

傅斯年圖書館善本書後設資料標準適用分析與建議報告

數位典藏國家型科技計畫--後設資料工作組

2002/12/04

一、前言

本報告為分析傅斯年圖書館善本書著錄時不同標準之適用性，將採用『解學士文集 十卷』為例子，分別套用『MARC21』及『EAD』兩套標準，提出其建檔、呈現及工作流程上的差異，並依據主題計畫目前的現狀與未來目的以及資源評估何者較為適用。

報告中將先就『解學士文集 十卷』的資料架構進行結構分析，進而分別就『MARC21』、『EAD』兩套標準，提出其建檔方式及呈現畫面之差異，而套用標準的方式又分成『MARC』、『EAD』、『MARC+EAD』，同時並提出分析說明與建議。

二、個案資料架構

『解學士文集 十卷』

第一層

題名：解學士文集 十卷
版心題名：解春雨先生文集 十卷
版本：明嘉靖間（1522-1566）重刻本
數量/高廣：10 冊；26 公分
序跋：明天順元年黃諫序，八年蔡朔序，明嘉靖四十一（壬戌）年羅洪先序
收藏印記：群碧樓，百靖齋，嘉靖刻本，鄭杰之印等印記
附註：傳圖全彩
排架號：檜木櫃 78-1
登錄號：180894-180903
主題：善本-集部-別集類；明刊本

第二層

第一冊
■ 文本前
羅洪先序
黃諫序
■ 卷一
目錄
應制集

第二冊
■ 卷二
目錄
四言古
五言古

第三冊
■ 卷三
目錄
七言絕

第三層

【卷二的內容】：

■ 目錄

■ 五言古 — 解氏世系歌

■ 四言古 — 題吳山吳子胥廟 乙丑
過曲江曲張九齡祠、...

各自連接該影像檔與電子全文

三、標準套用分析

在標準套用範例中將提出『MARC』、『EAD』、『MARC+EAD』這三種套用方式，分別提出其資料建檔、呈現畫面及工作流程的分析介紹。

1、資料建檔

以下就資料架構分析之後，分別提出「MARC」、「EAD」、「MARC+EAD」建檔所需欄位以供對照：

(1)「MARC」

245 00	\$a 解學士文集十卷
246 17	\$a 解春雨先生文集十卷
250	\$a 明嘉靖間(1522-1566)重刻本
300	\$a10 冊\$c26 公分\$b 線裝
500	\$a 明天順元年黃諫序,八年蔡朔序,明嘉靖四十一(壬戌)年羅洪先序
500	\$a 群碧樓,百靖齋,嘉靖刻本,鄭杰之印等印記
500	\$a 傳圖全彩
505	\$g 第一冊.\$t 印記;羅洪先序;黃諫序;蔡朔序;卷一目錄;卷一應制集.-- \$g 第二冊.\$t 卷二,目錄;四言古--解氏世系歌;五言古--墨菊 杏林精舍圖;
650	\$a 善本--集部--別集類
650	\$a 明刊本
852	\$h 檜木櫃\$i78-1\$8180894-180903

(2)「EAD」

第一層級		
欄位名稱	著錄資訊	EAD element
架號	78-1	<unitid>
題名	解學士文集十卷	<titleproper>
版心題名	解春雨先生文集十卷	<subtitle>
時期	明嘉靖間 (1522-1566)	<date>
版本	重刻本	<note>
裝訂	線裝	<physfacet>
範圍	10 冊	<extent>
尺寸	26 公分	<dimensions>
收藏印記	群碧樓,百靖齋,嘉靖刻本,鄭杰之印 等印記 註:須連結至「印記資料庫」	<separatedmaterial>
儲藏容器	檜木櫃	<container>
主題	善本—集部—別集類;明刊本	<subject>
傳記	解春雨先生	<biohist>
第二層級		
欄位名稱	著錄資訊	EAD element

第一冊	登錄號	180893	<unitid>
第三冊	登錄號	解學士文集十卷第一冊	<unitid>
第一冊	題名	解學士文集十卷第一冊	<unitid>
第三冊	登錄號	180895	<unitid>
	題名	解學士文集十卷第二冊	<unitid>
【 第四冊到第十冊 】			

第三層級 (以第二冊為例)

欄位名稱		著錄資訊	EAD element
目錄	唯一號碼	01	<unitid>
	題名	目錄	<unititle>
	唯一號碼	02	<unitid>
	題名	四言古	<unititle>
	唯一號碼	03	<unitid>
	題名	五言古	<unititle>
【 五絕至七言古 】			

第四層級 (以第二冊五言古為例)

欄位名稱	著錄資訊	EAD element
唯一號碼	01	<unitid>
題名	題吳山吳子胥廟	<unititle>
時間(天干地支)	乙丑	<unitdate>
影像檔	傳圖全彩 註：連結至「全文影像檔」	<daodes> <daoloc>

(3) 「MARC+EAD」

245 00 \$a 解學士文集十卷
246 17 \$a 解春雨先生文集十卷
250 \$a 明嘉靖間(1522-1566)重刻本
300 \$a10 冊\$c26 公分\$b 線裝
500 \$a 明天順元年黃諫序,八年蔡朔序,明嘉靖四十一(壬戌)年羅洪先序
500 \$a 群碧樓,百靖齋,嘉靖刻本,鄭杰之印等印記
500 \$a 傳圖全彩
650 \$a 善本--集部--別集類
650 \$a 明刊本
852 \$h 檜木櫃\$i78-1\$8180894-180903
856 42 \$3Finding aid\$uhttp://lcweb2.loc.gov/ammem/ead/jackson.sgm \$uhttp://jefferson.village.virginia.edu/pmc/contents.all.html

* 此為原 MARC 編目資料，但需加註 856 段，以連結至 EAD finding aid。原來的 505 段保留與否，則視館內政策決定，這部分可以經由 EAD finding aid 的撰寫呈

現。

2、呈現畫面

在呈現方面則是由各單位設計呈現方式，再經由語法設計及系統作業產生，若不論其美編設計，建檔之後的呈現方式大抵如下：

(1)「MARC」

題名：解學士文集十卷
版心題名：解春雨先生文集十卷
版本：明嘉靖間(1522-1566)重刻本
數量高廣：10 冊；26 公分
裝訂：線裝
附註：天順元年黃諫序,八年蔡朔序,明嘉靖四十一(壬戌)年羅洪先序； 群碧樓,百靖齋,嘉靖刻本,鄭杰之印等印記； 傳圖全彩
排架號：檜木櫃 78-1
登錄號：180894-180903
主題：善本--集部--別集類； 明刊本
入藏：第一冊：印記；羅洪先序；黃諫序；蔡朔序；卷一目錄；卷一應制集 第二冊：卷二，目錄；四言古--解氏世系歌；五言古--墨菊、杏林；

* 在呈現部分，若有「作者姓名」可連結至人名權威檔；

「主題」可各自連結至相關主題的查詢結果畫面；

「入藏」可各自連結至該影像檔或電子全文檔。

(2)「EAD」

A、電子書名頁
解學士文集十卷
解春雨先生文集十卷 明 嘉靖間 (1522-1566) 重刻本 檜木櫃 78-1 中央研究院 歷史語言研究所 傅斯年圖書館
就像書的前言一般，說明此批收藏的一般性資料，可包含以下幾點：
1、 解春雨先生生平概述：如 生卒、學歷、經歷 等。
2、 館藏資料類型與內涵：如 詩、傳、記、墓誌銘、祭文 等。
3、 館藏分類方式：如 共可分為十冊 等。
註記館藏範圍：
範圍：如 收藏成幾捲、放置於何種儲藏設備中、所佔空間多少 等。
提供連結，使用者有欲進一步了解時，可自行點選瀏覽：
1、 資料收藏資訊 用以提供資料何時入館等相關資訊。
2、 館藏位置 用以提供資料在館內放置在何處的相關資訊，如：樓、架，或登錄號起訖 等。
3、 傳記/歷史 如有關於作者生平更詳細的資料，可經由此連結提供更詳細的資訊。

4、[範圍與內容](#) 說明各冊下包含內容有哪些、代表的意義為何 等。

5、[館藏目錄清單](#) 用以提供詳盡的館藏目錄清單。

B、詳細內容(以下為虛擬清單)

解學士文集十卷

登錄號 - 項目 - 詩文

180894

第一冊

A：本文前

180894 01 羅洪先序

180894 02 黃諫序

180894 03 蔡朔序

B：卷一

180894 04 目錄

180894 05 應制集

180895

第二冊

180895 01 目錄

180895 02 四言古

180895 02 01 解氏世系歌

180895 03 五言古

180895 03 01 題吳山吳子胥廟，乙丑

180895 03 02 過曲江曲張九齡祠

180895 03 03 題胡忠簡公敕書

180895 03 04 青陽公祠

180895 03 05 題田景畫，四首

180895 03 06 墨菊，丙寅

180895 03 07 杏林精舍圖，為環州李允誠賦

180895 03 08 為康中正書

180895 03 09 送劉日昇南游庾嶺

180895 03 10 寄蕭公遜

180895 03 11 西行寄同鄉諸友

180895 03 11 南行寄友

180895 04 五絕

180895 05 五言律

180895 06 五言排律

180895 07 七言古

180896

第三冊

180896 01 目錄

【 至第十冊】

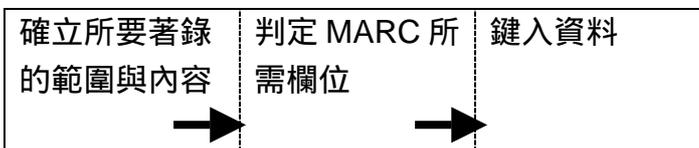
(3) 「MARC+EAD」

<p>題名：解學士文集十卷 版心題名：解春雨先生文集十卷 版本：明嘉靖間(1522-1566)重刻本 數量高廣：10 冊；26 公分 裝訂：線裝 附註：天順元年黃諫序,八年蔡朔序,明嘉靖四十一(壬戌)年羅洪先序； 群碧樓,百靖齋,嘉靖刻本,鄭杰之印等印記； 傳圖全彩 排架號：檜木櫃 78-1 登錄號：180894-180903 主題：善本--集部--別集類； 明刊本 詳細資訊 EAD SGML 格式</p>
--

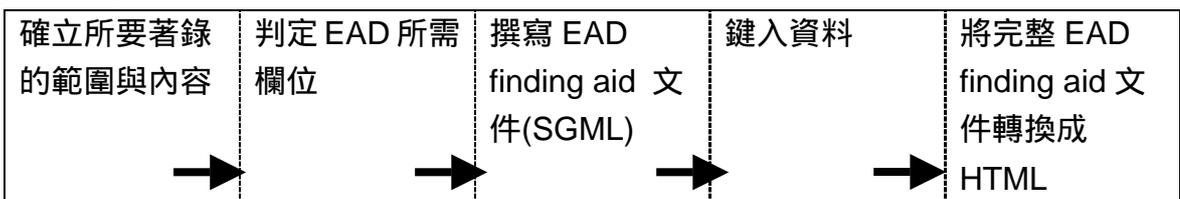
- * 在呈現部分，若有「作者姓名」可連結至人名權威檔；
- 「主題」可各自連結至相關主題的查詢結果畫面；
- 「詳細資訊」為 856 段，可連結至設計好的 html；
- 「EAD SGML 格式」為 856 段，可連結至 sgml。

3、工作流程

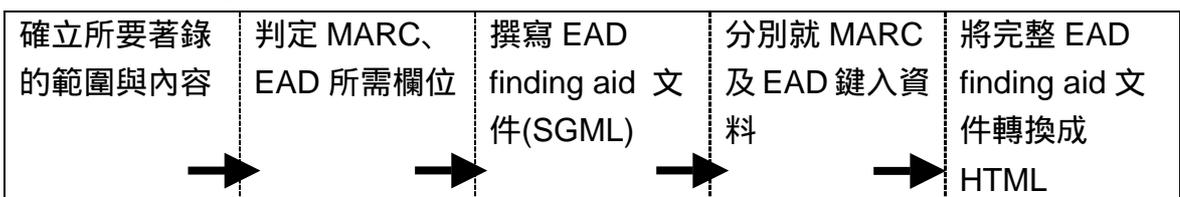
(1) 「MARC」



(2) 「EAD」



(3) 「MARC+EAD」



四、標準適用性分析

分析比較	MARC	EAD	MARC+EAD
------	------	-----	----------

<p style="text-align: center;">優點</p>	<p>1、為傳圖原本採用之標準，著錄人員對此著錄格式及規範會較習慣順手。</p> <p>2、著錄方式較為簡單明瞭。</p>	<p>1、針對描述資料可做較詳細且多元的著錄，如作者傳記，可以短文的方式作一詳細的介紹，不需再連結至人名權威檔，而且不受長度限制。</p> <p>2 由於 SGML 本身具有電子出版的特性，可藉由 EAD finding aid 的撰寫呈現具有階層性的館藏全貌。</p> <p>3、在交換機制中不受系統平台的限制。</p>	<p>除了有著採用 EAD 的優點之外，另有：</p> <p>1、在呈現部分可先呈現 MARC 提供摘要性的資料，再進一步連結至 EAD 部分，能清楚顯示其層次又不至於太過雜亂。</p> <p>2、可以讓使用者透過圖書館目錄檢索。</p>
<p style="text-align: center;">缺點</p>	<p>1、標準架構無層級，內容僅能著錄於 505 段，在呈現方面無法展現其層級關係，僅能以符號區分。</p> <p>2、許多善本資料較特殊的欄位如序跋、印記等僅能著錄於附註項，無法突顯其特性。</p> <p>3、對於資料的描述有長度的限制。</p>	<p>1、著錄人員須重新研讀 EAD 標準，耗費人力及時間。</p> <p>2、需要有熟悉且有能力的撰寫完整 SGML 文件的人手。</p> <p>3 在製作 SGML 部分需要有軟體的輔助，或許會需要額外的支出。</p> <p>4、著錄的資訊可能增加，著錄人員必須花費時間找出除資料本身的資訊，如作者傳記、資料的歷史以及其他的考證及參考資料。</p>	<p>1、著錄人員須重新研讀 EAD 標準，耗費人力及時間。</p> <p>2、需要有熟悉且有能力的撰寫完整 SGML 文件的人手。</p> <p>3 在製作 SGML 部分需要有軟體的輔助，或許會需要額外的支出。</p> <p>4、著錄的資訊可能增加，著錄人員必須花費時間找出其他資訊，而且必須同時著錄兩種標準。</p>

五、分析與建議

數位典藏之主要目的，除了保護珍貴資料避免再度受到人為破壞之外，此外，展現資料之全貌，促進資料之研究價值，提高資料之使用量，也是其重要目標之一。因此依據上述適用性

之比較評估，最適用的為「MARC+EAD」，因為如此不但可以較簡潔的方式呈現，也可使研究人員窺其全貌，未來在國際交換機制中，也可以輕易的與他人做交換，此外，MARC亦是傅斯年圖書館原本採用的著錄方式；但如此一來必定耗費更多人力及時間，故若傅圖希望採用此方法，可先行評估其人力及時間，在兩者皆許可的條件下，本工作組建議採用此種著錄格式。

而依據主題計畫目前的現狀，若確定採用此種方法，單純延續目前的工作進度看來，主要的工作項目可為：

- 1、完成目前 metadata 的建置；
- 2、檢視原有資料庫與目前 metadata 欄位的對應結果；
- 3、建置資料庫並轉入原有資料；
- 4、依 metadata 欄位輸入及更新資料；
- 5、檢視目前 metadata 與 EAD 和 MARC 的對應結果(mapping)；
- 6、撰寫 SGML 文件(除了原來 metadata 欄位外,尚需包含其他相關資料以及排版的部分)；
- 7、轉出原建置於資料庫的資料，並將 SGML 文件轉換成 HTML。

在這些工作項目中，還需評估一些必要的資源需求，如：

- 1、人力：需評估是否應增加工作人員；此外，工作人員(無論編目或是系統人員)需大致了解 EAD 以及 SGML 的應用。
- 2、設備：需評估原來的軟硬體是否符合需求；若不符，是否添購軟硬體？要購置哪些？
- 3、經費：需評估總預算以及可能增加的人事費、設備費、材料費等。

此外，因選用 EAD 需付出一定的人力與時間，加上傅圖的善本資料為數龐大，因此建議評估過後若確定採用 EAD，可採**循序漸進**的方式進行，先選定幾項資料進行，以下提供幾個要點作為參考：

- 1、檢視目前最完整的資料為何(如同一作者的著作)；
- 2、檢視目前最常被使用的資料為何；
- 3、檢視最重要的館藏為何；
- 4、檢視最有展示價值的館藏為何；

在選定預先做成 EAD 的資料之後，即可開始進行相關工作，不過此部分的工作不只是需要傅斯年圖書館相關工作人員的努力，亦需要 MAAT 以及相關系統人員共同的研析。

而在評估是否採用 EAD 時，本工作組提供 EAD Application Guidelines Appendix D：Implementation Checklist 之中譯作為參考，見附件。

EAD Application Guidelines for Version 1.0

附錄 D：應用核對清單(Implementation Checklist)

數位典藏國家型科技計畫後設資料工作組 譯

多數存藏所在應用 EAD 時，都需面對多種階段過程：

- 轉換舊的查檢工具
- 產生新的查檢工具
- 在網路上公佈查檢工具

在第二章、第四章、第五章詳細的說明每一階段的挑戰及可利用方法。以下的確認清單可協助存藏所運用 EAD 功能為架構時應思考的問題。

1. 評估查檢工具在目前檢索或參考服務環境下所扮演的角色。

a. 目前查檢工具如何被使用？

- 使用群為何？
- 被使用的為哪些範圍的項目(circumstances)？
- 這些被使用的項目，那些反映出最高層次的使用？
- 在查檢工具中，哪些資料類型最常被搜尋？
- 哪些疑問能使用查檢工具有效地解決？哪些不能？
- 線上的查檢工具能維持目前有效的檢索效能，並可能應付那些查檢工具尚未能有效滿足的部分？
- 線上查檢工具，能否為存藏所產生新的資料使用群？

b. 目前查檢工具的狀態？

- 目前查檢工具的實體格式(physical format)為何？
- 查檢工具的完整程度？對資訊的正確度有多少把握？
- 查檢工具的結構成分及其所包含的資料的一致性程度？這些成分標示的清楚度？
- 建構查檢工具所遵循的指引？
- 希望有多少查檢工具可直接或最後可轉換為 EAD？呈現多少頁的文字資料？
- 目前新產生的查檢工具所佔比例？

- c. 目前存藏所是否有產生 MARC 格式記錄，若有，則這些記錄格式和你的查檢工具的關係？
2. 如何處理轉換現行的查檢工具？
 - a. 現行查檢工具轉換的優先順序
 - 最重要的館藏
 - 最常使用的館藏(與最少使用的館藏輪流交替)
 - 最容易轉換的查檢工具(需要最少的處理)
 - 若為網路上利用，最能有效被利用的查檢工具
 - 分享多個存藏所的館藏，產生虛擬的查檢工具
 - 網路上可利用的電子資料的館藏集合
 - b. 採用的轉換方式
 - 機構內部轉換
 - 委托代理商
 - 參與提供轉換服務的合作計畫
3. 希望提供什麼方法讓使用者搜尋查檢工具？
 - a. 連結網站上的線上目錄
 - b. 藉由網站爬蟲(Web crawler)如 Alta Vista 或 Yahoo 搜尋網際網路
 - c. 直接進入機構網站瀏覽查檢工具
 - d. 在自身網站上使用搜尋引擎(search engine)
4. 產生 EAD 編碼的查檢工具及在網路上公佈時，所需的資源。
 - a. 人力需求配置
 - b. 需求人員的教育訓練
 - c. 所需的專業技術為何？若無法從存藏所內獲得，是否可從館內其他單位獲得？或能否加入已應用 SGML/XML 的聯合團體？是否能加入已發展的體系或分享資源和專業技術？
 - d. 需要那些專業文件及需要多少分？
 - EAD DTD files 或特定應用軟體的 DTD 版本(如.rls file for Author/Editor, .lgc file for WordPerfect)
 - EAD Tag Library
 - EAD Application Guidelines
 - 採聯盟方式進行時，所需的編碼指引(Encoding guideline)
 - e. 需要展的個別環境(Local convention)
 - 查檢工具所遵循的標準格式(standard format)
 - 著錄至各元素的標準規範
 - 控制展現查檢工具的樣式表(stylesheet)
 - 非標準權威來源所包含的檢索詞彙的權威形式

f. 產生或公佈新的查檢工具所需的軟體(並非每一查檢工具需完全採用以下工具)

- SGML/XML authoring package
- 具編輯 SGML/XML 功能或具有轉換附加程式的文書處理軟體
- Database
- 轉換 SGML 為 HTML 轉換軟體或 HTML 編輯工具(authoring tool)
- 轉換工具如 perl scripts、macros
- SGML/XML parser
- SGML/XML browser
- 樣式表編輯軟體(Stylesheet authoring software)
- 搜尋引擎(Search engine)

g. 產生或公佈新的查檢工具所需的硬體

- 電腦工作站(Computer workstation)
- 內部網路聯結(Local network connection)
- 網際網路聯結(Internet connection)
- 備份裝置(backup mechanism)
- 伺服器(Server)
- 印表機(Printer)

h. 如何掌握品質控制？

i. 如何維持及更新查檢工具？

j. 如何管理伺服器的運作及維修？

5. 分析每一過程需要的成本。

a. 那些費用是目前預算中有列項的？

b. 那些為新的支出費用？是一時的或會持續？

c. 那些費用是會隨時間而逐漸減少或增加？

d. 那些費用可從外部基金獲得？如補助金。

e. 是否有潛藏需進一步細查的費用？

f. 是否會因完成 EAD 應用，而節省某些部分費用？