



後設資料的規劃與建置作業介紹

講師：沈漢聰

中華民國101年3月30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

- 後設資料的簡介
- 後設資料的規劃與建置作業





一、後設資料的簡介





數位典藏基本流程與後設資料關係

數位典藏國家型科技計畫技術彙編：

數位典藏係指數位典藏品以數位形式典藏的過程，更精確的解釋為數位典藏品長期的儲存、維護及檢索取得（**accessibility**）。數位典藏品有些是原生的，本就是數位資料媒體（**born digital media**），有些則因典藏而改變成數位化格式的物件。數位典藏的主要目的是確保數位資料的可用性、持久性，及智慧整合性。

(來源：http://www.ndap.org.tw/2_techreport/index.php?pid=6)



數位典藏基本流程與後設資料關係

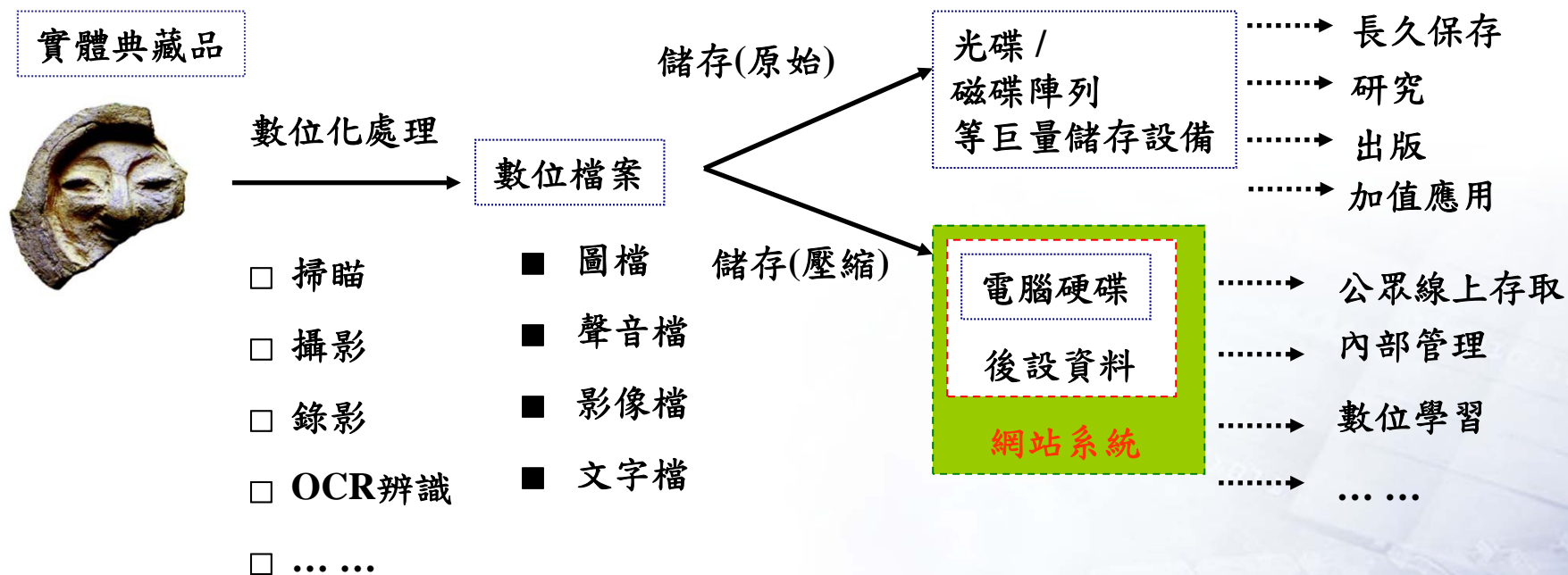
Wikipedia :

數位典藏 (digital archive) 指的是以數位方式處理(拍攝、全文輸入、掃描等), 並加上元數據(Metadata)的描述, 以數位檔案的形式儲存。

(來源 : <http://zh.wikipedia.org/wiki/%E6%95%B8%E4%BD%8D%E5%85%B8%E8%97%8F>)



數位典藏基本流程與後設資料關係





後設資料與後設資料標準簡介

『 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





後設資料與後設資料標準簡介

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。



後設資料與後設資料標準簡介

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，衣襬57.0cm

來源：中研院民族所台灣原住民文物計畫



後設資料與後設資料標準簡介

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



來源：故宮器物數位典藏子計畫



二、後設資料的規劃與建置作業





後設資料建置流程1/2

STEP 1：挑選典藏品

STEP 2：典藏品編碼

STEP 3：規劃後設資料欄位（建議以 Dublin Core為基礎）

STEP 4：進行典藏品描述，著錄欄位

STEP 5：彙入聯合目錄



STEP 1：挑選典藏品

- 檢視計畫的產出，挑選需要典藏的典藏品（原生為數位或非數位）。
- 清點典藏品之數量，整理成清冊。
- 歸納類型。
- 釐清各典藏品類型單元與單元間可能存在的關係（如：組、附件關係）。



STEP 1: 挑選典藏品

藏品類型範例-文字類



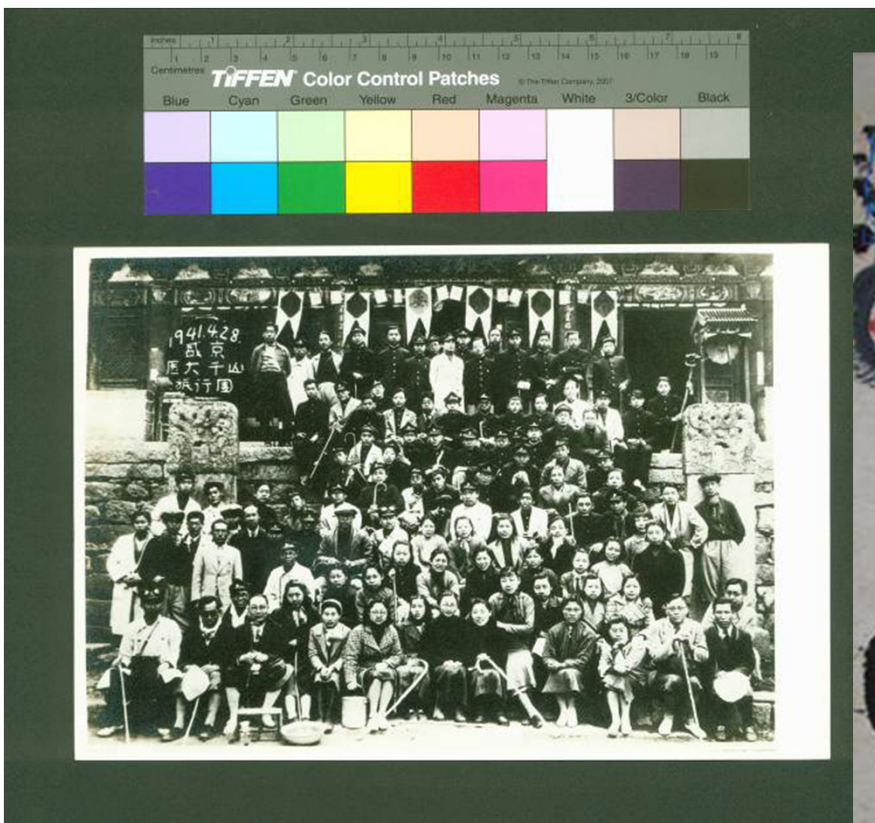
雜誌

報紙



STEP 1：挑選典藏品

藏品類型範例-靜態圖像類



照片



圖畫



STEP 1：挑選收藏品

藏品類型範例-聲音類



如：黑膠唱片、
錄音帶、CD、
數位錄音檔





STEP 1：挑選典藏品

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



STEP 1：挑選典藏品

藏品單元關係分析



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：中研院民族所台灣原住民文物計畫

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

藏品單元關係分析



• 範例1：



創作者、創作年代、創作背景都相同，藏品後設資料資訊大概雷同

欄位名稱		範例
名稱		墨寶
作品	編目層級	組件
	數量	2
	類型	書法

藏品單元關係分析

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	

來源：故宮器物數位典藏子計畫



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



STEP 2：典藏品編碼

- 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。



STEP 2：典藏品編碼

- 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



STEP 2：典藏品編碼

• 藏品編碼範例 - 智慧型編碼：

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)。



STEP 2：典藏品編碼

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



STEP 2：典藏品編碼

藏品識別碼設計與編碼

範例一、書籍與期刊

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。



STEP 2：典藏品編碼

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



STEP 2：典藏品編碼

- 藏品編碼範例 - 智慧型編碼：
 - 「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）

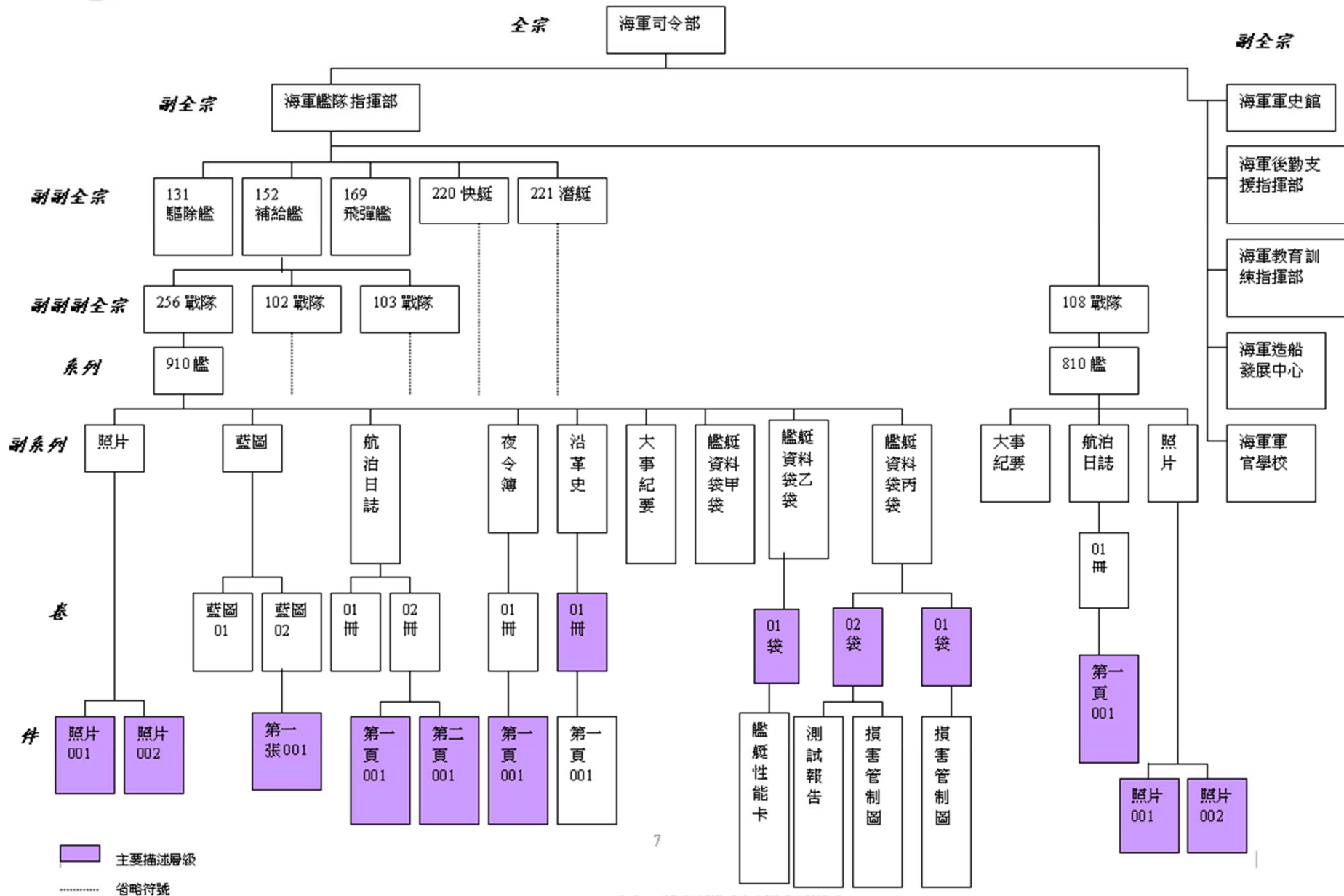


圖 2 海軍檔案層級架構圖



STEP 2：典藏品編碼

藏品識別碼設計與編碼



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



STEP 3

STEP 3

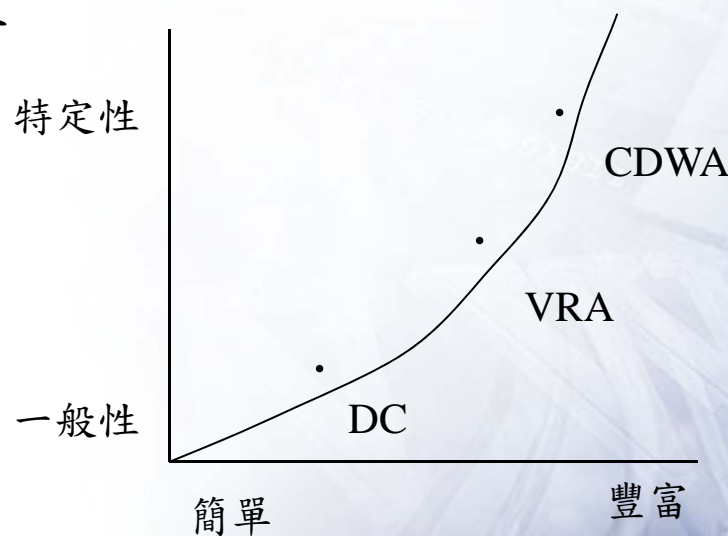
STEP 3

後設資料標準繁簡度

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





STEP 3

STEP 3

STEP 3

後設資料與後設資料標準簡介

博物館類：

CDWA、The CIDOC Information Categories、VRA Core Categories

檔案館類：

EAD、ISAD(G)、TEI

生物類：

Species 2000 Data Standard、HISPID

VS

一般性：Dublin Core



STEP 3

STEP 3

STEP 3

後設資料與後設資料標準簡介

藝術品描述類目

(CDWA , Categories for the Description of Works of Art) :

CDWA是由J.Paul Getty Trust下的Art Information Task Force(AITF)所發展的後設資料標準，主要的描述目標為藝術品、建築、物件群，其欄位的設計是希望可以滿足藝術史研究的需求，提供給研究者一致且可信的內容，共有27個類目（categories），細分了233個大小元素。

官方網站：http://www.getty.edu/research/conducting_research/standards/cdwa/

CDWA中文版：<http://www.sinica.edu.tw/~metadata/standard/CDWA/Element%20List.htm>

CDWA類目

- 物件/作品 (OBJECT/WORK)
- 分類 (CLASSIFICATION)
- 方位/配置 (ORIENTATION/ARRANGEMENT)
- 題名 (TITLES or NAMES)
- 階段 (STATE)
- 版本 (EDITION)
- 測量值 (MEASUREMENTS)
- 材質與技術 (MATERIALS and TECHNIQUES)
- 製作手法 (FACTURE)
- 形式描述 (PHYSICAL DESCRIPTION)
- 題款/標記 (INSCRIPTIONS/MARKS)
- 現況/鑑定歷史 (CONDITION/EXAMINATION HISTORY)
- 保存/處理歷史 (CONSERVATION/TREATMENT HISTORY)
- 創作 (CREATION)
- 所有權/收藏歷史 (OWNERSHIP/COLLECTING HISTORY)
- 版權/限制 (COPYRIGHT/RESTRICTIONS)
- 風格/時期/團體/運動 (STYLES/PERIODS/GROUPS/MOVEMENTS)
- 題材 (SUBJECT MATTER)
- 脈絡 (CONTEXT)
- 展覽/借展史 (EXHIBITION/LOAN HISTORY)
- 相關作品 (RELATED WORKS)
- 相關視覺文獻 (RELATED VISUAL DOCUMENTATION)
- 相關文本參考資料 (RELATED TEXTUAL REFERENCES)
- 批評性回應 (CRITICAL RESPONSES)
- 編目史 (CATALOGING HISTORY)
- 現藏地點 (CURRENT LOCATION)
- 描述註 (DESCRIPTIVE NOTE)

採用CDWA的計畫

- 器物類
 - 國立臺灣戲曲學院數位博物館—京劇文物典藏
 - 器物主題小組後設資料共通元素建議書 修訂版
 - 故宮文物數位典藏系統之開發與建立-故宮器物子計畫
 - 臺灣原住民數位典藏計畫：文物資料庫後設資料需求規格書v1.0
- 建築類
 - 大型歷史建築文物數位保存計畫
- 遺物、遺址
 - 臺灣史前文化教育推廣創意加值計畫：卑南遺址與卑南文化
- 書畫
 - 近現代珍貴數位典藏系統之研製
 - 國家歷史文物數位典藏計畫
 - 故宮文物數位典藏系統之開發與建立-故宮書畫子計畫
- 調查資料、標本、照片
 - 珍藏歷史文物數位典藏計畫
 - 分項一：考古發掘標本、照片、記錄與檔案數位典藏計畫考古發掘標本、照片、記錄與檔案需求規格書0.1版
 - 分項二：拓片與古文書數位典藏計畫
 - 分項四：民族學調查標本、照片與檔案數位典藏



STEP 3

STEP 3

STEP 3

後設資料與後設資料標準簡介

檔案編碼描述格式標準

(EAD , Encoded Archival Description) :

EAD是一種結構化的檔案檢索工具(finding aids)標準，其發展目的是為了支援檔案和手稿的收集保存，提供一個永久編碼標準，可機讀處理各種單位（如檔案館、圖書館、博物館與手稿館）所產生的目錄、登錄簿、索引等查檢工具，以利檔案資源易於在網路上取得。

官方網站：<http://www.loc.gov/ead/>

EAD 中文版：http://www.sinica.edu.tw/~metadata/ead/ead_titlepage_3.htm

採用EAD的計畫

- 近代歷史研究所-檔案館外交經濟重要檔案數位典藏計畫
- 歷史語言研究所－珍藏歷史文物數位典藏計畫-分項五：內閣大庫檔案著錄格式、權威檔建立
- 臺灣省諮議會及中研院臺灣史研究所－史料典藏數位化計畫-台史所-楊雲萍文書
- 臺灣省諮議會及中研院臺灣史研究所－史料典藏數位化計畫-臺灣省諮議會
- 國史館典藏國家檔案與總統文物數位化中程計畫
- 國史館臺灣文獻館-典藏日據與光復初期史料數位化計畫



STEP 3

STEP 3

STEP 3

後設資料與後設資料標準簡介

都柏林核心集

(DC , Dublin Core) :

Dublin Core後設資料標準，是一組簡單卻有效的核心元素集。為了要加速網路電子資源的整理與組織，並加強網路資源的找尋與檢索的精確性，因此集合了圖書館界、資訊科學等各領域專家，共同制定專為描述網路電子資源的一套後設資料格式。此核心集僅有15個元素。

官方網站：<http://dublincore.org/>

DC 中文版：<http://www.sinica.edu.tw/~metadata/standard/dublincore-chi.htm>



STEP 3

STEP 3

STEP 3

DC Application Profile

TELDAP 聯合目錄：

- 以Dublin Core 作為描述格式。
- Dublin Core The One-to-One Principle。
- 6個必填欄位。
- 欄位著錄格式：分隔、斷行符號
- 匯入系統之檔案格式：excel 檔、XML檔

都柏林核心集的最佳實務指引：<http://metadata.teldap.tw/project/filebox/metrics/DCPB.pdf>



STEP 3

STEP 3

STEP 3

Simple-DC 15個元素

- 標題 (Title)
- 主題 (Subject)
- 描述 (Description)
- 型式 (Type)
- 來源 (Source)
- 關聯 (Relation)
- 涵蓋範圍 (Coverage)
- 創作者 (Creator)
- 出版者 (Publisher)
- 貢獻者 (Contributor)
- 權利 (Rights)
- 日期 (Date)
- 格式 (Format)
- 識別碼 (Identifier)
- 語言 (Language)



STEP 3

STEP 3

STEP 3

後設資料與後設資料標準簡介

更多的後設資料標準介紹，
請參閱後設資料工作組網站之參考規範：

<http://metadata.teldap.tw>





STEP 3

STEP 3

STEP 3

參考規範



參考規範

| 依類型 | 依字母 | 中文版 | 標準比對

後設資料工作組依據所支援主題計畫的屬性，彙整國際上相關主題的後設資料參考資源，並提供主題計畫的應用典範，以做為使用者進行後設資料之分析規劃的參考規範。在此將後設資料的應用分為三大類：

- **整合性**：以公眾檢索為導向，以Dublin Core為代表。
- **學科導向**：依據所支援計畫的社群、學科為導向，目前有博物館、檔案館、善本、考古、語言/語料庫、地理空間資訊、影音/多媒體、數位學習、生物多樣性等九大類。
- **權威控制**：以「學科導向」為主，「權威控制」為輔，建立於人、事、時、地、物五大主軸間交錯的互動關係，以達到不同知識層次後設資料的分析與詮釋。

內容包括：

- **後設資料標準簡介**：含建議採用與其他相關的標準介紹，以類型、字母順序二種方式呈現。
- **中文版後設資料標準**：為本組所研譯之標準中文版。
- **主題計畫應用典範**：如主題計畫之後設資料元素與範例、主題相似的計畫間後設資料共通元素。
- **國際上相關計畫或組織**：各領域、國際性、區域性、或個別應用的相關計畫。
- **後設資料標準比對**：為不同後設資料格式間元素的對照表。

後設資料標準簡介－依類型

整合性

- 建議標準
 - Dublin Core 簡介 元素清單 都柏林核心案的最佳實務指引 [New](#)
 - METS 簡介

學科導向

博物館

- 計畫屬性
藝術/美學、視覺藝術、考古、人類學。
- 媒材類型
書畫、器物、金石拓片、簡牘、民俗文物等工藝作品。
- 建議標準
Categories for the Description of Works of Art (CDWA) 簡介
中文版 (v2.0 draft)
- 主題計畫元素清單與範例
 - 故宮書畫後設資料(V1.0/2001)與CDWA對照表(PDF)
 - 故宮器物後設資料(V1.2/2004)與CDWA對照表(PDF)

檔案館

- 計畫屬性
機構、團體的公文、文書；私人具紀念價值之文書、文物、照片等
- 建議標準
Encoded Archival Description (EAD) 簡介 中文版 (標準元素辭典 (EADTag Library) v2002)
- 共有元素一覽
範例取自內閣大庫、近史檔案等計畫
- 主題計畫元素清單與範例
 - 內閣大庫檔案著錄格式、權威檔建立
 - 近史所檔案館外交經濟重要檔案數位典藏計畫

<http://metadata.teldap.tw/standard/standard-frame.html>



問題與討論





製作群

後設資料工作組全體同仁





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
Tel : 02-27839766 轉 2231
Email : htshen@gate.sinica.edu.tw