的問題包含：(Moen, 2005)

1. 兩詮釋資料格式之間的語意 (semantic)、結構 (structural) 與資料轉換 (data conversion)。
2. 單向對映 (one-way) 或可反 (逆) 轉的對映 (reversible)。
3. 兩個元素之間的對映情形
 - (1) 一對一 (one-to-one)：來源資料元素對應至目標資料元素。
 - (2) 一對多 (one-to-many)：來源綱要架構 (schema) 中的元素對映至目標架構 (schema) 時，拆解成兩個以上的資料元素。
 - (3) 多對一 (many-to-one)：來源綱要架構 (schema) 中的元素在對映至目標綱要架構 (schema) 時，需將多個資料元素合併成一個資料元素。
 - (4) 一對零 (one-to-zero)：語意上沒有相容的元素。
4. 資料轉換 (Data conversion)
 - (1) 從包含較少格式到包含較多格式 (From less inclusive to more inclusive format)。
 - (2) 從未受控制到受控制的詞彙。

根據「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD) 兩詮釋資料格式結構與欄位分析為基礎，以下分別從結構轉換、語意轉譯以及詮釋資料格式欄位內容轉換等三方面分析「國民黨歷史檔案詮釋資料」對映轉換至「檔案描述編碼格式」(EAD) 之情形：

1. 結構轉換方面 (structural transform)
 - (1) 平面式 (flat) 描述結構轉換成階層式 (hierarchical) 描述結構：與「國民黨歷史檔案詮釋資料」平面式單一描述結構相較之下「檔

案描述編碼格式」(EAD) 屬於階層式多層次描述結構，因此在進行詮釋資料格式對映轉換時是由平面式詮釋資料格式對映轉換成階層式詮釋資料格式。

- (2) 描述粒度 (granularity) 層次：兩詮釋資料格式可表達描述粒度層次 (粗細) 之能力不同。「檔案描述編碼格式」(EAD) 可表達檔案資源的全宗 (全集) 層次描述、系列層次描述以及單件層次描述；而「國民黨歷史檔案詮釋資料」僅能表達單一描述粒度，即描述資源詳細之特徵。因此在進行兩詮釋資料格式對映轉換時必須對映轉換至相同的描述粒度層次中。

綜合上述結構轉換方面兩小點，圖 4-12 為「國民黨歷史檔案詮釋資料」對映到「檔案描述編碼格式」(EAD) 結構之示意圖，由平面式描述結構轉換成階層式描述結構且對映至相同的描述粒度層次：

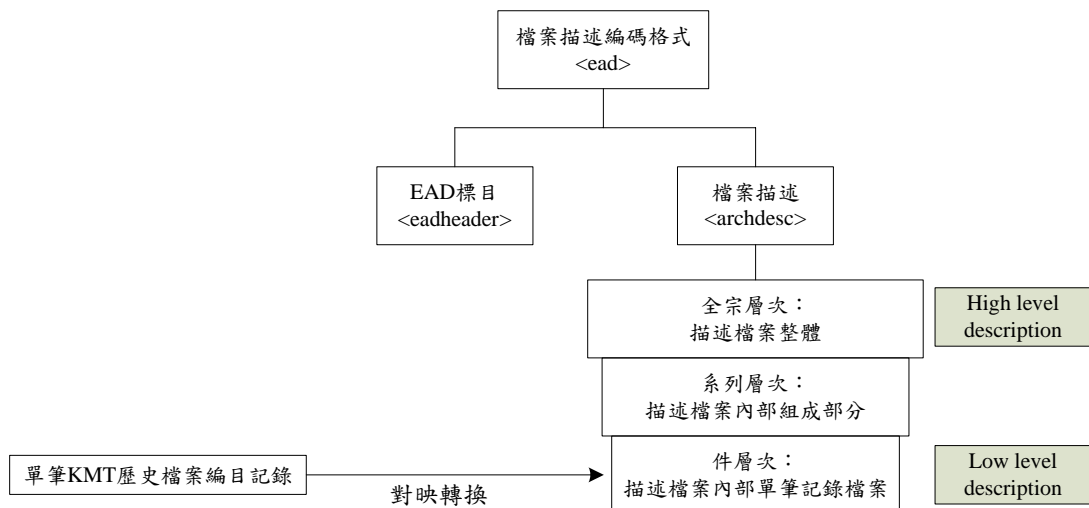


圖 4-12 「國民黨歷史檔案詮釋資料」與 EAD 結構轉換

資料來源：本研究

2. 語意轉譯方面 (semantic translation)

- (1) 描述項目欄位名稱：兩詮釋資料描述項目之欄位名稱不同，「國民黨歷史檔案詮釋資料」主要是依據國民黨歷史檔案之特性特質所設計，屬於地方性 (local) 之詮釋資料格式；而「檔案描述編碼格式」(EAD) 則是國際檔案界用於檔案檢索工具編碼的詮釋資料格式標準，因此在「國民黨歷史檔案詮釋資料」中較特定用來描述國民黨歷史檔案資料特性或特質之欄位在對映到「檔案描述編碼格式」(EAD) 中檔案描述的欄位時，由於欄位語意不相

容問題，產生某些欄位無法對映情況。兩詮釋資料格式欄位詳細的對映情形將於下面進行分析與說明。

3. 兩詮釋資料格式欄位內容轉換 (conversion)

- (1) 資料轉換內容控制部分：「國民黨歷史檔案詮釋資料」中的「標題關鍵字」、「創作者關鍵字」、「出版地關鍵字」以及「備註關鍵字」等欄位的關鍵字並未進行權威控制；而「檔案描述編碼格式」(EAD)中的權威控制欄位(元素)缺乏可應用在本研究實作對象進行檔案內容權威控制之權威控制參照檔。國外常採用檔案描述規則(Archives, Personal Papers, and Manuscripts, 簡稱 APPM)、美國國會圖書館標題表(Library of Congress Subject Headings, 簡稱 LCSH)或美國國會圖書館人名權威檔(Library of Congress Name Authority File, 簡稱 LCNAF)等作為檢索工具內容標準。

(二) 「國民黨歷史檔案詮釋資料」欄位對映 EAD 欄位

再瞭解詮釋資料格式欄位對映與轉換相關問題之後，進一步分析「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)兩詮釋資料格式欄位對映情形。

「國民黨歷史檔案詮釋資料」欄位對映到「檔案描述編碼格式」(EAD)結構中檔案描述之單件層次描述欄位可分為四種情形，分別為『語意相容可直接對映的欄位』、『加入屬性與值進行對映的欄位』、『無直接可對映的欄位』以及『不進行對映的欄位』，以下針對這四種對映情形進行說明：

1. 語意相容可直接對映的欄位：即「國民黨歷史檔案詮釋資料」中的欄位其語意可相容於「檔案描述編碼格式」(EAD)中的欄位。此情況所包含的欄位有：

➤ 標題－單元題名

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	標題	單元題名
標籤對映	<title>	<archdesc><dsc><c02 level="item"><did><unittitle

		encodinganalog="245\$a">
--	--	--------------------------

- 創作者姓名—名稱（原始創作者欄位中的資料並未區隔個人名稱、家族名稱或法人團體名稱，因此對映至「檔案描述編碼格式」（EAD）中的名稱欄位，確保欄位對映時語意上的最大相容性）

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	創作者姓名	名稱
標籤對映	<creatorName>	<archdesc><dsc><c02 level="item"><did><originati on label="Creator:"><name>

- 出版年月—單元日期

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	出版年月	單元日期
標籤對映	<issuedDate>	<archdesc><dsc><c02 level="item"><did><unitdate type="inclusive" encodinganalog="245\$f">

- 數量單位—範圍

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	數量單位	範圍
標籤對映	<quantity>	<archdesc><dsc><c02 level="item"><did><physdes c> <extent encodinganalog="300\$a">

➤ 語言－使用語言

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	語言	使用語言
標籤對映	<language>	<archdesc><dsc><c02 level="item"><did><langmat erial ><language encodinganalog="546">

➤ 開本（尺寸）－尺寸

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	開本（尺寸）	尺寸
標籤對映	<format>	<archdesc><dsc><c02 level="item"><did><physdes c><dimensions encodinganalog="300\$c">

➤ 數位物件－數位檔案物件（更為精確的來說是對映至「數位檔案物件」欄位中的「數位檔案物件描述」欄位的「段落」欄位）

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	數位物件	數位檔案物件
標籤對映	<digitalObject>	<archdesc><dsc><c02 level="item"><did><dao linktype="simple" href="Text"><daodesc><p>

- 加入屬性與屬性值進行對映的欄位：於「檔案描述編碼格式」（EAD）檔案描述欄位（或稱元素）中加入特定的屬性（attribute）與值（value），讓欄位更為專指（specific）再進行欄位的對映。此情況所包含的欄位

有：

➤ 類別－單元識別（屬性名稱="屬性值"）

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	類別	單元識別（標籤屬性="類別"）
標籤對映	<classification>	<archdesc><dsc><c02 level="item"><did><unitid label="classification" encodinganalog="099">

➤ 類別序號－單元識別

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	類別序號	單元識別（標籤屬性="類別序號"）
標籤對映	<classificationNumber>	<archdesc><dsc><c02 level="item"><did><unitid label="classificationNumber" encodinganalog="099">

➤ 角色－名稱*

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	角色	名稱（角色屬性="角色"）
標籤對映	<character>	<archdesc><dsc><c02 level="item"><did><originati on label="Creator:"><name role="character">

➤ 裝訂－實體面

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	裝訂	實體面（類型屬性="裝訂"）
標籤對映	<binding>	<archdesc><dsc><c02 level="item"><did><physdes c> <physfacet encodinganalog="300\$b" type="binding">

➤ 技法－實體面

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	技法	實體面（類型屬性="技法"）
標籤對映	<technique>	<archdesc><dsc><c02 level="item"><did><physdes c> <physfacet encodinganalog="300\$b" type="technique">

➤ 原件/重製－實體面

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	原件/重製	實體面（類型屬性="原件/ 重製"）
標籤對映	<duplication>	<archdesc><dsc><c02 level="item"><did><physdes c><physfacet encodinganalog="300\$b" type="duplication">

➤ 保存型式－實體面

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	保存型式	實體面 (類型屬性="保存型式")
標籤對映	<pattern>	<archdesc><dsc><c02 level="item"><did><physdesc><physfacet encodinganalog="300\$b" type="pattern">

➤ 出版地關鍵字－權威控制：地理名稱

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	出版地關鍵字	地理名稱 (角色屬性="主題")
標籤對映	<issuedPlaceKeyword>	<archdesc><dsc><c02 level="item"><controlaccess><controlaccess><geogname role="subject" encodinganalog="651" source="Text" rules="local">

3. 無直接可對映的欄位：語意上沒有相容的欄位，但原始欄位中的檔案特質描述資料可提供使用者作為判斷檔案內容的有用資訊，則將此情況中的欄位對映至「檔案描述編碼格式」(EAD)的「附註」(Note)元素並加入標籤(LABEL)屬性與屬性值之後的「段落」(Paragraph)元素裡。此情況所包含的欄位有：

➤ 出版者－附註

	國民黨歷史檔案詮釋資料	檔案描述編碼格式

欄位對映	出版者	附註（出版者）
標籤對映	<publisher>	<archdesc><dsc><c02 level="item"><did><note label="publisher"><p>

➤ 出版地—附註

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	出版地	附註（出版地）
標籤對映	<issuedPlace>	<archdesc><dsc><c02 level="item"><did><note label="issuedPlace"><p>

➤ 版本—附註

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	版本	附註（版本）
標籤對映	<edition>	<archdesc><dsc><c02 level="item"><did><note label="edition"><p>

➤ 卷期—附註

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	卷期	附註（卷期）
標籤對映	<volume>	<archdesc><dsc><c02 level="item"><did><note label="volume"><p>

➤ 備註—附註

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	備註	附註（備註）
標籤對映	<note>	<archdesc><dsc>< c02 level="item"><did><note label="note"><p>

➤ 來源—附註

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	來源	附註（來源）
標籤對映	<source>	<archdesc><dsc>< c02 level="item"><did><note label="source"><p>

➤ 保存情況—附註

	國民黨歷史檔案詮釋資料	檔案描述編碼格式
欄位對映	保存情況	附註（保存情況）
標籤對映	<condition>	<archdesc><dsc>< c02 level="item"><did><note label="condition"><p>

4. 不進行對映的欄位：原始「國民黨歷史檔案詮釋資料」中的「標題關鍵字」、「創作者關鍵字」以及「備註關鍵字」等欄位的資料並未明確區隔出人名、機關組織名、主題與地名等關鍵字，因此在對映轉換到「檔案描述編碼格式」（EAD）中的「權威控制」（Controlled Access Headings）欄位（元素）時，產生內容轉換上的問題；若原始關鍵字欄位中的資料內容有進一步區隔出人名、機關組織名、主題或地名等關鍵字時，則關鍵字對映情況為一對多。

詳細的「國民黨歷史檔案詮釋資料」欄位對映到「檔案描述編碼格式」(EAD) 檢索工具中檔案描述之單件層次描述欄位之表格，請參考附錄八。

除了上述欄位語意對映的四種情況之外，「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD) 兩詮釋資料格式在元素之間對映情形包含有「一對一關係」、「多對一關係」以及「一對零關係」等三種對映關係。

第四節 自動全集轉換實作

本章第三節分析兩詮釋資料格式欄位語意對映或稱語意匹配 (linguistic matching) 以及兩個元素之間的對映情形，請參照圖 4-13 左半部所示，欄位語意相容對映之分析結果做為第四節撰寫 XSLT 轉換規則文件以進行自動全集轉換實作時之依據。

在本研究中，轉換所指的是由「國民黨歷史檔案詮釋資料」編目 XML 檔透過 XSLT 轉換規則文件轉換成「檔案描述編碼格式」(EAD) 檢索工具 XML 檔，即詮釋資料格式結構轉換 (transformation)，請參照圖 4-13 右半部所示。

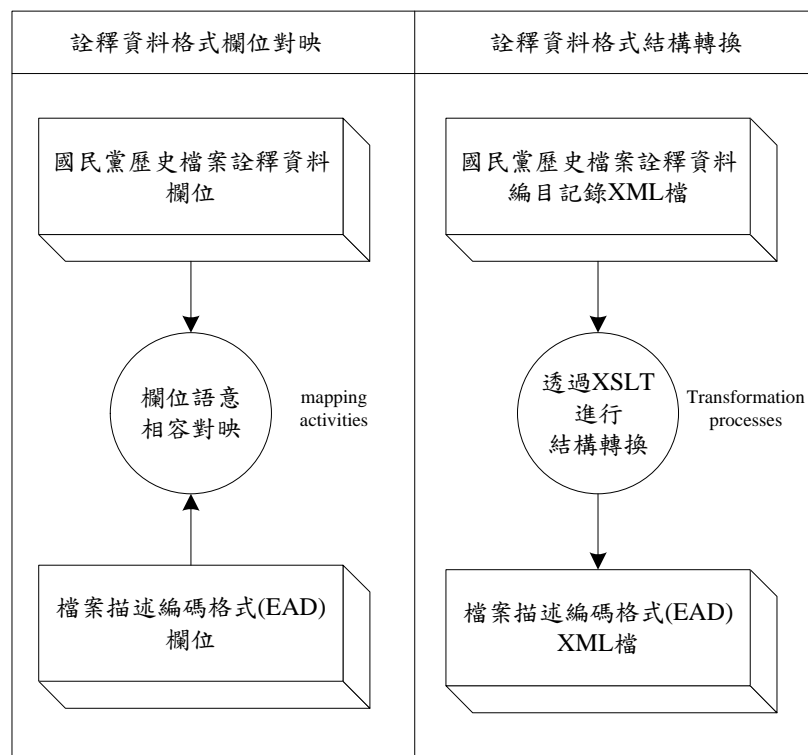


圖 4-13 「國民黨歷史檔案詮釋資料」對映與轉換到 EAD

資料來源：本研究

一、XSLT 轉換詮釋資料格式結構

可延伸樣式表語言轉換 (XSLT) 是一種樹狀結構導向的轉換語言，透過 XSLT 可將 XML 文件樹狀架構轉換成另一個新的架構，本研究即透過可延伸樣式表語言轉換 (XSLT) 將單件層次描述的國民黨歷史檔案單筆編目紀錄 XML 檔轉換成全集層次描述的「檔案描述編碼格式」(EAD) 編目紀錄 XML 檔。

(一) 本研究運用 XSLT 轉換詮釋資料格式結構流程

本研究運用 XSLT 轉換詮釋資料格式結構流程如下：

1. 設定來源樹 (source tree)：XSLT 轉換的輸入文件為來源樹，在本研究中指的是以「國民黨歷史檔案詮釋資料」編目之單筆國民黨歷史檔案編目紀錄 XML 檔。
2. 撰寫 XSLT 對映轉換規則文件：XSLT 對映轉換規則文件記錄「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD) 兩詮釋資料格式之間的資料轉換關係 (用於比對元素、建立元素)、規範從一個詞彙轉換到另一個詞彙的過程。
3. 設定結果樹 (result tree)：透過 XSLT 轉換後的輸出為結果樹，在本研究中指的是「檔案描述編碼格式」(EAD) 編目紀錄 XML 檔。
4. 測試與驗證轉換結構後之結果：由「國民黨歷史檔案詮釋資料」編目紀錄 XML 檔轉換成「檔案描述編碼格式」(EAD) 編目紀錄 XML 檔後，透過「檔案描述編碼格式」(EAD) 之文件類型定義檔 (DTD) 驗證轉換後的檔案檢索工具是否為有效的 (valid)。

本研究運用 XSLT 轉換詮釋資料格式結構之流程圖請參考圖 4-14：

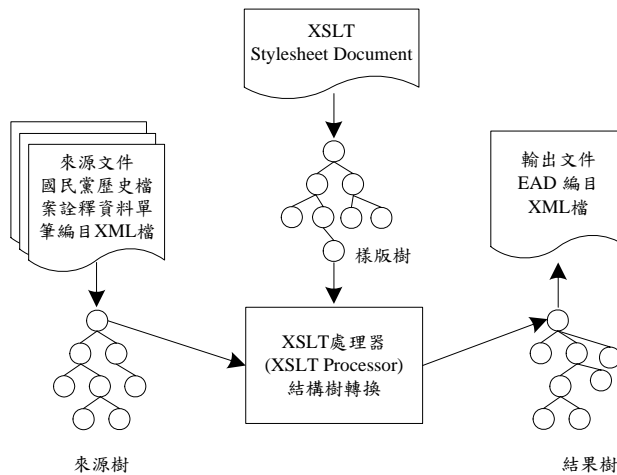


圖 4-14 透過 XSLT 轉換詮釋資料格式結構流程

資料來源：本研究

(二) XSLT 轉換規則文件結構

轉換「國民黨歷史檔案詮釋資料」單筆編目紀錄 XML 檔到「檔案描述編碼格式」(EAD) 編目紀錄 XML 檔之 XSLT 轉換規則文件本身也是一份 XML 文件，其結構包含 XML 路徑語言、新增元素與轉換語彙，以下針對該份文件之結構進行簡要說明，完整詳細的 XSLT 轉換規則編碼文件請參考附錄九：

1. XML 路徑語言 (XML Path Language, 簡稱 XPath)：在 XSLT 轉換規則文件中，主要是透過 XML 路徑語言 (XPath)，在 XML 文件找尋資料，其語法可用來指出文件架構或資料之位置，指定需要進行轉換之內容。
2. 新增元素、屬性與屬性值：在 XSLT 轉換規則文件中為轉換後的結果樹加入新資料元件，即當來源樹的某一節點符合 XSLT 轉換規則文件中的某個規則模式時，該規則所涉及的內容將在結果樹中被建立起來。例如在本研究中，從「國民黨歷史檔案詮釋資料」編目紀錄 XML 檔轉換成「檔案描述編碼格式」(EAD) 時編目紀錄 XML 檔，透過 XSLT 轉換規則文件新增「檔案描述編碼格式」(EAD) 中的「EAD 標目」、「檔案描述」等元素及其相關子元素、屬性與屬性值。
3. 轉換語彙：在 XSLT 轉換規則文件中包含了一組規則，用來規範從一個詞彙轉換到另一個詞彙的過程，定義出兩份不同 XML 文件之間資料元素的對映關係。在本研究中指的是「國民黨歷史檔案詮釋資料」

欄位對映轉換成「檔案描述編碼格式」(EAD) 檔案描述欄位的對映轉換規則。

二、自動全集轉換

再撰寫 XSLT 轉換規則文件後，本研究透過 XMLSPY 應用軟體進行單筆國民黨歷史檔案編目記錄 XML 檔轉換至一筆「檔案描述編碼格式」(EAD) 編目記錄 XML 檔，即在「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD) 之間交換編目記錄資料。

(一) 單筆國民黨歷史檔案編目記錄轉換實作

以中國國民黨(組織全宗)的「中央改造委員會檔案」(系列)中的一筆檔案紀錄(單件)為例，進行單件層次描述編目紀錄轉換為全集層次描述編目記錄。

圖 4-15 是以「國民黨歷史檔案詮釋資料」編目的單筆檔案描述記錄及其 XML 編碼檔：

類別	6.43
類別序號	52
標題	中改會第52次工作會議紀錄
標題關鍵字	中改會；會議記錄
出版年月	1951/10/02
出版地	台北
出版地關鍵字	台北
技法	毛筆
技法	油印
原件/重製	原件
數量單位	24張
備註	附台灣各界1元獻機運動競賽辦法
備註關鍵字	台灣各界1元獻機運動競賽辦法


```

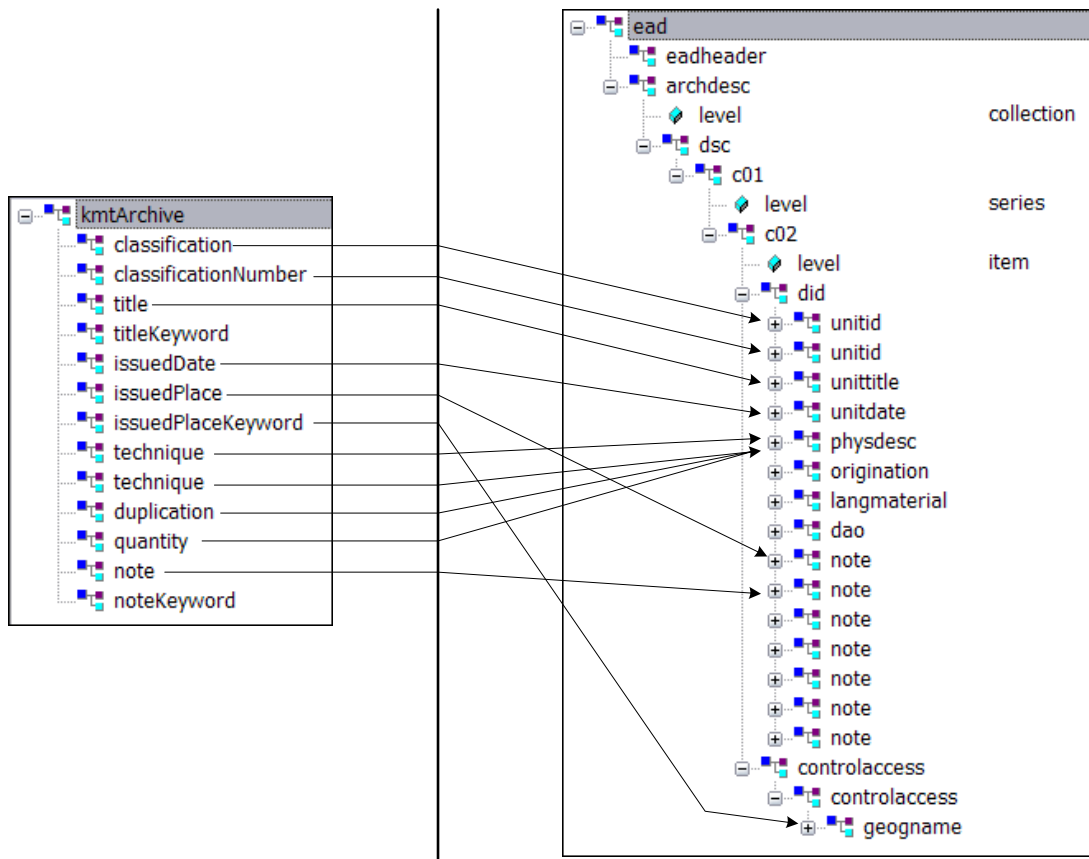
<?xml version="1.0" encoding="UTF-8"?>
<kmtArchive>
  <classification>6.43</classification>
  <classificationNumber>52</classificationNumber>
  <title>中改會第52次工作會議紀錄</title>
  <titleKeyword>中改會；會議記錄</titleKeyword>
  <issuedDate>1951/10/02</issuedDate>
  <issuedPlace>台北</issuedPlace>
  <issuedPlaceKeyword>台北</issuedPlaceKeyword>
  <technique>毛筆</technique>
  <technique>油印</technique>
  <duplication>原件</duplication>
  <quantity>24張</quantity>
  <note>附台灣各界1元獻機運動競賽辦法</note>
  <noteKeyword>台灣各界1元獻機運動競賽辦法</noteKeyword>
</kmtArchive>

```

圖 4-15 單筆「國民黨歷史檔案詮釋資料」編目記錄及其 XML 編碼檔

資料來源：本研究

透過 XMLSPY 應用軟體載入 XSLT 轉換規則文件與單筆國民黨歷史檔案編目紀錄 XML 檔，進行自動全集單筆轉換實作，請參考圖 4-16 自動轉換示意圖：



XSLT轉換規則文件

圖 4-16 本研究自動全集轉換示意圖

資料來源：本研究

表 4-4 為單筆國民黨歷史檔案編目紀錄轉換至「檔案描述編碼格式」(EAD) 中單件層次描述編目紀錄結果：

表 4-4 單筆國民黨歷史檔案編目紀錄轉換至 EAD 單件層次描述編目紀錄

元素名稱		編目記錄
描述識別<did>		
單元識別 (標籤屬性="類別") <unitid>		6.43
單元識別 (標籤屬性="類別序號") <unitid>		52
單元題名<unittitle>		中改會第 52 次工作會議紀錄
單元日期<unitdate>		1951/10/02

	實體描述<physdesc>	
	數量單位<extent>	24 張
	實體面 (類型屬性="技法") <physfacet>	毛筆
	實體面 (類型屬性="技法") <physfacet>	油印
	實體面 (類型屬性="原件/重製") <physfacet>	原件
	附註 (標籤屬性="出版地") <note>	
	段落	台北
	附註 (標籤屬性="備註") <note>	
	段落	附台灣各界 1 元獻機運動競賽辦法
	權威控制<controlaccess>	
	權威控制<controlaccess>	
	地理名稱(角色屬性="主題")<geogname>	台北

單筆「國民黨歷史檔案詮釋資料」編目紀錄 XML 檔透過 XSLT 轉換規則文件轉換成「檔案描述編碼格式」(EAD) 編目紀錄 XML 檔之完整檔案階層描述請參照附錄十。

(二) 自動全集轉換實作結果分析

由於「國民黨歷史檔案詮釋資料」欄位是對映至「檔案描述編碼格式」(EAD) 中檔案單件層次描述欄位，因此透過 XSLT 轉換規則文件轉換詮釋資料格式結構與資料後，「檔案描述編碼格式」(EAD) 中用來描述全宗、系列以及用來描述檢索工具等欄位是需要藉由人工方式填入資料，請參考圖 4-17，再填妥這些欄位的資料後，才能成為一筆完整的檔案資源全集描述。

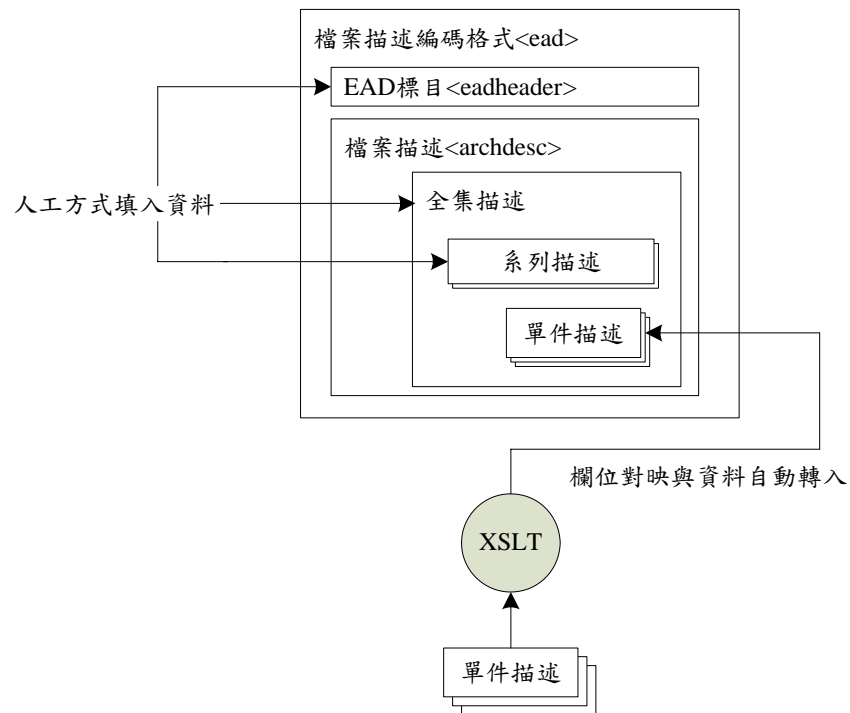


圖 4-17 「國民黨歷史檔案詮釋資料」欄位對映 EAD 單件層次描述欄位

資料來源：本研究

以下列出詮釋資料格式結構與資料轉換成「檔案描述編碼格式」(EAD)後，需再藉由人工方式填入資料的欄位，分為兩個部分，第一個部分為「EAD 標目」中的欄位及子欄位，第二個部分為「檔案描述」之全宗描述欄位以及系列描述欄位，以下逐一說明，並填入相關描述資料：

1. 「EAD 標目」中的欄位及子欄位—「EAD 標目」是組成「檔案描述編碼格式」(EAD)的必備結構。表格 4-5 列出「EAD 標目」中需藉由人工方式填入資料的欄位與相關編目記錄：

表 4-5 「EAD 標目」編目記錄

元素名稱	編目紀錄 (測試性編目記錄)
EAD 標目<eadheader>	封包元素，不需著錄內容
EAD 識別<eadid>	中國國民黨歷史檔案 EAD 檢索工具文件 唯一識別編碼
文件描述<filedesc>	不需著錄內容

標題敘述<titlestmt>	封包元素，不需著錄內容
檢索工具正題名 <titleproper>	中國國民黨歷史檔案記錄指南
背景描述<profiledesc>	不需著錄內容
建立<creation>	中國國民黨歷史檔案 EAD 檢索工具由蕭伯瑜編碼
日期<date>	2006
使用語言<language>	不需著錄內容
語言<language>	繁體中文

2. 「檔案描述」中全宗描述欄位與系列描述欄位。表格 4-6 列出檔案全宗描述中需藉由人工方式填入資料的欄位。表格 4-7 列出檔案系列描述中需藉由人工方式填入資料的欄位。

表 4-6 檔案全宗描述編目記錄

元素名稱	編目記錄 (測試性編目記錄)
檔案描述<archdesc> (全宗層次)	封包元素，不需著錄內容
描述識別<did>	封包元素，不需著錄內容
單元識別<unitid>	中國國民黨歷史檔案全宗登錄號
單元題名<unittitle>	中國國民黨歷史檔案紀錄全集
單元日期<unitdate>	1914-2006
典藏機構<repository>	中國國民黨
實體描述<physdesc>	封包元素，不需著錄內容
數量單位<extent>	中國國民黨歷史檔案全宗 (全

		集) 包含三百萬件史料。
起源<origination>		中國國民黨
資料之語言<langmaterial>		不需著錄內容
	語言<language>	繁體中文
傳記或歷史<bioghist>		不需著錄內容
	標題<head>	中國國民黨歷史沿革
	段落<p>	中國國民黨始源於孫中山先生於 1894 年創立的興中會；興中會於 1905 年改組為同盟會；同盟會於 1912 年改組為國民黨；國民黨於 1914 年改組為中華革命黨；中華革命黨於 1919 年改組為中國國民黨，黨名至今未變。自 1894 年創黨至今 2006 年，中國國民黨已有 112 年的歷史。
範圍與內容<scopecontent>		不需著錄內容
	段落<p>	分為「一般」(從興中會時期至 1969 年的檔案)、「漢」檔(即「漢口檔案」，包括國民黨武漢時期第一屆及第二屆中央執委會檔案)、「會」檔(各種會議記錄檔案)、「中」檔(中央委員會相關檔案)、「部」檔(聯俄容共時期士農工商內政五部的檔案)、「監」檔(國民政府時期中監會)、「政」檔(有關內政及各項國政業務的檔案)、「特」檔(各種特種檔案)、「環」檔(環龍路檔案)、「稚」檔(吳稚暉專檔)與「膺」檔(迪鷹專檔)等十

		一類。
控制取得情況<accessrestrict>		不需著錄內容
	段落<p>	開放作為研究之用
其他檢索工具<otherfindaid>		不需著錄內容
	段落<p>	提供檔案卡片目錄
相關資料<relatedmaterial>		不需著錄內容
	段落<p>	中國國民黨一百週年大事年表（全二冊）、中國國民黨與中華民國（全一冊）
書目<bibliography>		不需著錄內容
	段落<p>	中國國民黨史述（全五冊）、中國國民黨與中國民主憲政（全一冊）、中國國民黨黨史研究論文選集（全五冊）
權威控制<controlaccess>		封包元素，不需著錄內容
	權威控制<controlaccess>	封包元素，不需著錄內容
	個人名稱(主題)<persname>	孫中山、蔣中正、蔣經國
	個人名稱（創作者）<persname>	孫中山、蔣中正、蔣經國
	權威控制<controlaccess>	封包元素，不需著錄內容
	家族名稱(主題)<famname>	蔣介石、宋子文、孔祥熙
	家族名稱（創作者）<famname>	蔣介石
	權威控制<controlaccess>	封包元素，不需著錄內容
	法人團體名稱（主題）	中國國民黨

	<corpname>	
	法人團體名稱 (創作者) <corpname>	中國國民黨
	權威控制<controlaccess>	封包元素，不需著錄內容
	地理名稱 (主題)	中國、台灣
	權威控制<controlaccess>	封包元素，不需著錄內容
	主題<subject>	興中會、同盟會、國民黨、中華革命黨、中國國民黨
	權威控制<controlaccess>	封包元素，不需著錄內容
	題名<title>	中國國民黨黨史資料與研究

表 4-7 檔案系列描述編目記錄

元素名稱		編目記錄 (測試性編目記錄)
	附屬成分描述<dsc>	封包元素，不需著錄內容
	成分 (系列層次) <c01>	封包元素，不需著錄內容
	描述識別<did>	封包元素，不需著錄內容
	單元識別<unitid>	「中央改造委員會檔案」6.43 登錄號
	單元題名<unittitle>	中央改造委員會檔案 6.43
	單元日期<unitdate>	1950-1952
	資料之語言 <langmaterial>	不需著錄內容
	語言 <language>	繁體中文

範圍與內容<scopecontent>	不需著錄內容
段落<p>	「中央改造委員會檔案」涵蓋的主題範圍：1950年8月5日，國民黨中央改造委員會成立，通過《中央改造委員會組織大綱》，蔣介石任該委員會主席，張其昀為秘書長，周宏濤為副秘書長。改造運動的主要內容有：黨員歸隊、黨員整肅、刷新黨政關係、開展民眾運動、加強社會調查、地方黨務改造、反共抗俄總動員等。1952年10月，國民黨召開第七次黨員代表大會，宣告改造完成，歷時2年多。
編排<arrangement>	不需著錄內容
段落<p>	「中央改造委員會檔案」依原有單件及其題名保存。
控制取得情況<accessrestrict>	不需著錄內容
段落<p>	開放作為研究使用
控制使用狀態<userrestrict>	不需著錄內容
段落<p>	使用者取得「中央改造委員會檔案」中的實體檔案後，由於文件脆弱，不得影印與複製。

單筆「國民黨歷史檔案詮釋資料」編目紀錄 XML 檔透過 XSLT 轉換規則文件轉換成「檔案描述編碼格式」(EAD) 編目紀錄 XML 檔之後，再藉由人工方式將「EAD 標目」與「檔案描述」中闕漏資料的欄位填入資料，則一筆「檔案描述編碼格式」(EAD) 編目紀錄包含有檔案的全宗描述、系列描述以及單件描述。

第五節 自動全集轉換實作相關問題探討

由單件層次描述轉換成全集層次描述實作過程中，難免會遭遇一些問題，研究者將問題分為兩方面進行探討，一方面是由單件層次描述建立全集層次描述相關問題；另一方面則為「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)兩詮釋資料格式互通性相關問題，以下針對這兩個方面進行探討：

一、由單件層次描述建立全集層次描述相關問題

「國民黨歷史檔案詮釋資料」主要用於描述單筆檔案之詳細特徵，提供使用者進行檔案的辨識，然而前提為使用者已經瞭解典藏機構中典藏資源的範疇與特色以及欲檢索的檔案之某一確切特徵或特質，例如檔案的創作者；「檔案描述編碼格式」(EAD)主要以全宗作為描述的單位，提供全宗所涵蓋的範圍、包含的系列、單件數量、相關使用限制情況或典藏資源位址等摘要資訊，目的為揭示深層隱藏的典藏資源，特別是當使用者不瞭解典藏機構的典藏範圍與特色時，透過建立檔案的全宗描述，可幫助使用者在查檢全宗檔案之初，就可對此全宗檔案建立起一個概略性的了解，或幫助管理者了解典藏品的相關典藏狀況。

以本章第三節詮釋資料格式對映與第四節詮釋資料格式轉換實作為基礎，發現「國民黨歷史檔案詮釋資料」現有的欄位大部分皆能對映至「檔案描述編碼格式」(EAD)中，但無法架構起國民黨歷史檔案的全宗層次描述或系列層次描述，原因包含：

1. 兩詮釋資料格式欄位的對映主要是以單件為主，因此當「國民黨歷史檔案詮釋資料」編目記錄在轉換成「檔案描述編碼格式」(EAD)編目記錄後，件的上一層次的描述例如件所屬的系列描述以及最高層次的全宗描述，皆須透過人工方式補填描述資料。
2. 「檔案描述編碼格式」(EAD)中常透過「傳記或歷史」、「範圍與內容」以及「編排」等欄位建立檔案的全宗層次描述；而「國民黨歷史檔案詮釋資料」現有欄位都是針對單筆檔案記錄藏品特質的描述，缺少針對單份檔案紀錄背景描述欄位，因此無法由目前現行的「國民黨歷史檔案詮釋資料」欄位建立起「檔案描述編碼格式」(EAD)之全宗層次描述或系列層次描述。

要由現行的「國民黨歷史檔案詮釋資料」單件描述欄位建立起「檔案描述

編碼格式」(EAD)中完整的全宗描述或系列描述是不可能的；但可藉由改善目前某些欄位以建立起部分的全宗描述或部分的系列描述，例如：

權威控制欄位：將現行「國民黨歷史檔案詮釋資料」中的「標題關鍵字」、「創作者關鍵字」、「出版地關鍵字」以及「備註關鍵字」等欄位作進一步的權威控制，在轉入「檔案描述編碼格式」(EAD)時可以用來建立全宗描述或系列描述的權威控制相關資訊，為資源檢索者提供重要的檢索點。

二、兩詮釋資料格式之互通性相關問題

詮釋資料格式之間的互通性是透過『語意的互通』以及『結構與語法的互通』來達成，換句話說，必須同時滿足共通語意和共通語法這兩個條件才可能達成互通性。

本研究在第三節建立「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)兩詮釋資料格式之間欄位的對映即是屬於語意互通性的方法；在第四節透過XSLT轉換規則文件轉換單筆「國民黨歷史檔案詮釋資料」編目記錄XML檔至「檔案描述編碼格式」(EAD)編目記錄XML檔結構，即屬於結構和語法互通性的方法。

以下將從語意及語法分析本研究中「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)兩詮釋資料格式之互通性。

(一) 語意的互通

語意(semantics)即欄位的含意。語意的互通是在對比兩種詮釋資料格式時發現語意相同或相容的元素，彼此可以相互對映。

本研究中「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)兩詮釋資料格式在欄位語意對映方面可分為以下四種情況，分別為：

1. 語意相容可直接對映的欄位。
2. 「檔案描述編碼格式」(EAD)欄位加入屬性與屬性值，讓欄位更為專指，再進行對映。
3. 語意上沒有相容的元素，無直接可對映的欄位，則對映至「檔案描述編碼格式」(EAD)的「附註」元素加入標籤屬性與屬性值之後的「段

落」元素。

4. 不進行對映轉換欄位：「國民黨歷史檔案詮釋資料」的「標題關鍵字」、「創作者關鍵字」以及「備註關鍵字」等欄位中的資料，並未明確區隔人名、機關組織名、主題與地名等關鍵字，因此在對映轉換到「檔案描述編碼格式」(EAD)之「權威控制」元素時，產生無法對映的問題，則不進行對映。

除了上述欄位語意對映的四種情況之外，「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD)兩詮釋資料格式在元素對映之間包含有「一對一關係」、「多對一關係」以及「一對零關係」等三種對映關係。

(二) 語法的互通

語法 (Syntax) 即用來傳達意義 (meaning) 以及結構 (structure) 的文法 (grammar)。可延伸標籤語言 (XML) 以一種開放的自我描述方式定義了資料結構，在描述資料內容的同時能突出對結構的描述，從而體現出資料之間的關係。

「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD) 兩詮釋資料格式皆採用 XML 作為語法的規範，即以 XML 作為資料和內容的架構並作為兩詮釋資料格式之間交換的語法。XML 提供「國民黨歷史檔案詮釋資料」與「檔案描述編碼格式」(EAD) 兩詮釋資料格式在語法層次上的互通性。

(三) 運用 XSLT 轉換詮釋資料格式技術

根據 W3C 文件，XSLT 被定義為『用來轉換 XML 文件到另一個 XML 文件的一種語言』，本研究運用 XSLT 做為詮釋資料格式轉換技術，轉換「國民黨歷史檔案詮釋資料」編目紀錄 XML 檔為「檔案描述編碼格式」(EAD) 編目紀錄 XML 檔。

不同於一般平面式詮釋資料格式轉換至平面式詮釋資料格式，本研究是由平面式詮釋資料格式結構轉換成階層式詮釋資料格式結構，因此有許多元素、屬性與屬性值必須透過 XSLT 轉換規則文件進行新增。

全集描述不同於單件描述的一大特點為，它可以動態加入單件層次描述至全集描述中；本研究透過 XSLT 轉換規則文件僅轉換單筆「國民黨歷史檔案詮

釋資料」編目記錄至一筆「檔案描述編碼格式」(EAD)編目記錄中，若能透過批次轉換的機制，一次將多筆「國民黨歷史檔案詮釋資料」編目記錄轉入一筆「檔案描述編碼格式」(EAD)編目記錄中，更可呈現出檔案全宗描述的精神。

第六節 EAD 檔案全集描述與全球化全集檢索

一、EAD 與「RSLP 全集描述綱要架構」

以「檔案描述編碼格式」(EAD)建立起的檔案全集描述，僅適用於檔案領域中全集資料的交換與檢索，如何達成全球化的全集檢索，讓使用者可以一次發現圖書館、博物館與檔案館等典藏機構中的全集資訊是很重要的事。

「RSLP 全集描述綱要架構」(RSLP CD schema)是目前很重要的全集層次詮釋資料格式，它對於現行的全集描述計畫有顯著且重大的影響，因此本節進一步探討「檔案描述編碼格式」(EAD)全集描述與「RSLP 全集描述綱要架構」中描述全集屬性(attributes)之間的對映關係。

在「檔案描述編碼格式」(EAD)中主要是以「檔案描述」的第一層來描述全集(全宗)，因此本研究就「檔案描述」第一層中較為重要且常用來描述全集之元素與 RSLP 全集描述綱要中的全集屬性進行對映。

以下進一步分析「檔案描述編碼格式」(EAD)與「RSLP 全集描述綱要架構」兩種詮釋資料格式在全集描述方面的對映關係。

二、EAD 對映「RSLP 全集描述綱要架構」

(一) EAD「描述識別」元素對映「RSLP 全集描述綱要架構」

「檔案描述編碼格式」(EAD)主要是透過「描述識別」元素進行全集特徵的描述，例如描述全集的題名、全集的典藏機構、全集含括的日期範圍或全集包含的單件數量等等資訊。

表 4-8 為「檔案描述編碼格式」(EAD)中「描述識別」元素與「RSLP 全

集描述綱要架構」全集描述對映分析：

表 4-8 EAD 「描述識別」元素對映「RSLP 全集描述綱要架構」

	EAD	RSLP CD schema
全集的識別	單元識別 archdesc/did/unitid 單元識別/@國碼 archdesc/did/unitid/@countrycode 單元識別/@典藏機構代碼 archdesc/did/unitid/@repositorycode	全集識別 (Identifier) [Collection] dc:identifier
全集的名稱	單元題名 archdesc/did/unittitle	全集題名 (Title) [Collection] dc:title
全集的日期	單元日期 archdesc/did/unitdate	內容日期範圍 (Contents Date Range) [Collection] cld:contentsDateRange
		時間 (Time) [Collection] dcq:temporal
全集典藏機構名稱	典藏機構 archdesc/did/repository/corpname	組織名稱 (Organisation name) [Administrator] vcard:org
全集典藏機構地址	典藏機構/地址 archdesc/did/repository/address	郵政位址 (Postal address) [Location] cld:address
全集數量	實體描述/數量單位	實體特徵 (Physical)

	archdesc/did/physdesc/extent	Characteristics) [Collection] dc:format
全集實體面	實體描述/實體面 archdesc/did/physdesc/physfacet	實體特徵 (Physical Characteristics) [Collection] dc:format
全集的起源	起源/個人名稱 archdesc/did/origination/persname	名稱 (Name) [Collector] vcard:fn
	起源/法人團體名稱 archdesc/did/origination/corpname	組織名稱 (Organisation name) [Collector] vcard:org
全集語言	資料之語言 archdesc/did/langmaterial/language	語言 (Language) [Collection] dc:language
全集摘要	摘要 archdesc/did/abstract	描述 (Description) [Collection] dc:description
全集附註	附註 archdesc/did/note	註釋 (Note) [Collection] ead:note
全集實體位置	實體位置 archdesc/did/physloc	無相容可對映
全集資料特殊細節	資料特殊細節 archdesc/did/materialspec	無相容可對映

EAD「描述識別」元素對映「RSLP全集描述綱要架構」結果分析：

語意方面：「描述識別」元素中用來描述全集特徵的子元素大部分都能夠對映「RSLP 全集描述綱要架構」語意相容的元素，唯有「實體位置」元素以及「資料特殊細節」元素沒有語意相容的元素可供對映。

元素與元素之間的關係：一對一、一對零、多對一。

(二) EAD 全集摘要描述元素對映「RSLP 全集描述綱要架構」

在「檔案描述編碼格式」(EAD) 中常透過「傳記或歷史」、「範圍與內容」、「控制取得狀態」以及「權威控制」等元素來描述全集之典藏背景、編排方式或取得權限等相關資訊，以下分析這些元素與「RSLP 全集描述綱要架構」之間的對映關係，請參考表 4-9：

表 4-9 EAD 全集摘要描述元素對映「RSLP 全集描述綱要架構」

	EAD	RSLP CD schema
全集傳記或歷史	傳記或歷史 archdesc/bioghist	代理者歷史 (Agent History) [Collector] cld:agentHistory
全集範圍與內容	範圍與內容 archdesc/scopecontent	描述 (Description) [Collection] dc:description
全集編排	編排 archdesc/arrangement	無相容可對映
全集採訪資訊	採訪資訊 archdesc/acqinfo	保管歷史 (Custodial History) [Collection]cld:custodialHistory
全集保管歷史	保管歷史 archdesc/custodhist	保管歷史 (Custodial History) [Collection]cld:custodialHistory
全集鑑定資訊	鑑定資訊 archdesc/appraisal	保管歷史 (Custodial History) [Collection] cld:custodialHistory

全集增加狀態	自然增長 archdesc/accruals	增加狀態 (Accrual Status) [Collection] cld:accrualStatus
全集處理資訊	處理資訊 archdesc/processinfo	無相容可對映 * 向上兼容到[Collection] dc:description
全集的控制取得狀態	控制取得狀態 archdesc/accessrestrict	存取控制 (Access Control) [Collection] cld:accessControl
全集的控制使用狀態	控制使用狀態 archdesc/userrestrict	存取控制 (Access Control) [Collection] cld:accessControl
全集其他可取得之類型	其他可取得之類型 archdesc/altformavail	無相容可對映 * 向上兼容到[Collection] dc:description
全集建議引用	建議引用 archdesc/prefercite	全集題名 (Title) [Collection] dc:title
全集其他檢索工具	其他檢索工具 archdesc/otherfindaid	目錄或描述 (Catalogue or description) [Collection] cld:hasDescription
全集相關資料	相關資料 archdesc/relatedmaterial	相關全集 (Associated collection) [Collection] cld:hasAssociation
全集附件資料	附件資料 archdesc/separatedmaterial	相關全集 (Associated collection) [Collection] cld:hasAssociation
全集相關書目	書目 archdesc/bibliography	相關出版物 (Associated publication) [Collection] cld:hasPublication

全集其他的描述 資料	其他的描述資料 archdesc/odd	註釋 (Note) [Collection] cld:note
---------------	-----------------------------	--

EAD 全集摘要描述元素對映「RSLP 全集描述綱要架構」結果分析：

語意方面：「檔案描述編碼格式」(EAD) 中較常用來描述全集範圍、背景內容特徵等元素大部分都能夠對映至「RSLP 全集描述綱要架構」中語意相容的元素，僅「編排」元素、「處理資訊」元素以及「其他可取得之類型」元素沒有語意相容的元素可供對映。

元素與元素之間的關係：一對一、一對零、多對一。

(三) EAD「權威控制」元素對映「RSLP 全集描述綱要架構」

「檔案描述編碼格式」(EAD) 透過「權威控制」元素指出全集資料中的關鍵檢索點，以下分析 EAD 中的「權威控制」元素與「RSLP 全集描述綱要架構」之間的對映關係，請參考表 4-10：

表 4-10 EAD「權威控制」元素對映「RSLP 全集描述綱要架構」

	EAD	RSLP CD schema
全集之主題權威 控制	主題 archdesc/controlaccess/subject	概念 (Concept) [Collection] dc:subject
全集之個人名稱 權威控制	個人名稱 archdesc/controlaccess/persname	名稱 (Name) [Collection] cld:agentName
全集之法人團體 名稱權威控制	法人團體名稱 archdesc/controlaccess/corpname	名稱 (Name) [Collection] cld:agentName
全集之地理名稱 權威控制	地理名稱 archdesc/controlaccess/geogname	地點 (Place) [Collection] dcq:spatial

EAD「權威控制」元素對映「RSLP全集描述綱要架構」之結果分析：

語意方面：「檔案描述編碼格式」(EAD)中較常用來描述全集中重要的個人名稱、法人團體名稱或地理名稱等「權威控制」元素大部分都能夠對映至「RSLP全集描述綱要架構」語意相容的元素。

元素與元素之間的關係：一對一、多對一。

第七節 檔案全集呈現介面

不同於單件檢索工具提供單筆檔案及其內容詳細的描述資訊(屬於解析性檢索工具);全集檢索工具僅提供全集整體性的描述資訊,不深入描述構成該全集的單筆檔案資訊(屬於整體性檢索工具)。

透過「檔案描述編碼格式」(EAD)建立的全集描述檢索工具既屬於整體性檢索工具也是階層性檢索工具,因此,除了能呈現全集的整體描述資訊之外,也能提供全集與單件之間關連以及單件相關資訊,請參考圖 4-18。使用者可在全集層次開始他們的檢索,當識別出符合所需的全集之後,再往下移動到系列層次或單件層次進行準確的檢索。

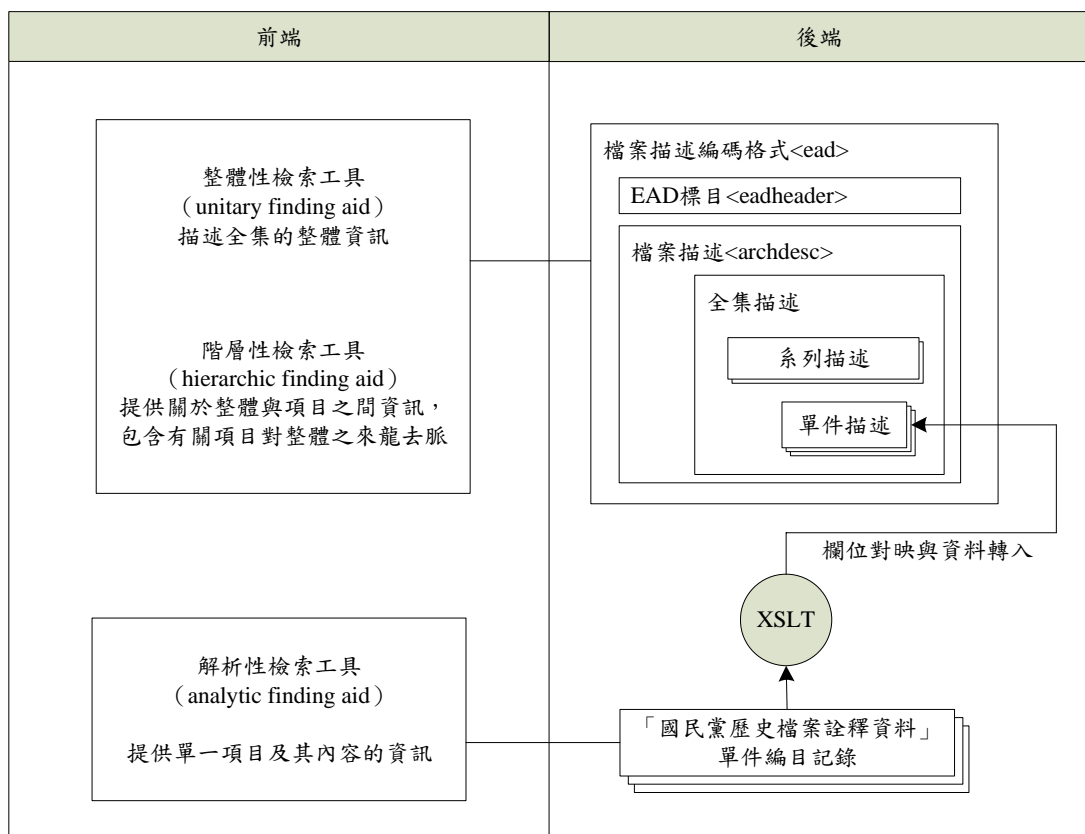


圖 4-18 整體性檢索工具與解析性檢索工具呈現示意圖

資料來源：本研究

一、現行檔案全集描述呈現介面分析

以下首先分析國外三個採行「檔案描述編碼格式」(EAD)作為檔案全集描述編碼之網站，瞭解目前檔案全集呈現介面之特徵，接著歸納合理的檔案全集呈現介面所應包含的功能。

(一) 美國「加州線上檔案」

1. 「加州線上檔案」全集呈現介面功能分析

圖 4-19 為美國「加州線上檔案」(Online Archive of California)全集呈現介面：

The screenshot shows the OAC website interface. At the top is the OAC logo and navigation links: OAC Home | About | Search | Site Map | Help. Below this is a breadcrumb trail: Finding Aids > UC Los Angeles > Special Collections, Young Research Library. The main heading is "Finding Aid for the Collection of Material about Abraham Lincoln".

The interface is divided into several sections:

- View options:** 3. Standard (selected), Entire finding aid.
- Search within this document:** 4. A search input field with a "GO" button.
- Contents:** 5. A list of links: Descriptive Summary (selected), Administrative Information, Scope and Content, Organization and Arrangement, Indexing Terms, Container List. Below this is a detailed list of materials: Books, Pamphlets, and Other Printed Materials, [v.d.], Letters, Speeches, Writings of Abraham Lincoln, Writings About Abraham Lincoln, Miscellaneous Publications Not Directly Related to Lincoln, Lincoln Organizations, Museums, and Memorials, Photographs and other Pictorial Materials Relating to Lincoln.
- Descriptive Summary:** 2. Title: Collection of Material about Abraham Lincoln, (1809-1865), 1850-1950. Collection number: 2021. Extent: 9 boxes (4.5 linear ft.), 1 oversize package. Repository: University of California, Los Angeles. Library. Dept. of Special Collections. Los Angeles, California 90095-1575. Physical location: Stored off-site at SRLF. Advance notice is required for access to the collection. Please contact the UCLA Library, Department of Special Collections Reference Desk for paging information. Language: English.

圖 4-19 「加州線上檔案」全集描述呈現介面

資料來源：Online Archive of California (2006). Online Archive of California.
Retrieved November 28, 2006, from <http://www.oac.cdlib.org/>

「加州線上檔案」全集描述呈現介面可分為五個區塊，以下針對這五個區塊中所呈現的內容與提供的功能進行分析：

(1) 全站主要導覽系統：幫助使用者瞭解目前所在的全集位置以及可以到網站中的哪裡。

(2) 全集摘要性描述資料：提供全集題名與全集資料所涵蓋日期、全集識別號、全集所包含的實體數量、全集典藏機構、全集實體位置以及全集資料語言等相關資訊。

(3) 全集描述呈現模式：提供使用者「標準瀏覽」以及「完整瀏覽」兩種全集呈現模式。

(4) 進階搜尋系統：提供使用者針對目前所在的全集作進一步的檢索，例如使用者在全集層次識別該全集為符合所需，可再進一步透過搜尋系統檢索該全集所包含的單件。

(5) 輔助性導覽系統：提供使用者目前所在全集的相關資訊，包含全集的摘要性描述連結、全集管理資訊描述連結、全集的範圍與內容描述連結、全集的組織與編排方式連結、全集的索引詞彙連結以及組成全集之更小成分資訊連結。

2. 「加州線上檔案」全集呈現介面特徵分析

「加州線上檔案」全集呈現介面之特徵包括：

- 檔案編排架構與內容呈現方面：可完整呈現檔案編排階層架構，讓使用者可以瞭解檔案層級以及各層級中檔案的歸檔方式與整體脈絡架構。
- 檔案查詢方面：提供使用者進階查詢功能，即當使用者進入某個全集之後，可透過進階查詢功能快速找到符合所需的單件檔案。
- 檔案描述方面：以檔案全集（全宗）為主要的描述單位，除了提供全集摘要性描述資料之外，也提供系列、單件與數位影像檔的描述資料。
- 檔案瀏覽方面：有提供數位影像檔以供連結。

(二) 英國「檔案中心」

1. 「檔案中心」全集呈現介面功能分析

圖 4-20 為英國「檔案中心」(Archives Hub) 檔案全集描述呈現介面：

The screenshot displays the Archives Hub interface for the 'Glasier Papers ca.1879-1975' collection. The page is divided into three main sections:

- Table of Contents (Left):** Lists various document types with a large number '3' indicating the count. The items include: General correspondence, Public letters, Family letters, Related papers, Diaries (with sub-items for John Bruce Glasier and Katharine Bruce Glasier), Notes (with sub-items for John Bruce Glasier notebooks, John Bruce Glasier loose notes, Katharine Bruce Glasier notebooks, Katharine Bruce Glasier loose notes, and Various authors notes), Newscuttings, Loose cuttings, Newspapers, Scrapbooks, and Printed ephemera.
- Main Content (Right):** Provides detailed metadata for the collection, including Author (Challinor, Emma and Clare Tunstall), Held at (University of Liverpool, Special Collections and Archives), Reference and contact details (GB 141 GP), Title (Glasier Papers), Dates of Creation (ca.1879-1975), Extent (49 boxes 2 portfolios and 62 printed books), Name of Creator (John Bruce Glasier (1867-1960) and Katharine Conway (1867-1960) later Katharine Bruce Glasier), Level of Description (fonds), Revisions (October - November 2004, File Modified), and Published by (University of Liverpool Library, Special Collections and Archives November 2004). Below this is a 'Biographical Note' starting with "...many thousands have made their first contact with Socialist principles through the writings and speeches of the Glasiers" (Labour Party Pamphlets *Introducing Labour's Special Campaigners*).
- Footer (Bottom):** Contains navigation links: Home, Quick Search, Browse Indexes, Subject Finding, and Advanced Search. It also features a 'Print Results' button, a 'Print Description' button, and the 'cheshire | 2' logo.

圖 4-20 「檔案中心」全集描述呈現介面

資料來源：Archives Hub (2006). Archives Hub. Retrieved November 28, 2006, from <http://www.archiveshub.ac.uk/>

「檔案中心」全集描述呈現介面可分為三個區塊，以下針對這三個區塊中所提供的功能與呈現的內容進行分析：

(1) 全站主要導覽系統：指引使用者可往何處去，例如指引使用者回到首頁、檢索求助或瀏覽索引。

(2) 檔案全集描述：呈現檔案全集整體摘要性描述資訊。

(3) 輔助性導覽系統：提供使用者目前所在全集之階層性相關資訊。

2. 「檔案中心」全集描述呈現介面特徵

「檔案中心」全集呈現介面之特徵包含以下幾個特點：

- 檔案描述方面：大部分僅提供全宗（全集）描述資訊，少部分會提供完整的全宗描述、系列描述以及單件描述。
- 檔案編排架構與內容呈現方面：若提供完整的全宗描述時，可呈現出檔案編排階層架構。
- 檔案查詢方面：當使用者進入一個全宗之後，並無提供進階查詢功能。
- 檔案瀏覽方面：無提供數位影像檔。

(三) 美國「西北數位檔案」

1. 「西北數位檔案」全集呈現介面功能分析

圖 4-21 為美國「西北數位檔案」(Northwest Digital Archives, 簡稱 NWDA) 全集描述呈現介面：

The screenshot shows the NWDA website interface. At the top, there is a navigation bar with links for search, about NWDA, participating institutions, tools for members, contact us, and home. The main content area is titled "Guide to the Sons of Union Veterans of the Civil War Records 1913-1978" and is compiled by Geoffrey B. Wexler. The page is divided into two main sections: "Overview of the Collection" (labeled 2) and "Content Description" (labeled 3). The "Overview of the Collection" section provides details about the repository (Oregon Historical Society Research Library), the sponsor (National Endowment for the Humanities), and the collection number (Mss 2937). The "Content Description" section lists various categories such as Historical Note, Arrangement, Administrative Information, and Subjects.

Contents	
OVERVIEW OF THE COLLECTION	
HISTORICAL NOTE	
CONTENT DESCRIPTION	3
ARRANGEMENT	
ADMINISTRATIVE INFORMATION	
Acquisition Information	
Separated Materials	
USE OF THE COLLECTION	
Restrictions on Access	
Restrictions on Use	
Preferred Citation	
SUBJECTS	
DETAILED DESCRIPTION OF THE COLLECTION	

1

search about NWDA participating institutions tools for members contact us home

Guide to the Sons of Union Veterans of the Civil War Records 1913-1978

Finding aid compiled by Geoffrey B. Wexler

© 2006

Overview of the Collection **2**

Repository: Oregon Historical Society
Research Library

1200 SW Park Avenue
Portland, Oregon 97205
Phone: 503-306-5240
Fax: 503-219-2040
E-mail: libreference@ohs.org

Sponsor: Funding for encoding this finding aid was provided through a grant awarded by the National Endowment for the Humanities.

Collection Number: Mss 2937

圖 4-21 「西北數位檔案」全集描述呈現介面

資料來源：Northwest Digital Archives (2006). Northwest Digital Archives.
Retrieved November 28, 2006, from <http://nwda.wsulibs.wsu.edu/>

「西北數位檔案」全集描述呈現介面可分為三個區塊，以下針對這三個區塊所呈現的內容與提供的功能進行分析：

(1) 全站主要導覽系統：引導使用者可往何處去。

(2) 檔案全集描述：呈現詳細且完整的檔案資源全集描述資料。

(3) 輔助性導覽系統：提供使用者目前所在全集的相關資訊，包含：全集之概要連結、全集歷史附註連結、全集內容描述連結、全集的編排方式連結、全集的管理資訊連結、全集的使用限制資訊連結、全集中相關主題關鍵字索引連結以及全集的詳細描述連結。

2. 「西北數位檔案」全集描述呈現介面特徵

「西北數位檔案」全集呈現介面之特徵包括：

- 檔案編排架構與內容呈現方面：可呈現檔案編排階層架構。
- 檔案查詢方面：當使用者進入一個全集之後，並無提供進階查詢功能。
- 檔案描述方面：提供完整的全宗描述、系列描述、單件描述以及數位影像檔描述。
- 檔案瀏覽方面：有提供數位影像檔以供連結。

二、合理的檔案全集呈現介面

檔案全集由於其特殊的編排方式與描述慣例不同於一般全集，因此呈現介面表達方式與一般全集有所不同。

(一) 合理的檔案全集呈現介面表達

合理的檔案全集呈現介面需表達與反映檔案在典藏與管理上所重視的「整

體編排」與「描述的完整性」。

以檔案在典藏與管理上重視的三大原則、多層次描述以及上述三個現行檔案全集呈現介面分析為基礎，以下歸納合理的檔案全集呈現介面需如何表達：

1. 檔案編排架構與內容呈現方面：檔案編排方面最重要的就是尊重檔案的「全宗原則」、尊重檔案的「來源原則」以及尊重檔案的「原始順序原則」。因此檔案全集在呈現介面上必須能夠表達出這三個重要原則，透過這三個重要原則的實踐，讓使用者可以看到檔案的層級（可以瞭解單件檔案的起源，起源原則），以及各層級中實體文件的歸檔方式（排列反映出一個組織或個人記錄它們最初的工作順序，原始順序原則）。
2. 檔案描述方面：由於檔案處理過程強調層級的處理與層級之間的關係，以維護檔案原始歷史證據，因此，需呈現「多層次描述」以反映各檔案層級之間的關係。
3. 檔案查詢方面：可支援某特定項目的查詢檢索。

（二）合理的檔案全集呈現介面需提供的功能

綜合上述分析，歸納合理的檔案全集呈現介面需提供給使用者的功能包含：

1. 全集瀏覽方面：提供使用者全集簡略描述與全集詳細描述兩種瀏覽模式：
 - (1) 全集簡略描述瀏覽模式：僅提供全集摘要性描述資料。例如，全集題名與全集資料所涵蓋日期、全集識別號、全集所包含的實體數量、全集典藏機構以及全集資料之語言等資訊。
 - (2) 全集詳細描述瀏覽模式：呈現一筆完整且詳細的「檔案描述編碼格式」(EAD) 編目記錄。
2. 導覽系統方面：幫助使用者方便瀏覽全集內容，由於一份「檔案描述編碼格式」(EAD) 檢索工具中有時包含數以萬計的單件檔案，若不提供使用者相關的導覽，使用者將容易迷失於其中。導覽系統可進一步分為：
 - (1) 全站導覽：引導使用者可往何處去。包含：指引使用者回到網站首頁、提供關於本網站之歷史、網站地圖以及提供相關協助。

- (2) 局部導覽：提供使用者目前所在全集的相關資訊。包含：全集的摘要性描述、全集管理資訊描述、全集的範圍與內容描述、全集的組織與編排方式、全集的索引詞彙以及組成全集之更小成分資訊。
3. 搜尋系統方面：當使用者在全集層次開始他們的檢索，識別符合所需的全集之後，為了幫助使用者快速找出該全集當中符合所需的系列或單件，必需提供使用者可以檢索目前所在全集之功能。
4. 數位物件連結模式方面：當全集中包含有數位物件時，可先透過局部導覽，提示使用者該全集當中包含的數位物件之數量以及顯示數位物件所在之位置，方便使用者進行數位物件的連結使用。