



後設資料簡介與建置流程

講師：沈漢聰

中華民國一百年九月十四日

數位典藏與數位學習國家型科技計畫 後設資料工作組研製

大綱

- 一. 後設資料的簡介
- 二. 後設資料建置流程
 - 描述單元（藏品粒度）的決定
 - 規劃藏品之唯一識別碼
 - 決定描述主體
 - 後設資料標準選用
 - 訂定欄位著錄規範
- 三. 後設資料對照服務介紹

一、後設資料簡介

後設資料簡介

『 Metadata 』 一詞之各式中譯：

- 元資料
- 元數據
- 超資料
- 詮釋資料
- 後設資料

後設資料簡介

- 後設資料(Metadata)的定義：
- 資料中的資料 (Data about data) 。
- 機器可解讀的網頁資訊。(from W3C)
- 有關資料的結構性資料。(from Dublin Core)
- 後設資料是有關一個數位典藏品的資料，通常由典藏品創作者或提供者來建立。一般而言，後設資料會伴隨著數位典藏品出現或埋置於檔案標頭中，因此，後設資料可以作為資訊儲存與檢索系統很有用的基礎。(from 美國圖書館學會)
- …… 超過30種的定義。

後設資料簡介

什麼是後設資料(Metadata)?



中文品名：豹皮衣

分類編號：10162

關係族社：排灣族

尺寸：

衣長67.0cm，出手29.0cm，
出手高39.0cm，胸寬51.5cm，
衣襠57.0cm

後設資料簡介

什麼是後設資料(Metadata)?



品名：玉辟邪

原品名：漢玉獸

尺寸：長13.2cm、高5.7cm

保存狀況：良好

典藏位置：玉器庫房



後設資料簡介

什麼是後設資料(Metadata)?



品名：玉辟邪
時代：漢代(206B. C. - A. D. 220)
質材：玉石
色澤：玉色白裡透黑
文物評等：古物



來源：故宮器物數位典藏子計畫

後設資料簡介

清 錦地團花蒔繪三層小屨盒



後設資料

授權與引用資訊

連結到原始資料網站

資料識別：

故玉004103D000001000

資料類型：

木竹漆器

描述：

二屨及最上層格內置各式小玉器，盒外黑漆地上均以平蒔繪菱格形梅花紋為地紋，其上高蒔繪折枝花卉團花，團花內地文或平目地或梨地，適合多種技法，十分華麗。中層屨平蒔繪的技法和其他地方明顯不同，可能是清宮修補。

日期：

1644 A.D.-1911 A.D.

格式：

長8.3公分 寬7.8公分 高17.7公分

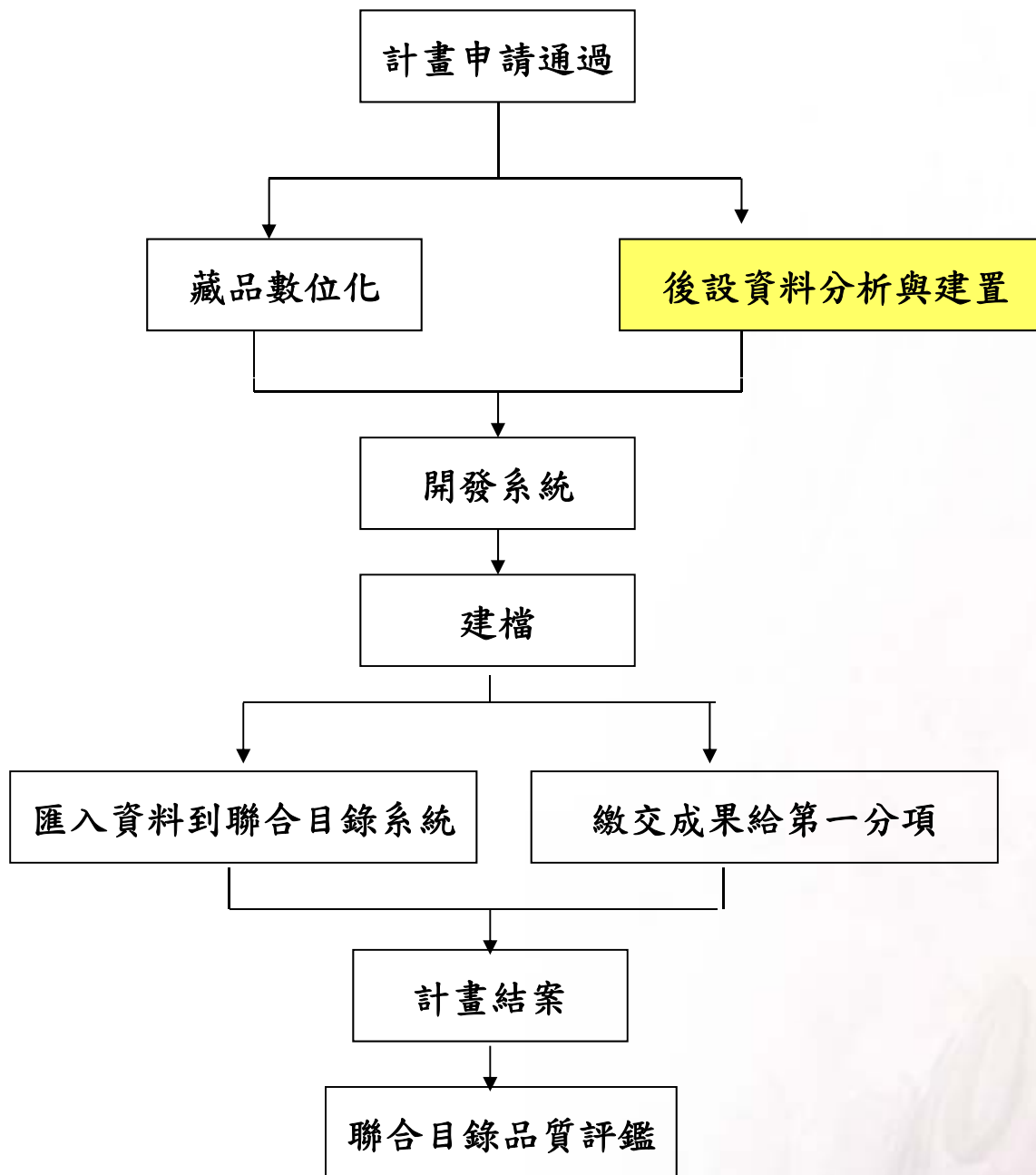
關聯：

《清宮蒔繪—院藏日本漆器特展》

管理權：

國立故宮博物院

二、後設資料建置流程



數位典藏計畫執行簡單示意圖

計畫通過國科會審核，開始執行....

清查計畫要數位化的藏品，整理成清冊。

進行藏品分類。

如：文物、照片、手稿、檔案、報導、影帶、...等。

決定每種類型的描述單元（藏品粒度）。

如：一本？一章？一篇？（期刊、書籍...）

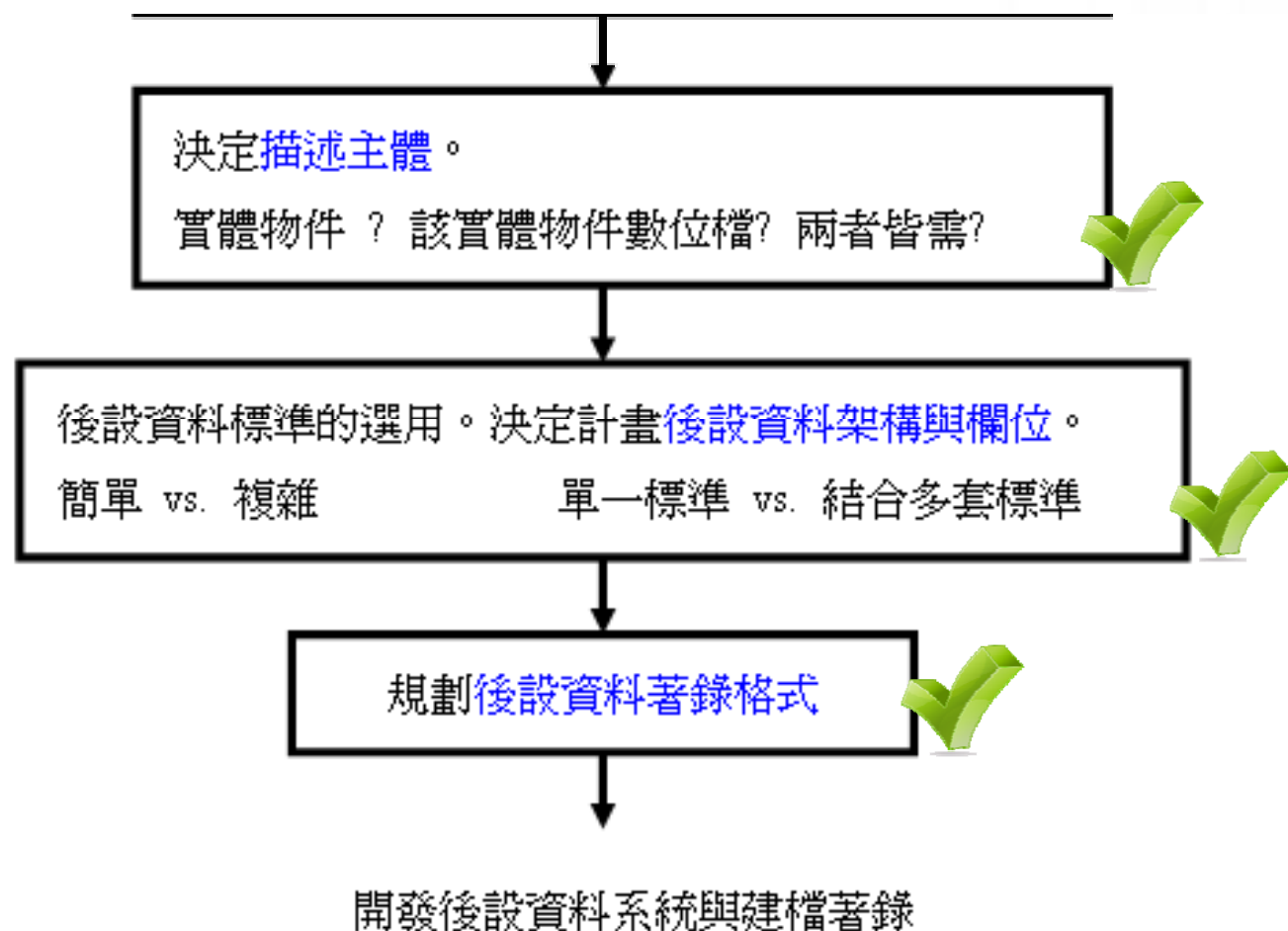
如：一組？單件？（文物、器物...）

如：整卷？分段？（影帶、錄音...）



依據藏品粒度大小的決定來設計唯一識別碼的編碼原則。實體物件與數位物件應分開編碼。





後設資料建置流程1—釐清藏品類型

- 清點即將進行數位化的實體典藏品，歸納出典藏品之類型清單。
- 釐清各典藏品類型單元與單元間可能存在的組、附件關係。
- 規劃整體的需求面向，與預期未來的應用。

藏品類型範例-文字類



圖畫



雜誌

藏品類型範例-文字類



海軍艦艇航泊日記資料頁

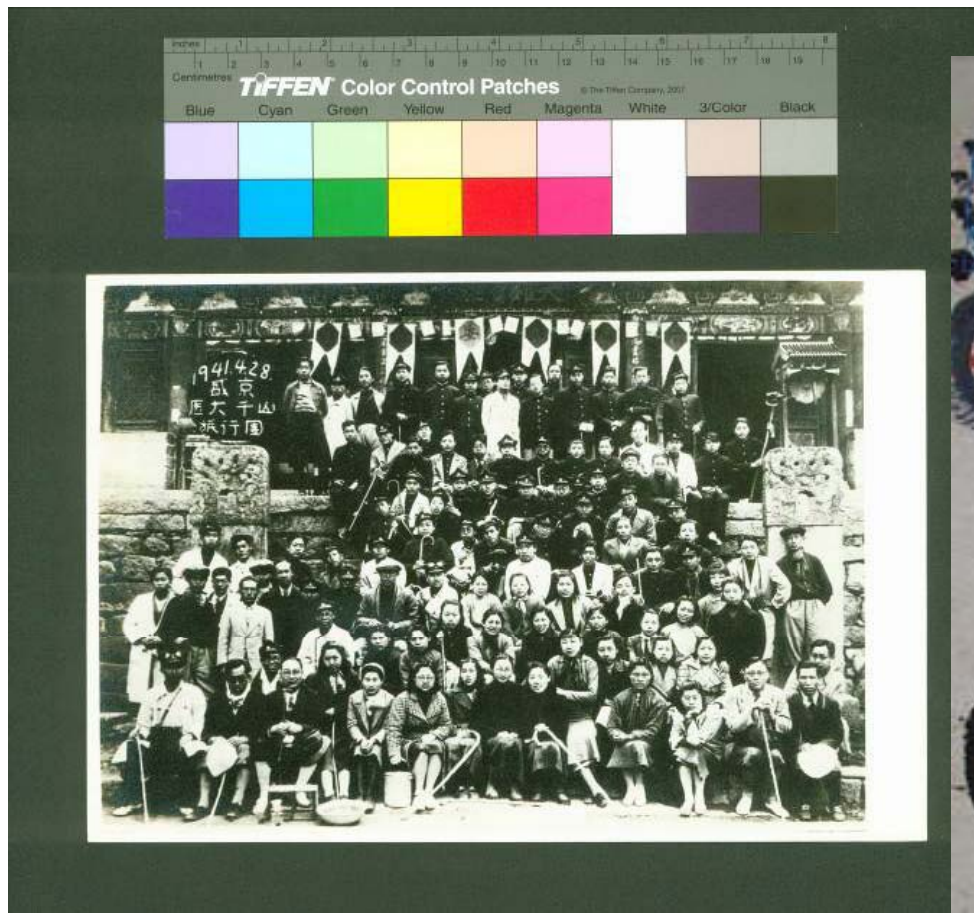
中華民國八十二年五月廿一日

日期	艦名	航泊地點	航程	航速	航向	航時	航程	航速	航向	航時	航程	航速	航向	航時
0800	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來
1400	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來
2000	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來	自來

檔案

報紙

藏品類型範例-靜態圖像類



照片



圖畫

藏品類型範例-聲音類



如：黑膠唱片、
錄音帶、CD、
數位錄音檔

藏品類型範例-動態影像類



如：錄影帶、
DV帶、LD、VCD、
DVD、數位錄影
檔

藏品類型範例-實體物件類



建築



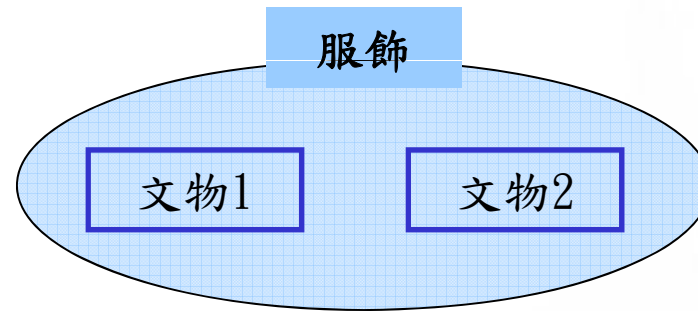
文物



標本

藏品描述單元與群組關係分析

範例1：



套件：每件藏品獨立存在是有意義的，與其他藏品合置也具意義，但只需單獨著錄每件藏品。

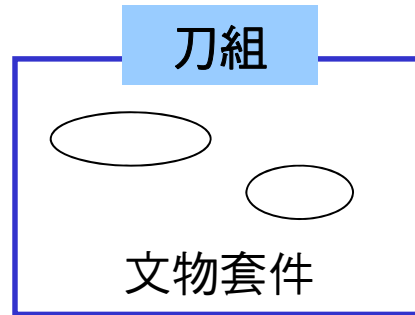
元素名稱		範例1	範例2
物件名稱		泰雅族女子上衣	泰雅族女子腰裙
分類編號		10020	10021
相關 藏品	套件 描述	屬於泰雅族女子服飾，另有同套腰裙 <u>10021</u> 。	屬於泰雅族女子服飾，另有同套上衣 <u>10020</u> 。



來源：
<http://www.hles.tpc.edu.tw/nine/sample/document/art/13>

藏品描述單元與群組關係分析

範例2：



組件：兩件或兩件以上的藏品合置在一起較具意義，故以整組著錄即可。

欄位名稱		範例
物件名稱		刀(有鞘)
分類編號		50048
相關藏品	組件描述	刀與其附屬刀鞘，兩者為一完整之刀組。



來源：中研院民族所台灣原住民文物計畫

來源：中研院民族所台灣原住民文物計畫

藏品描述單元與群組關係分析

國外範例：

欄位名稱	範例- Polyptych with the Madonna and Child, Saint James Major, and Various Saints	
編目層級		
數量	1	
類型	polyptych	Altarpiece
	panel painting	
組件	10 panels	
	數量	10
	類型	Panels
註釋		
引文		

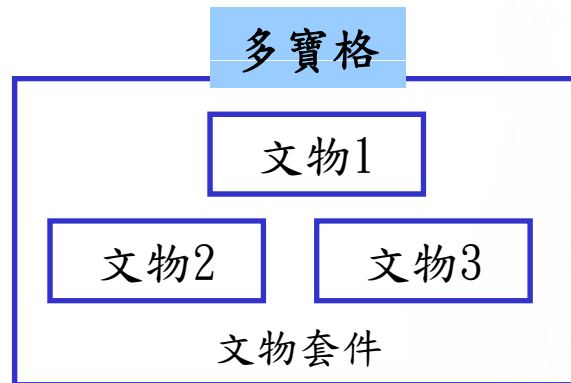


From: CDWA

http://www.getty.edu/research/conducting_research/standards/cdwa/3_cataloging_examples/index.html



藏品描述單元與群組關係分析

範例3：



套組件：每件藏品獨立存在是有意義的，與其他藏品合置也深具意義，故除了單筆著錄每件藏品之外，也需再另外著錄一筆整組的資訊。

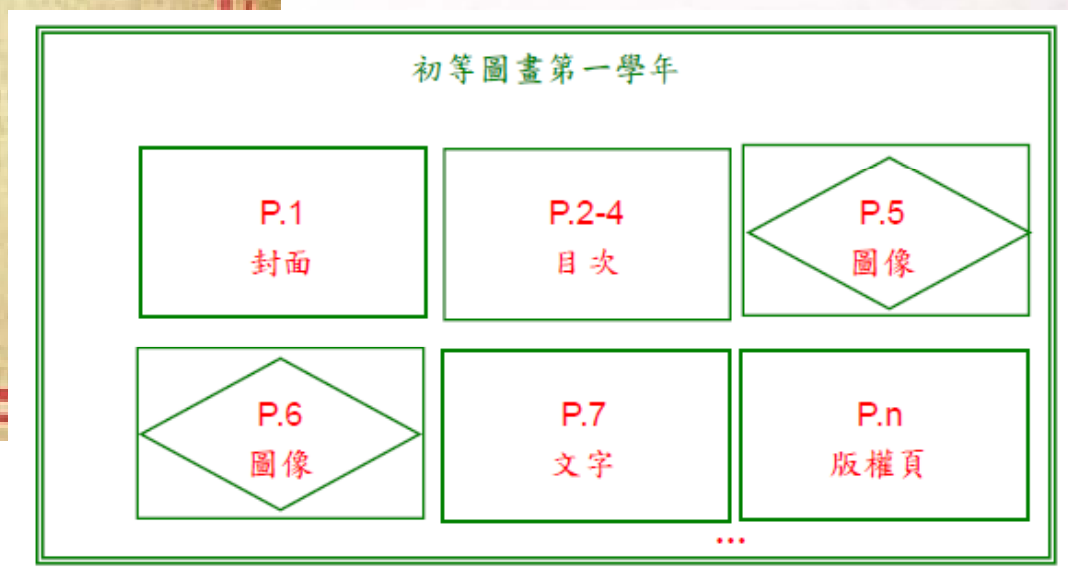
藏品描述單元與群組關係分析

欄位名稱		例1--清 雕 紫檀龍紋多 寶格 故雜 001374N0000 00000		例2- 明 青玉四耳活尊 故玉 005759N000000000	
編 目 層 級	類別	單件	附木座		
	說明	附各式珍玩四十七件，分別著錄	存放於雕紫檀龍紋多寶格小櫃（故雜001374），分別著錄。		
	數量	1	2		
	相關 藏品	故畫003848N000000000 故玉005759N000000000 故玉005760N000000000 故玉005761N000000000 故玉005762N000000000 故瓷018217N000000000	故雜001374N000000000		

藏品描述單元確認



一本？一章？



藏品描述單元確認



一本？一篇？

台灣人權促進會會訊創刊號

封面

版權頁
(目次
頁)

篇 1

「掃黑
與人權」
座談會

篇 N

...

藏品描述單元確認

海軍艦艇航泊日記資料頁									
日期	艦名	航次	航向	航速	航程	航時	航高	航寬	航深
15.1.23	150	28	23	10/2	01	3			
15.1.24	150	28	24	10/2	01	3			
15.1.25	150	28	25	10/2	01	3			
15.1.26	150	28	26	10/2	01	3			
15.1.27	150	28	27	10/2	01	3			
15.1.28	150	28	28	10/2	01	3			
15.1.29	150	28	29	10/2	01	3			

0800	白雲潭海軍艦艇航泊日記	28.23	1945
1400	同上		
2000	同上		

檔案層級：全宗(副全宗)、系列(副系列)、案卷、件

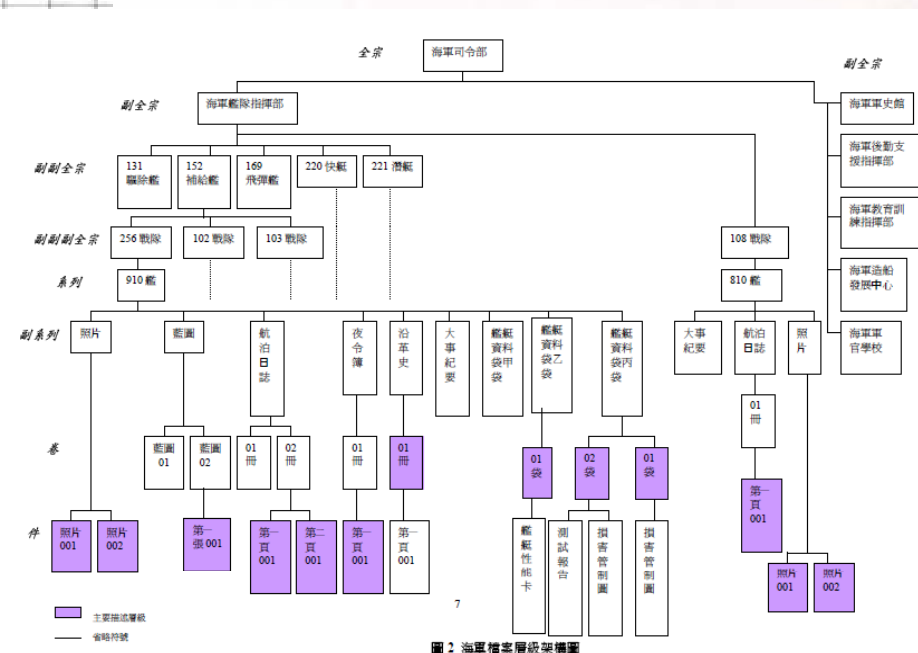


圖 2 海軍檔案層級架構圖

藏品描述單元確認



一張唱片?正反面?一首歌曲?

藏品描述單元確認



一捲錄影帶?一片DVD?一段影片?

藏品識別碼設計與編碼

進行識別碼編碼前須完成下列工作：

- ✓ 確認藏品類型與藏品描述單元
- ✓ 清點各藏品類型與藏品單元數量
- ✓ 依據藏品類型、單元之數量進行識別碼架構設計

藏品識別碼設計與編碼

識別碼設計通用原則：

- 所有使用的識別碼應具有其意義。
- 識別碼應該是簡短且易於辨識其類別的。
- 盡量只使用英文字母與數字進行編碼，在同一編碼類型中如有空格需求時，使用”_”進行區隔，在不同編碼類型間如有間隔需求則使用”-”進行區隔。
- 識別碼內容應該儘可能的使用縮寫。
- 所有資源使用的位數應該統一，如未使用之位數則補0(例如編號第1、11、111的資源，在識別碼呈現上應為001、011、111)。

藏品識別碼設計與編碼

- 藏品編碼類型：

- 非智慧型編碼

- 僅設定編碼位數，主要以流水號的方式依序編號，編碼不含任何藏品相關資訊。

- 智慧型編碼

- 依循一套編碼的規則與格式，通常把常用藏品資訊轉化成藏品編碼格式之一部份，讓藏品管理員快速分辨藏品類型。

藏品識別碼設計與編碼

- 藏品編碼範例－非智慧型編碼：
台灣外省人生命記憶與敘事資料庫(II)－揭開「白色封印」數位典藏計畫，以「識別碼」作為典藏系統之主鍵元素，主鍵元素具有識別紀錄之功能。「識別碼」可由系統自動產生且在建檔時需檢查此欄位是否有重複建檔。「識別碼」之組成格式為：6碼流水號
- 例如：000001、000065

藏品識別碼設計與編碼

- 藏品編碼範例 - 智慧型編碼：

- 故宮器物「文物統一編號」，共有兩種格式：

- 格式一：

- 兩個中文字 + 六個阿拉伯數字 + 一個英文字元 + 九個阿拉伯數字

- 範例：故琺000975N000000000

- 格式二：

- 三個中文字 + 四個阿拉伯數字 + 一個英文字元 + 九個阿拉伯數字

- 範例：中日陶0569N000000000

藏品識別碼設計與編碼

範例一、書籍與期刊(日治時代台灣小公學校美術教課書數位典藏收集計畫)

識別碼架構：A00-00-000

說明：

1. 確認藏品類型與單元後共有美術教科書、美術教育專題、期刊美術相關三類藏品。
2. 清點各藏品類型與藏品單元數量後，每類藏品不超過100本、每本不超過100篇、每篇不超過1000頁。
3. 編碼設計分為「本－篇－頁」三區。
4. 依照清點數量給定編碼長度。
5. 以“本”區第一碼為藏品分類，A代表「美術教科書」類；B代表「美術教育專題」類；C代表「期刊美術相關」類。

藏品識別碼設計與編碼



識別碼：A01

說明：整本教科書的
識別碼

識別碼：A01-00-001

說明：該本教科書的
第1頁

藏品識別碼設計與編碼



識別碼：B01-01

說明：該篇專題的識別碼

識別碼：B01-01-001

說明：該篇專題的第1頁

範例二、檔案

- 「海軍歷史文物數位典藏－陽字型驅逐艦數位典藏計畫（I）」藏品編碼共22碼，編碼方式如下：

共八區：

海軍一級單位（1碼）+海軍二級單位（2碼）+艦隊（3碼）
+戰隊（3碼）+艦號（4碼）+資料類型碼（2碼）+卷（4碼）+件（3碼）

範例：1-01-146-268-0914-03-0072-005

（洛陽軍艦航泊日誌單日本，第72卷，第5頁）

藏品識別碼設計與編碼

海軍艦艇航泊日記資料頁

日期	艦名	航泊地點	航泊時間	航泊次數
10/1	268	150.1	28.28	1/2
10/1	268	150.4	28.28	1/2
10/1	268	150.6	28.28	1/2
10/1	268	150.4	28.26	1/2
10/1	268	150.5	28.24	1/2
10/1	268	150.7	28.25	1/2

0800 自海軍司令部海軍艦隊司令部海軍艦隊司令部
1400 同上
2000 同上

268	4700	1946
471	4700	1946
81360	4700	1946

識別碼：1-01-146-268-0914-03-0072-005

說明：海軍司令部-海軍艦隊指揮部-146艦隊-268戰隊-0914號艦-航泊日誌-第72冊-第5件

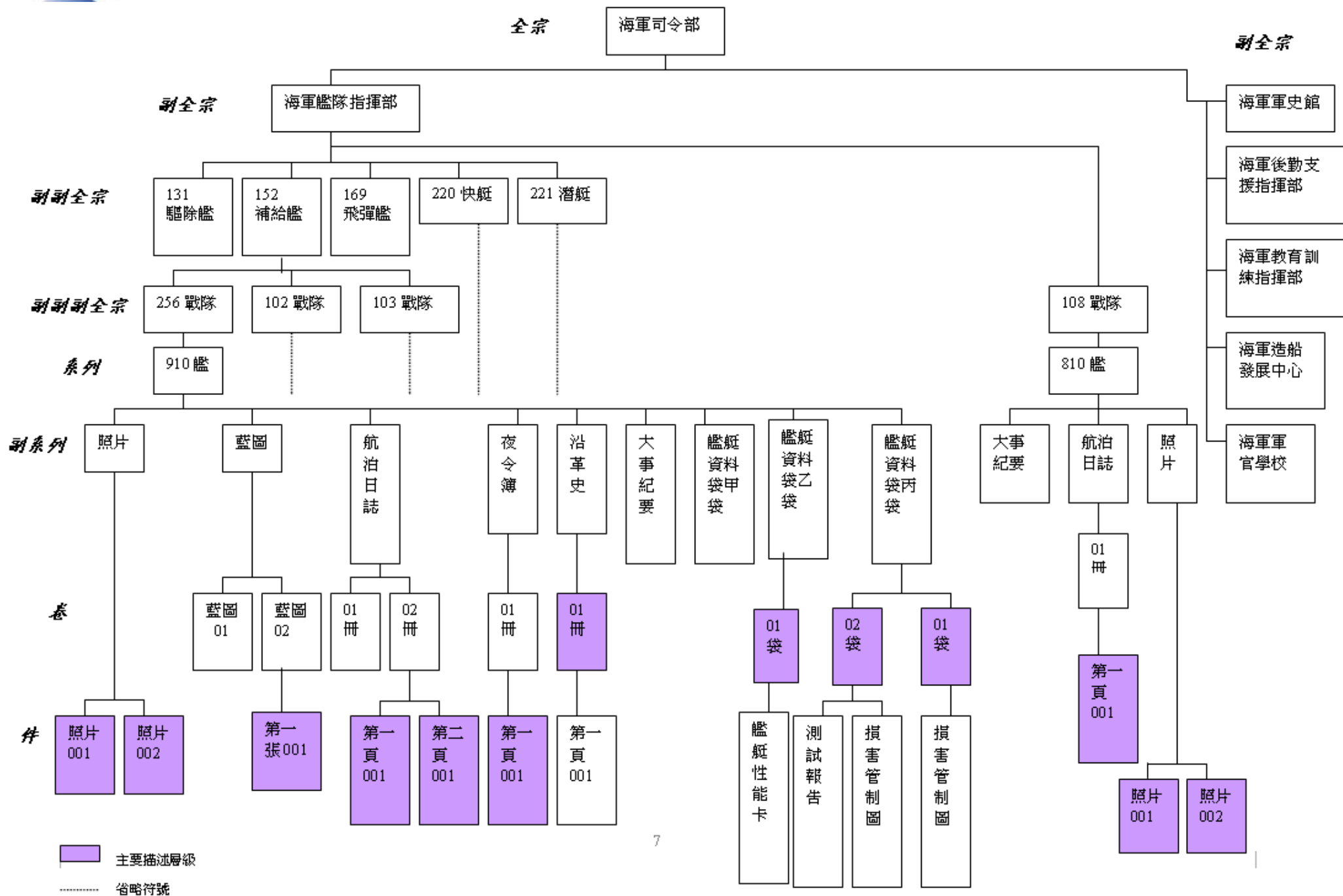
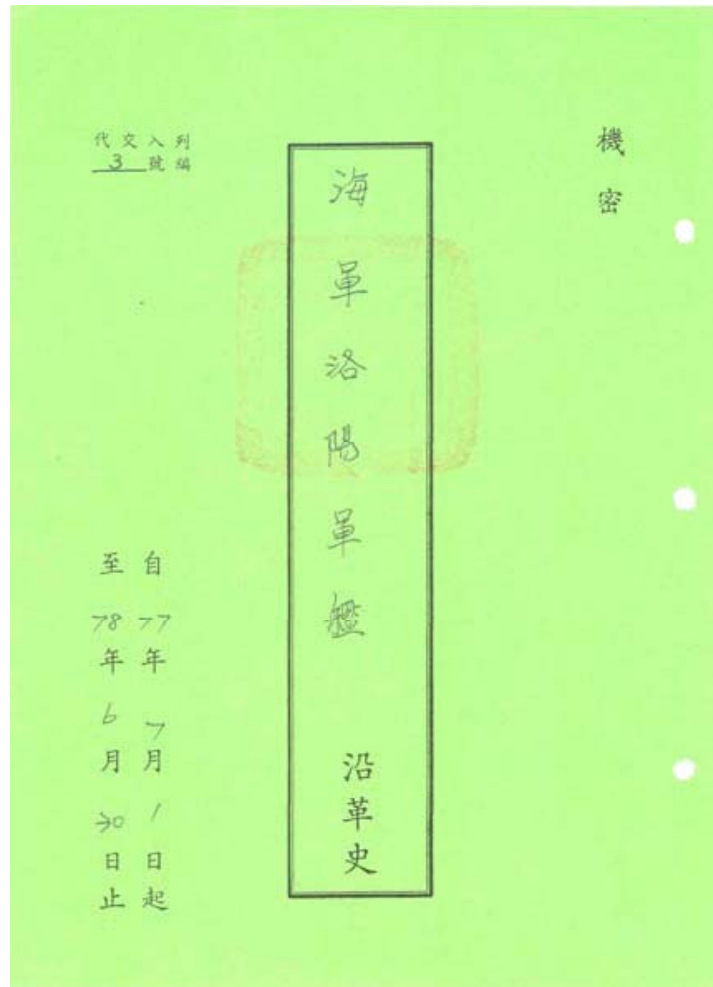


圖 2 海軍檔案層級架構圖

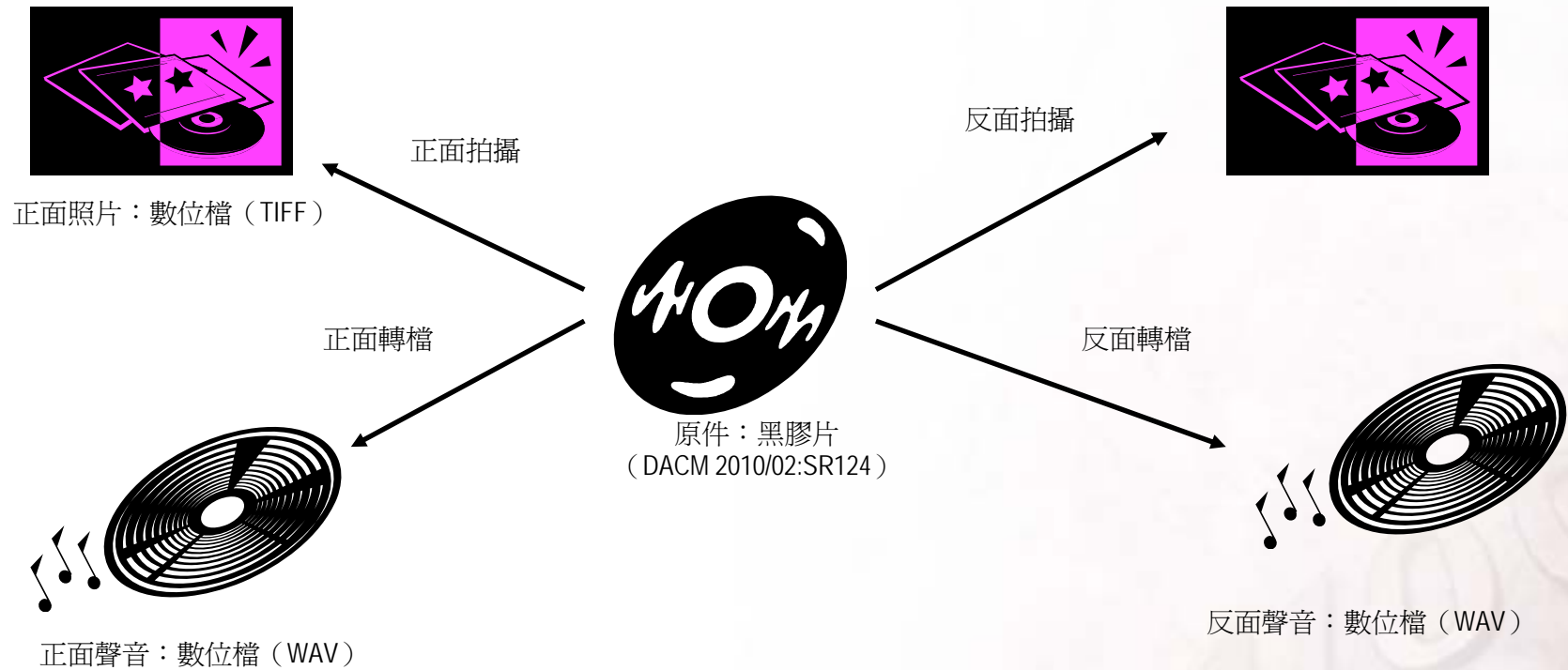
藏品識別碼設計與編碼



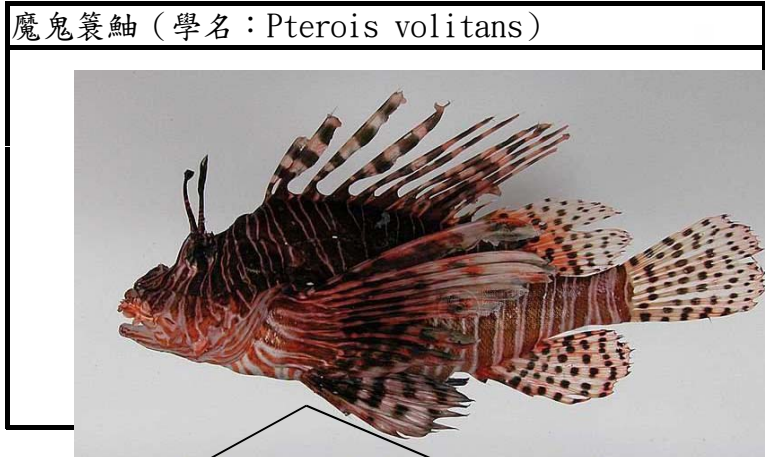
識別碼：1-01-010-011-
0017-05-0002-001

說明：海軍司令部-海軍
艦隊指揮部-010艦隊-
011戰隊-0017號艦-沿革
史-第2冊-第1件

日治時期臺灣曲盤數位典藏計劃



選擇不同的描述主體，呈現不同的後設資料



原件

標本號：FRIP21012
類型：臺灣東部魚類標本
鑑定者：林沛立
標本採集方式：手釣
鑑定時間：2005/07/26
TL(全長)：285mm
重量：314g
原件使用權：不開放

數位檔

照片編號：51593
檔案類型：圖片
數位照相機：RICOH Caplio 500SE
拍攝時間：2005/08/01
檔案大小：1888353 bytes
檔案格式：image/jpeg
數位檔案使用權：可供教學使用

選擇不同的描述主體，呈現不同的後設資料



照片原件

數位化



數位檔

Title : 蘭醫師夫婦
Creator : 不詳
Date : 1954 (拍攝日期)
Format : 6.44*8.85cm (照片原件尺寸)
Source : 無
Rights : 彰化基督教醫院院史館 (文物典藏地點)
Description : 蘭大衛醫生夫婦於英國倫敦合影
Type : 靜態圖像

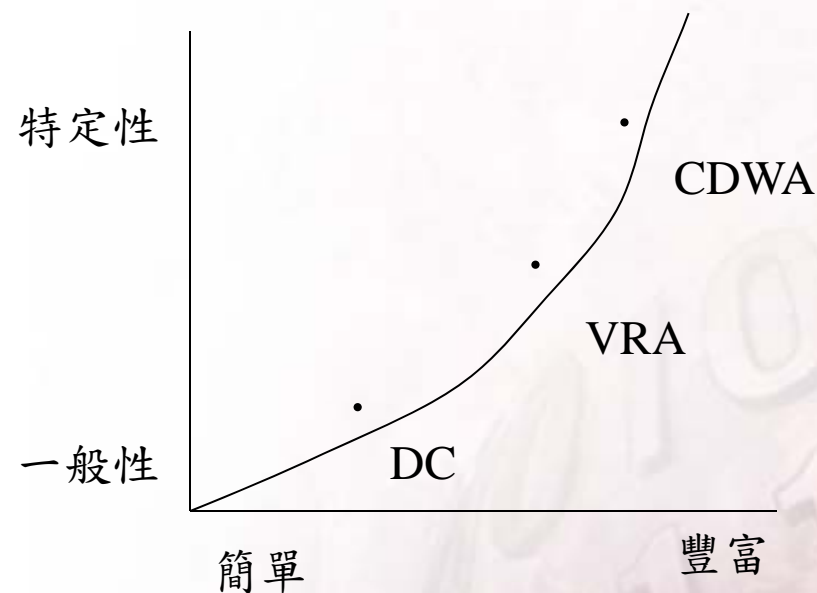
Title : 蘭醫師夫婦 (數位檔)
Creator : 劉耀仁 (掃描者)
Date : 2010-12-06 (掃描日期)
Format : tiff (檔案格式) 350ppi (解析度) 200% (放大倍率)
Source : 蘭醫師夫婦 (原件)
Rights : 彰化基督教醫院院史館 (數位典藏地點)
Description : 蘭大衛醫生夫婦於英國倫敦合影
Type : 靜態圖像

後設資料標準簡介與選用

後設資料標準詳簡類型探討：

不同的後設資料格式支援不同目的的需求。

- 簡單 *vs* 豐富
- 一般性 *vs* 特定性



後設資料標準簡介與選用

標準策略選擇策略

如何選擇適用的標準？

- 需求符合性
- 標準的成熟度
- 標準應用的輔助工具（指引文件、軟體工具……）
- 標準的聲譽（Reputation）與目前採用狀況
- 日後的發展與維護

後設資料與後設資料標準簡介

後設資料的組成結構：

— 資料結構 (*data structure*)

如：*CDWA*這套標準之*metadata*。

— 資料內容 (*data content*)

如：*Cataloguing Cultural Objects*。

— 資料值 (*data value*)

如：*CDWA*中的*Authorities*，包括*Place/Location Identification*、*Subject Identification*、*Creator Identification*等。



後設資料與後設資料標準簡介

編目紀錄 建檔者: 黃XX

建檔時間: 2009/11/23 3:45pm

資料結構 – 後設資料的欄位結構與編排



後設資料與後設資料標準簡介

編目紀錄 建檔者：黃XX

建檔時間：2009/11/23 3:45pm

資料內容—欄位值填寫的內容範圍



後設資料與後設資料標準簡介

編目紀錄 建檔者：黃XX

建檔時間：2009/11/23 3:45pm

資料值—欄位內容值的填寫格式

訂定欄位著錄規範

各類型藏品可參考之著錄規範

- 圖書館類：AACR、RDA、LCSH、LCC
- 檔案館類：APPM、DACS
- 博物館類：CCO
- 生物類：ABCD Schema

訂定欄位著錄規範

其他建議採用之欄位著錄格式與規範：

- 日期相關欄位：建議使用符合ISO8601之「YYYY-MM-DD」與「hh:mm:ss」格式
- 數位檔案格式：建議使用MIME Type列表作為欄位內容著錄之依據。MIME Type列表可參見後設資料工作組網站

http://metadata.teldap.tw/elearning/doc/MIME_Type.pdf

訂定欄位著錄規範

- 資料類型：建議使用DCMI control vocabulary

資料值	定義	描述對象
圖像	非文字的視覺呈現。	繪畫、地圖、動畫
動態影像	當連續放映時，給予動態印象的一系列視覺呈現。	電影、電視節目
靜態圖像	靜態的視覺呈現。	繪畫、版畫、照片
文字	主要由文字構成的閱讀資源。	書籍、報紙、信件
聲音	以聆聽為主要意圖的資源。	音樂、演講
互動資源	需要透過與使用者之互動，才能了解、執行或體驗的資源。	網頁
軟體	原始碼或編譯形式的電腦程式。	
合集	資源的聚集。	展覽
資料集	以明確結構編碼的資料。	清單、資料庫
服務	提供單一或多種功能的系統。	影印服務、館際互借
事	某段時間發生的非持續性事件。	展覽、會議、表演
實體物件	無生命、立體的物件或物質。	雕塑、建築、標本

六、後設資料對照服務介紹

後設資料工作組發展沿革

- Metadata Architecture and Application Team , MAAT
- <http://metadata.teldap.tw/>



The screenshot shows the MAAT website homepage with a dark blue background. At the top left is the MAAT logo and the text '中央研究院 後設資料工作組' (Academia Sinica Metadata Working Group) and '數位典藏與數位學習國家型科技計畫' (National Science Foundation Digital Library and Digital Learning National Science Foundation Project). A 'SITE MAP' button is on the top right. A central navigation menu lists: 最新消息 (Latest News), 計畫簡介 (Project Introduction), 參考規範 (Reference Standards), 工作計畫 (Work Plan), 實務規劃 (Practical Planning), 系統工具 (System Tools), 計畫成果 (Project Results), FAQ, 聯絡與建議 (Contact and Suggestions), and 數位典藏與數位學習國家型科技計畫 (National Science Foundation Digital Library and Digital Learning National Science Foundation Project). An 'RSS' button is next to the first item. At the bottom right is a large circular graphic with the text 'Metadata Architecture and Application Team' and a globe icon. The footer contains the project name, copyright notice, email address 'metadata@gate.sinica.edu.tw', and the update date '08/10/2009 10:51:18'.



後設資料服務介紹

聯合目錄資料匯入協助服務：

依據數位典藏國家型科技計畫第一分項拓展臺灣數位典藏計畫之管考作業，各項參與計畫須於完成後將資料匯入聯合目錄，後設資料工作組為配合此項流程，亦提供聯合目錄資料匯入服務，協助各項參與計畫進行都柏林核心集（Dublin Core）的正確比對。同時研擬品質評鑑指標，進一步確保聯合目錄資料呈現之品質。

服務申請機制：

http://metadata.teldap.tw/design/mapping/service_apply_mapping.php

後設資料對照服務流程

- 一. 於後設資料工作組網路填寫服務申請單
- 二. 後設資料工作組聯絡窗口寄出授權同意書
- 三. 回覆授權同意書與提供十筆後設資料範例(包含圖片或影音檔)，如有網站請一併提供網址
- 四. 聯絡窗口回覆計畫助理此次對照服務負責同仁
- 五. 由負責同仁與計畫助理就比對服務內容與疑問進行溝通
- 六. 負責同仁寄發比對報告初稿，計畫助理於七日內回覆確認
- 七. 由聯絡窗口正式發布，主題計畫將比對報告、後設資料、藏品電子檔、網站原始連結一併交予聯合目錄。

後設資料服務申請機制

實務規劃

[聯合目錄匯入協助](#)

[後設資料服務申請](#)

[後設資料生命週期作業模式](#)

[作業表單及填寫範例](#)

鑑於多數主題計畫皆有後設資料方面的需求，為能加速與維持一定水準的品質，我們依據「數位學習與數位典藏國家型科技計畫」的實際需求情形，參酌電腦科學之系統分析與人文社會科學的內涵分析等方法，設計「[後設資料生命週期作業模式](#)」，做為本組後設資料的服務程序，並依此方法轉化為系統化的服務模式。除了加速數位典藏國家型科技計畫逾八十個主題計畫的後設資料應用與系統開發外，同時也達成品質保證、品質一致性與計畫管理等多重目的。

因應二期計畫之調整與各計畫之需求，後設資料工作組之任務目標與服務內容有所調整。請有後設資料服務需求之各計畫或使用者，改以線上方式提出相關之需求與進行服務申請，將視各主題計畫需求，提供不同的服務內容與方式，協助各主題計畫完成後設資料的設計、建置及國際的標準化與接軌。

■ [後設資料生命週期作業模式](#)

「[後設資料生命週期作業模式](#)」分四個階段十項程序，以系統化的程序和流程，服務主題計畫的後設資料需求：

- **需求評估與內涵分析：**需求訪談、計畫相關標準與個案觀察、深入分析後設資料需求、確認後設資料策略暨標準間的互通性。
- **後設資料需求功能書：**研製後設資料功能需求書、評估後設資料系統。
- **後設資料系統：**研製參考規範、開發後設資料系統。
- **服務與評估：**後設資料服務、後設資料作業評估。

聯合目錄匯入協助服務 – 圖 1

聯合目錄匯入協助服務申請表	
※ 為必填欄位。	
計畫或專案名稱：	※ 中文： <input type="text"/>
	※ 英文： <input type="text"/>
執行機構：	<input type="text"/>
執行期間：	自民國 <input type="text"/> 年 <input type="text"/> 月 <input type="text"/> 日起至民國 <input type="text"/> 年 <input type="text"/> 月 <input type="text"/> 日
服務申請人（TELDAP項下所屬計畫請填計畫主持人）：	
※ 姓名：	<input type="text"/> (中文)
	<input type="text"/> (英文)
職稱：	<input type="text"/>
※ 聯絡電話：	<input type="text"/>
※ E-mail：	<input type="text"/>
聯絡窗口： 姓名：	<input type="text"/>
	聯絡電話： <input type="text"/>
	<input type="text"/>

聯合目錄匯入協助服務 – 圖 2

聯絡窗口： 姓名：

聯絡電話：

E-mail：

1. 是否為數位典藏與數位學習國家型科技計畫項下計畫？ 是 否
(答「是」者，請續答下題；答「否」者，可不必回答下題)

2. 計畫所屬之分項計畫？

- 第一分項：拓展臺灣數位典藏計畫
- 第二分項：數位技術研發與整合計畫
- 第三分項：數位核心平台計畫
- 第四分項：數位典藏與數位學習之學術與社會應用推廣計畫
- 第五分項：數位典藏與數位學習之產業發展與推動計畫>
- 第六分項：數位教育與網路學習計畫
- 第七分項：語文數位教學計畫
- 第八分項：數位典藏與數位學習之海外推展暨國際合作計畫計畫

3. 計畫書上傳：

或提供網址下載：

(計畫書中若有個人隱私及不宜公開之資訊，如：身份資料、預算分配等，請刪除後再行上傳。)

聯合目錄匯入協助服務 – 圖 3

4. 貴專案或計畫是否有製作任何系統開發需求規格書(包含著錄欄位名稱、欄位定義、範例、代碼表及各種系統需求)?

若有請提供檔案或網址。

請上傳檔案(若超過一個檔案請壓縮為zip或rar檔)

 瀏覽...

或網址：

承上題，若無，請提供「著錄欄位名稱及欄位定義」相關檔案。

請上傳檔案(若超過一個檔案請壓縮為zip或rar檔)

 瀏覽...

或網址：

5. 請提供貴專案或計畫後設資料欄位與Dublin Core之對照表。

請上傳檔案(若超過一個檔案請壓縮為zip或rar檔)

 瀏覽...

或網址：

6. 請提供貴專案或計畫後設資料範例及其數位圖檔共10筆。

請上傳檔案(若超過一個檔案請壓縮為zip或rar檔)

 瀏覽...

或網址：

7. 貴專案或計畫是否已完成系統建置？若有請提供網址（及必要的測試帳號）。

網址：



問題與討論



The End

歡迎指正

- 數位典藏與數位學習國家型科技計畫
數位技術研發與整合計畫
後設資料工作組
- 中研院計算中心後設資料工作組
- 沈漢聰
htshen@gate.sinica.edu.tw
02-27839766 轉 223