

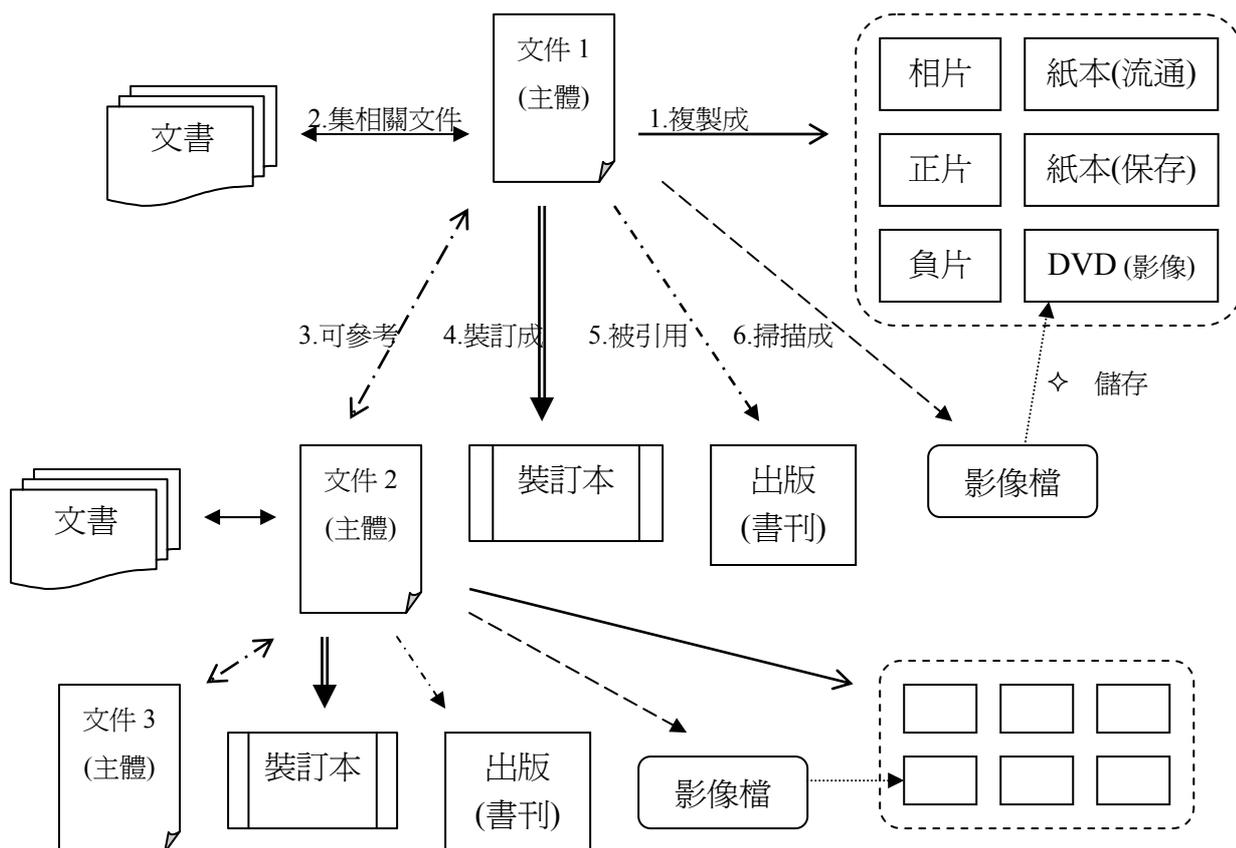
臺史所古文書需求分析與後設資料標準應用研析

中央研究院 後設資料工作組

92/07/10

一、古文書關係結構分析

分析古文書之藏品元素結構關係，可得以下關係圖：



館藏之著錄以文件為單位，文件又因為管理需求、館藏整理、研究應用、內涵關聯性之不同而衍生出 6 種關係：

1. 因保存及流通上的需要，文件會複製成 6 種不同形式之重製品。
2. 若一批文件是與某一個人有關的，則除著錄個別文件名稱外，也會該批文件一個統一的文書名稱。
3. 文件可能有其他相關文件可供參考，著錄時著錄相關文件之文件編號，以提供連結。
4. 館藏整理時，曾依主題聚集相關文件，裝訂成冊，例：「淡水文書」，著錄時著錄裝訂本之索書號，及文件所在頁碼。
5. 文件被引用刊載於出版書刊中，著錄書刊之書目資料。
6. 文件會掃描成影像檔，需提供連結。影像檔會另存成 DVD 以做保存。

二、系統需求

1. 研究與管理並重：除了希望數位化內容可供研究者使用外，因為古文書會複製成多種媒體以供保存，因此也需要能滿足對於各種媒體的管理需求。
2. 書目資料不重複：因各種媒體對於內容的描述上都是一樣的，因此不希望檢索時出現多筆內容相同的書目紀錄。

三、後設資料標準應用情形評估分析

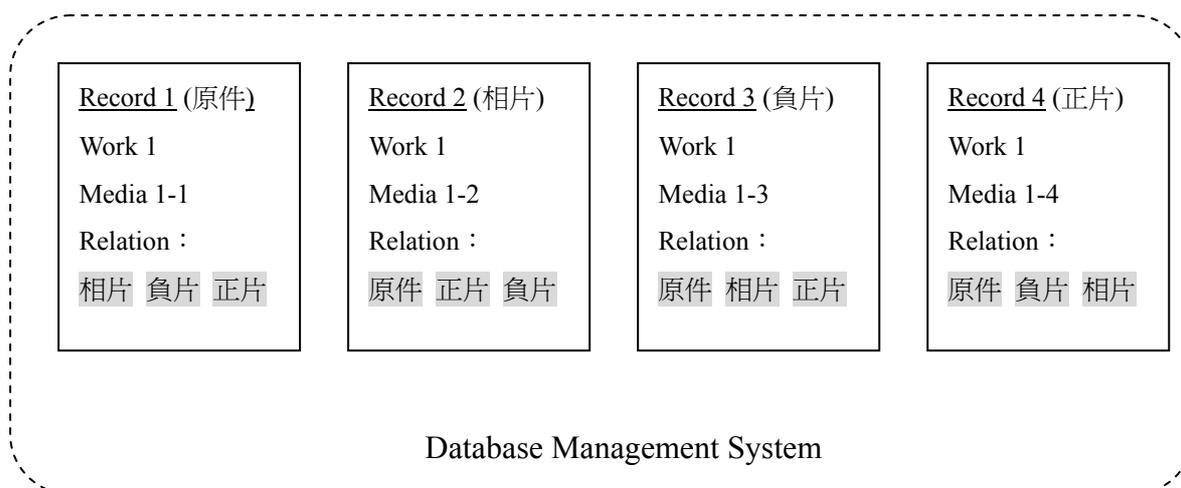
1. MICI

MICI(中文詮釋資料交換格式)是根據 Dublin Core 發展而成的 metadata 標準，因此，在應用上也承襲了 Dublin Core 的 1:1 原則，即一筆紀錄(record)僅描述一個物件(資源)，其他相關資源則藉由 Relation(關聯)來與該物件做連結。

因此，以臺史所古文書的個案看來，應用 MICI 時，應將一件文件著錄為一筆資料，而其餘相關資源，包括：複製、裝訂、參考、引用、掃描等關係，皆應分別著錄為個別的紀錄，再透過 Relation(關聯)相互連結。例：就複製品而言，文件會複製成六種不同形式，因此連同原件必須建立 7 筆紀錄，再透過 Relation 互相參照關聯，在資料庫管理系統中，將呈現出下圖之結構：

Work：知識內容

Media：媒體(載體)相關紀錄



此應用方式必須建立多筆不同媒體但資料內容相同的書目紀錄，因此並不符合臺史所的需求。

2. MARC 21

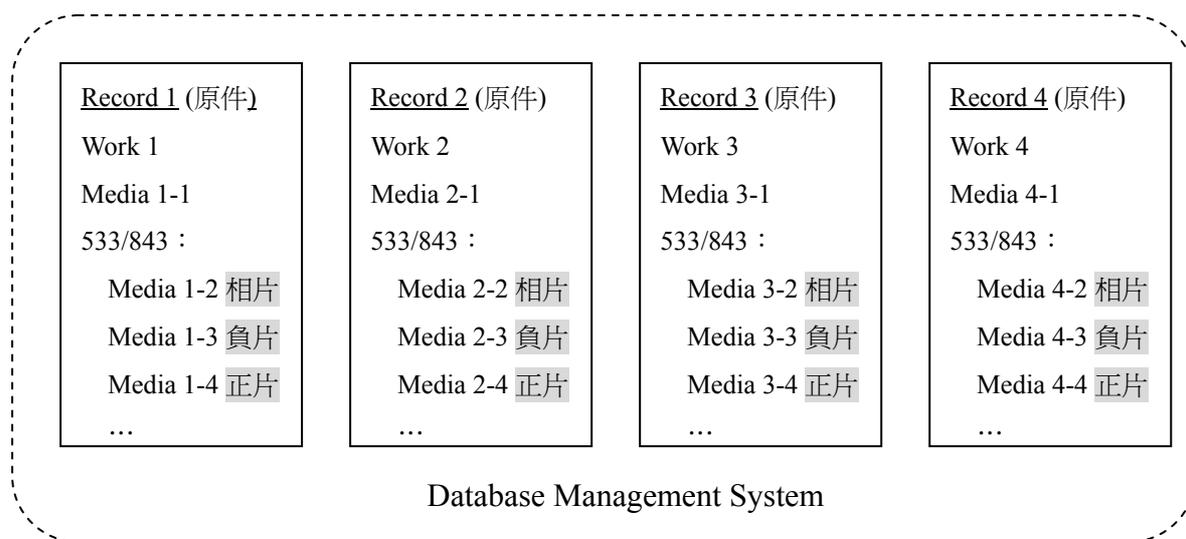
MARC 可以滿足描述物件與其他物件間多種關係的需求，當以原件文件作為著錄單元時：

- 各項複製品可放在 533 複製品註或 843 複製品註。
- 文件如與其他文件間具有相關性而有「文書名稱」或「裝訂本」時，可在 440 集叢項作描述。
- 文件有其他相關文件可供參考時，相關文件資訊可放在 787 連接款目段，以連結到相關文件。
- 文件被引用而需著錄相關出版書刊之書目資料時，可放在 581 出版品註。
- 文件需連結影像檔時，影像檔的相關資訊及使用限制可著錄於 856。

因此，應用 MARC 時，在資料庫管理系統中，將呈現出如下結構：

Work：知識內容

Media：媒體(載體)相關紀錄



此標準之應用結果，是以原件做為描述的主體，而將各式複製品及各項關聯資訊視為是原件的附屬描述資訊，因此原件與複製品將明顯被劃分為主從的關係。

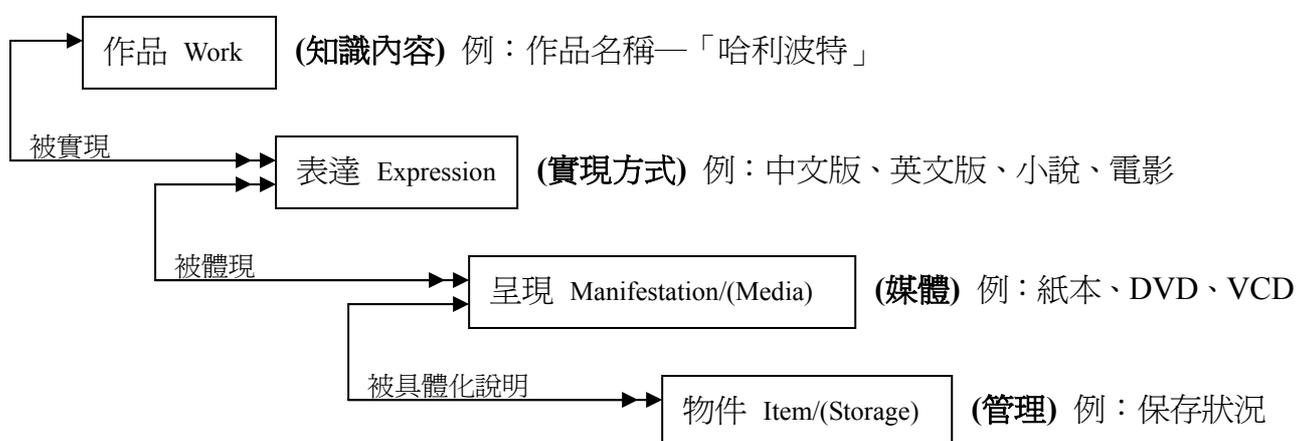
3. EAD

應用 EAD 最重要的目的是在表達檔案整體結構的層級關係，古文書的館藏結構如果不具層級關係，則應用 EAD 的意義不大，故對 EAD 之於古文書的應用情形，便不加以探討。

四、導入 IFLA FRBR Model 概念

古文書室館藏最特殊的地方在於其擁有多種不同類型之複製品，為達到研究與管理並重的目的，避免重複著錄多筆資料內容相同之紀錄造成人力及時間上的浪費，以及檢索時產生大量資料內容相同的紀錄所帶來的困擾，Metadata 工作組嘗試將 IFLA FRBR Model 的概念導入本計畫中，希望藉由 FRBR Model 的功能，將不同層次的描述屬性區分開來，再進一步透過標準比對的動作，找出適用的後設資料標準。

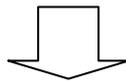
FRBR (Functional Requirements for Bibliographic Records, 書目紀錄功能需求) Model 由國際圖書館學會聯盟(IFLA)於 1998 年提出，關係圖如下所示：



將 FRBR Model 導入古文書計畫後，便可將需求元素依四個層次的屬性分析出來，內容描述的元素屬於作品層次，作品的實現方式則包括文書、照片、明信片等幾種，而呈現的媒體除了原件之外，也包含各式複製品，最後則是呈現的各項管理元素。

詳細元素區分請見下頁

Work+ Expression	文件名稱 其他名稱(土語名稱) 藏品類型 文書語文		
	內容： 全文資料檔名(文書) 立契人/發文者(私文書/公文書) 姓名 身份/職稱 團體/單位 收文者(公文書) 姓名 身份/職稱 團體/單位 關係人(契約-證人、代筆人) 姓名 身份/職稱 文書角色 作成時間(契約)/事件時間(照片/明信片) 朝代 年號 年 月 西元記年	內容： 照片明信片原有說明(照片/明信片) 必要註記(照片/明信片) 創作者/作者(照片/明信片) 姓名 身份/職稱 團體/族群 製作方式 作成時間(契約)/事件時間(照片/明信片) 朝代 年號 年 月 西元記年	
	作成地點/事件地點(契約/片/明信片) 土名 今地名 相關地方 相關族群 印記 種類 內容	稽核項-插圖 主題 主要主題 次要主題 關鍵詞 文書形式類別	相關參照註 文書名稱 出版 書刊名 ISBN/ISSN 出版資訊(地、者、年)



Manifestation (media)	原件與否 手稿與否 資料格式(原件、紙本-流、紙本-保、相片、負片、正片、影像檔、DVD)			
	原件： 稽核項 頁數 高廣尺寸 顏色 材質 附件 裝訂本 索書號 頁碼	影像： 檔案格式 檔案大小 解析度	DVD： (無相關紀錄)	其他： 數量 規格尺寸



Item (storage)	原件： 典藏登錄號 舊有編碼 典藏地+排架號 入藏 來源 取得方式 採集者 採集地 價格 同批入藏資料代號- (批號) 其他註記 原件取得日期 保存狀況 重製品相關持有紀錄 展覽 展覽名稱 展覽地點 展覽主辦單位 展覽日期 編目/管理紀錄 編目者 編目日期 核對更新者 資料管理者 最近一次修改紀錄- (修改者+修改時間) 權限範圍 典藏單位 典藏單位國家 著作權所有人 使用權限	影像： 檔案編號-(檔案名稱) 儲存主機名稱位置 URI 入藏日期 開放使用日期	DVD： 編號 典藏地+排架號 入藏日期	其他： 典藏號 典藏地+排架號 入藏日期 開放使用日期(流、相)
-------------------	--	---	-------------------------------	--

五、建議欄位結構

為簡化著錄層次，我們依據研究與管理兩種不同導向將作品(work)與表達(expression)合併成為內容層次，而呈現(Manifestation)與物件(Item)合併為實體層次，並且將內容層次與實體層次的元素先予以結構化如下表所示：

(說明：顏色標示者代表調整過後的結構)

■ 內容層次

內容 (work)				備註	建議說明
原始結構		建議欄位			
文件名稱		文件名稱			
其他名稱(土語名稱)		其他名稱			
藏品類型		藏品類型		文書/照片/明信片	
文書形式類別		文書形式類別			
文書語文		語文			文書語文調整為語文，因此欄位不單指文書語文，也同時為明信片上所記載之語文。
內容	全文資料檔名(文書)	內容	電子全文	連結電子全文	電子全文用以連結全文電子檔。
	照片名片原有說明(照片/明信片)		內容說明文		內容說明文可直接紀錄原有文字或註記說明。
	必要註記(照片/明信片)				
立契人/發文者(私文書/公文書)	姓名	主要創作者	姓名		立契人、發文者、創作者、作者分別為私文書、公文書、照片、明信片之主要創作者，可將這些項目合併，再以角色來辨別主要創作者在不同類型藏品中所扮演的角色。
	身份/職稱		角色	立契人/發文者/創作者/作者/拍攝者/製作者/繪製者	
	團體/單位		身份/職稱		
創作者/作者(照片/明信片)	姓名		團體/單位/族群		
	身份/職稱				
	團體/族群				
	製作方式				
收文者(公文書)	姓名	關係人	姓名		收文者、關係人分別為與公文書、契約內容有相關的人物，可合併為關係人，再以角色來辨別關係人在不同類型藏品內容中所扮演的角色。
	身份/職稱		角色	收文者/證人/代筆人	
	團體/單位		身份/職稱		
關係人(契約-證人、代筆人)	姓名		團體/單位/族群		
	身份/職稱				
	文書角色				
作成時間(契約/事件時間)(照片/明信片)	朝代	時間	類別	作成時間/事件時間	時間可分為作成時間及事件時間，如果以「作成時間/事件時間」為一欄，
	年號		中曆	朝代	
	年			年號	
	月			年	

內容 (work)					
原始結構		建議欄位		備註	建議說明
	西元記年		月		不易辨別時間類型，故建議以類別一欄來區分，並將時間分為中曆與西曆。
			西曆		
作成時間/事件 時間(契約/照片 明信片)	土名	地點	類別	作成地點/事件地點	地點可分為作成地點及事件地點，如果以「作成地點/事件地點」為一欄，不易辨別該地點的意義，故建議以類別一欄來區分。
	今地名		土名		
	相關地方		今地名		
			相關地方		
相關族群		相關族群			
插圖		插圖			
印記	種類	印記	種類		
	內容		內容		
主題	主要主題	主題	主要主題		
	次要主題		次要主題		
關鍵詞		關鍵詞			
相關參照註		相關參照註			
文書名稱		文書名稱			
出版	書刊名	出版	書刊名		建議將出版資訊中不同的資訊分開著錄。
	ISBN/ISSN		ISBN/ISSN		
	出版資訊(地、 者、年)		出版地		
			出版者		
			出版年		

■ 實體層次

實體(Manifestation+ Item)					
實體-原件					
原始結構		建議欄位		備註	建議說明
原件與否		原件與否			
手稿與否		手稿與否			
資料格式		資料格式			
典藏登錄號		典藏登錄號			
舊有編碼		舊有編碼			
典藏地+排架號		典藏位置	典藏地		「典藏地+排架號」包含兩項資訊，定義為典藏位置，建議可將欄位名稱調整為典藏位置，並將典藏地及排架號分項著錄。
			排架號		
稽核項	頁數	稽核項	頁數		
	高廣尺寸		高廣尺寸		
	顏色		顏色		
	材質		材質		
	附件		附件		

裝訂本	索書號	裝訂本	索書號		
	頁碼		頁碼		
入藏	來源	入藏	來源		原件取得日期亦屬入藏資訊之一，建議合併載入藏資訊中。
	取得方式		取得方式		
			取得日期		
	採集者		採集者		
	採集地		採集地		
	價格		價格		
	同批入藏資料代號		同批入藏資料代號		
其他註記	其他註記				
原件取得日期					
保存狀況		保存狀況			
重製品相關持有紀錄		重製品相關持有紀錄			
展覽名稱		展覽	名稱		
展覽地點			地點		
展覽主辦單位			主辦單位		
展覽日期			日期		
權限範圍	典藏單位	權限範圍	典藏單位		
	典藏單位國家		典藏單位國家		
	著作權所有人		著作權所有人		
	使用權限		使用權限		
編目/管理紀錄	編目者	編目/管理紀錄	編目者		最近一次修改紀錄中需紀錄修改者+修改時間，建議可與核對更新者合併為更新者及更新日期。
	編目日期		編目日期		
	核對更新者		更新者		
	資料管理者		更新日期		
	最近一次修改紀錄		資料管理者		
實體-影像					
原始結構		建議欄位		備註	建議說明
原件與否		原件與否			
手稿與否		手稿與否			
資料格式		資料格式			
檔案編號-(檔案名稱)		檔案編號			
檔案格式		檔案格式			
檔案大小		檔案大小			
解析度		解析度			
儲存主機名稱位置		儲存主機名稱位置			
URI		URI			
入藏日期		入藏日期			
開放使用日期		開放使用日期			
實體-DVD					
原始結構		建議欄位		備註	建議說明
原件與否		原件與否			
手稿與否		手稿與否			
資料格式		資料格式			
編號		編號			
典藏地+排架號		典藏位置	典藏地		「典藏地+排架號」包含兩項資訊，定義為典藏位置，建議可將欄位名稱調整為典藏位置，並將典藏地及排架號分項著錄。
			排架號		
入藏日期		入藏日期			

實體- (紙-流、紙-保、相片、負片、正片)				
原始結構	建議欄位		備註	建議說明
原件與否	原件與否			
手稿與否	手稿與否			
資料格式	資料格式			
典藏號	典藏號			
典藏地+排架號	典藏位置	典藏地		「典藏地+排架號」包含兩項資訊, 的定義為典藏位置, 建議可將欄位名稱調整為典藏位置, 著錄時填入典藏地及排架號。
		排架號		
數量	數量			
規格尺寸	規格尺寸			
入藏日期	入藏日期			
開放使用日期	開放使用日期			

問題：

1. 請確認是否每件複製品皆須辨識「原件與否」及「手稿與否」？
2. 「實體-原件」的部分是否有其他相關資訊是其他複製品也需要著錄到的，例如：保存狀況、展覽、權限範圍等？

六、元素清單與標準比對(MICI、MARC21)

爲了解本計劃需求元素與標準比對的情形，我們將結構化之後的元素依內容及實體兩個層次列出各層次之元素清單，與 MICI 跟 MARC21 做比對，希望藉由標準比對的結果，找出適用本計畫之後設資料標準。

■ 內容層次

內容 (work)					
建議欄位		MICI	MARC21	備註	
文件名稱		題名-主要題名	245\$a		
其他名稱		題名-其他題名	246\$a		
藏品類型		資源類型-自訂藏品類型	245\$k		
文書形式類別		資源類型-自訂藏品類型	245\$k		
語文		語文-作品語文	546\$a		
內容	電子全文	關連-has format	856\$f		
	內容說明文	簡述-內容摘要/內容提要	520\$a		
主要創作者	姓名	創作者-姓名	100\$a		
	角色	創作者-角色	100\$e/100\$g		
	身份/職稱		100\$c		
	團體/單位/族群	創作者-團體/族群	100\$u		
關係人	姓名	其他貢獻者-姓名	600\$a		
	角色	其他貢獻者-角色	600\$e/600\$g		
	身份/職稱		600\$c		
	團體/單位/族群	其他貢獻者-團體/族群	600\$u		
時間	類別	時空涵蓋範圍-時間類別-事件時間	648\$3		
	中曆	朝代	時空涵蓋範圍-時間類別-事件時間	648\$a	
		年號	日期-製作日期		
		年			
		月			
西曆	時空涵蓋範圍-時間類別-事件時間	648\$a			
地點	類別	時空涵蓋範圍-空間類別-事件地點	651\$3		
	土名	時空涵蓋範圍-空間類別-事件地點	651\$a	兩標準均有用以描述相關地名的欄位，但皆無法明確分辨土名、今地名及相關地方。	
	今地名	時空涵蓋範圍-空間類別-事件地點	651\$a		
	相關地方	時空涵蓋範圍-空間類別-涵蓋範圍等級	522\$a		
相關族群	主題和關鍵詞-控制詞彙-團體/族群	650\$a	兩者均有控制詞彙的意義，但不能辨識此欄位說明的是「相關族群」。		
插圖		簡述-稽核項-插圖	300\$b		
印記	種類	簡述-印記種類	500\$a		
	內容	簡述-印記種類	500\$a		
主題	主要主題	主題與關鍵詞-控制詞彙-主要主題	650\$a		
	次要主題	主題與關鍵詞-控制詞彙-次要主題	650\$a		
關鍵詞		主題與關鍵詞-關鍵詞彙	653\$a		
相關參照註		關連-相關參照 references	787\$w		

文書名稱	(關連-is part of)	440\$a	Is part of 為連結相關資源的欄位，相較之下，MARC 可正確對應到 440\$a 集叢名。
出版	書刊名	(關連-is reference by)	581\$a
	ISBN/ISSN		581\$z
	出版地		
	出版者		
	出版年		
共 18 大項	含子元素共 39 項	實際滿足 31 項(79%)	實際滿足 36 項(92%)

■ 實體層次

實體(Manifestation+ Item)元素清單									標準比對		
建議欄位	原件	影像	DVD	紙-流	紙-保	相片	正片	負片	MICI	MARC21	備註
原件與否	✓			✓	✓				原件與否		
手稿與否	✓										
資料格式	✓	✓	✓	✓	✓	✓	✓	✓	資料格式-媒體類型	245\$h	
舊有編碼	✓										
典藏登錄號	✓								識別資料-登錄號	852\$j	
DVD-編號			✓						識別資料-登錄號	852\$j	
其他複製品-典藏號				✓	✓	✓	✓	✓	識別資料-登錄號	852\$j	
影像	檔案編號	✓							識別資料-登錄號	856\$f	
	檔案格式		✓						資料格式-媒體類型	856\$q	
	檔案大小		✓						資料格式-檔案大小或數量單位	856\$s	
	解析度		✓								
	儲存主機名稱位置		✓							856\$n	
	URI		✓							識別資料-URI	856\$u
其他複製品	數量	✓		✓	✓	✓	✓	✓	資料格式-檔案大小或數量單位-數量	300\$a	
	規格尺寸	✓		✓	✓	✓	✓	✓	資料格式-檔案大小或數量單位-尺寸	300\$c	
稽核項	頁數	✓							資料格式-檔案大小或數量單位-數量	300\$a	
	高廣尺寸								資料格式-檔案大小或數量單位-尺寸	300\$c	
	顏色								簡述-稽核項-色澤	300\$b	
	材質	✓							簡述-稽核項-材質	300\$b	
	附件	✓							簡述-稽核項-附件	300\$e	
裝訂本	索書號	✓							(關連-is part of)	(440-)	Is part of 為連

實體(Manifestation+ Item)元素清單										標準比對		
建議欄位		原件	影像	DVD	紙-流	紙-保	相片	正片	負片	MICI	MARC21	備註
	頁碼	√										接相關資源的欄位，因此沒有正確欄位可供比對；440為叢書項，但亦無正確欄位可比對到索書號及頁碼。
入藏	來源	√								簡述-入藏-來源	541\$a	
	取得方式	√								簡述-入藏-方式	541\$c	
	取得日期 (入藏日期)	√	√	√	√	√	√	√	√	日期-取得日期 入藏日期	541\$d	
	採集者	√										
	採集地	√								簡述-採集或遺址資訊-採集或遺址位置	541\$b	
	價格	√								簡述-入藏-價格	541\$h	
	同批入藏資料代號	√									541\$e	
	其他註記	√								簡述-附註項	500\$a	
保存狀況		√								簡述-保存狀況	583\$l	
開放使用日期			√		√		√					
重製品相關持有紀錄		√									(535\$a) (535\$3)	MICI 無相關欄位可供記載；MARC 有535 原件/副本所在註可著錄持有人，及持有之資料類型。建議可將此項資訊區分成持有者及持有紀錄。
典藏位置	典藏地	√		√	√	√	√	√	√		852\$b	
	排架號	√		√	√	√	√	√	√	識別資料-排架號	852\$c	
展覽	名稱	√								簡述-展覽-展覽名稱	585\$a	
	地點	√									585\$a	
	主辦單位	√									585\$a	
	日期	√									585\$a	
權限範圍	典藏單位	√								權限範圍-收藏者-名稱	852\$a	
	典藏單位國家	√								權限範圍-收藏者-國家	852\$n	
	著作權所有人	√								(權限範圍)	540\$b	MICI 雖沒有為著作權所有人及使用權限規劃出子欄位，但權限範圍可包含此二者的意義。
	使用權限	√								(權限範圍)	540\$a	
編目/管理紀錄	編目者	√									(040\$a)	
	編目日期	√								日期-編目日期	008 00-05	
	更新者	√									(040\$d)	

實體(Manifestation+ Item)元素清單									標準比對		
建議欄位	原件	影像	DVD	紙-流	紙-保	相片	正片	負片	MICI	MARC21	備註
更新日期	√								日期-最後修改日期	005	
資料管理者	√										
共 19 大項	含子元素共 48 項								實際滿足 29 項(60%)	實際滿足 36 項(75%)	

七、綜合比較說明與建議：

※ 元素與標準比對之滿足率：

1. 內容層次：

- 共包含 39 項元素，MICI 可確實比對到其中 31 項，，MARC 可確實比對到其中 36 項，可滿足 92%的元素需求。
- 比較兩者標準無法比對到的項目如下表：

無法滿足的項目			
MICI		MARC	
主要創作者	身份/職稱		
關係人	身份/職稱		
文書名稱			
出版	書刊名	出版	
	ISBN/ISSN		
	出版地		出版地
	出版者		出版者
	出版年		出版年

2. 實體層次：

- 共包含 48 項元素，MICI 可確實比對到其中 29 項，可滿足 60%的元素需求，MARC 可確實比對到其中 36 項，可滿足 75%的元素需求。
- 比較兩者標準無法比對到的項目如下表：

無法滿足的項目			
MICI		MARC	
		原件與否	
手稿與否		手稿與否	
舊有編碼(文件)		舊有編碼(文件)	
影像	解析度	影像	解析度
	儲存主機名稱位置		
裝訂本	索書號	裝訂本	索書號
	頁碼		頁碼
入藏	採集者	入藏	採集者
	同批入藏資料代號		
重製品相關持有紀錄		重製品相關持有紀錄	

典藏位置	典藏地		
展覽	地點		
	主辦單位		
	日期		
權限範圍	著作權所有人		
	使用權限		
編目/管理紀錄	編目者	編目/管理紀錄	編目者
	更新者		更新者
	資料管理者		資料管理者

3. 建議方案：

(1). 建議採用的標準應方式：

採用 MARC21+FRBR Model

原因：a.元素語意上的滿足率(元素與 MARC21 標準比對)：MARC21 可滿足大部分的元素需求。

b.結構呈現：將 FRBR Model 導入 MARC21 之應用方式是可行的，此法可取代原來只應用 MARC21 時原件與複製品間之主從附屬關係，進而將各種複製品提升到與原件平等的地位。

(2). 其他方案

採用 MICI

a. 元素語意上的滿足率(元素與 MICI 標準比對)：MICI 同樣可滿足大部分的元素需求。

b. 結構呈現：選擇採用此標準，必須遵循一件媒體建立一筆記錄的原則(DC 1:1 原則)，此方式之應用結果將產生多筆資料內容相同的書目紀錄，同時也會使得在建檔的作業上也必須付出相當的人力與時間成本。