

# 後設資料(Metadatas)標準與 個案的實際應用

---

中央研究院計算中心

陳淑君

[sophy@sinica.edu.tw](mailto:sophy@sinica.edu.tw)

2002年8月2日

數位典藏專業培訓課程

# 本課程預期目標

---

- 深入了解二種後設資料標準與應用原則
- 獲取個案應用後設資料的方法
- 學習如何著手進行後設資料作業規劃
- 實際規劃並著錄所屬主題計畫的後設資料個案

# 本課程數位典藏/博物館計畫後 設資料需求與屬性<sub>1</sub>

社群

圖書館、檔案館、博物館、標本館...

學科

文學、人類學、物理治療、歷史、政治、  
動/植物學、藝術、佛學、礦物學、化石、  
工業技術、語言學...

資料類型

善本古籍、器物、拓片、檔案、田野資料、  
織品服飾、布袋戲文物、科技文物...

功能/目的

整理、搜尋、檢索、管理、描述、交換、  
利用、保存、國際接軌、  
形成完整數位典藏系統...

# 本課程數位典藏/博物館計畫後設資料需求與屬性<sub>2</sub>

---

- 目前採用的後設資料標準
  - DC (Dublin Core)
  - DC based
  - EAD (Encoded Archival Description)
  - CDWA
  - CIDOC
  - OLAC
  - ?...

# 本課程數位典藏/博物館計畫後設資料需求與屬性<sub>3</sub>

---

- **Metadata**知識需求
  - 後設資料與數位典藏間的關係
  - 後設資料的建置流程
  - 後設資料的分析方法
  - 後設資料與權威檔、索引典的關係
  - 後設資料與著錄規則
  - 後設資料著錄實務
  - 後設資料系統開發技術
  - 後設資料的調整原則

# 後設資料與數位典藏的關係

- **數位典藏**目前面臨的重要問題：
  - **Discovery**:如何找到龐大的數位典藏品?
  - **Longevity**:如何確保數位典藏品的長久保存與取得?
  - **Interoperability**:如何讓數位典藏品達到互通與共享?
- **後設資料**是解決上述問題的一種方法

# 後設資料與數位典藏的關係

- 一項內涵結構化的標準與技術
- 傳統的內涵資訊把文本（data/content）與呈現程序（presentation/procedure）溶合一起
- 在數位典藏中，若將兩者分開處理：
  - 對內涵進行結構化分析與描述、發展
  - 多樣化的呈現樣式
- 如此，我們的數位典藏品將可視使用者的需求與目的，以多元化的方式重新組合、呈現、再使用(reuse)與轉換(transformation)，及有效的知識組織與檢索(access)

# 後設資料的類型

## ■ 描述性後設資料

- 描述一項文件或資源的內涵與關聯性，以便發現與辨識資源。如：書目記錄、Dublin Core.

## ■ 結構性後設資料

- 給予數位典藏品實質的結構，以便瀏覽、檢索與呈現上述資源。如：讓電子全文具翻頁功能、全文與相關影像的適切連結、書的章節結構。

## ■ 管理性後設資料

- 為俾便長久管理、使用、觀看與保存數位化資源的相關資料。如：檔案格式、數化解析度、多媒體資訊壓縮、像素尺寸、智財權管理資訊。



# 後設資料的組成結構

---

- 資料結構/語意 (Data Structures/Semantics)
  - 如：DC, CDWA, VRA Core Categories, EAD, MARC...
- 資料值/詞彙 (Data Values/Vocabularies)
  - 如：AAT, ULAN, LCSH
- 資料內容/語法 (Data Content/Syntax)
  - 如：AACR2, ISAD

# 後設資料標準種類與清單列舉<sub>1</sub>

---

- 書目性
  - MARC (Machine Readable Cataloging)
  - DUBLIN CORE
  - GILS (Government Information Locator Service/ Global Information Locator Service)
  - TEI Headers (Text Encoding Initiative)

# 後設資料標準種類與清單列舉<sub>2</sub>

---

- 影像或實體物件
  - Categories for the Description of Works of Art (CDWA)
  - VRA Core Categories version 3.0
  - NISO Data Dictionary for Technical Metadata for Digital Still Images
  - CIDOC

# 後設資料標準種類與清單列舉<sup>3</sup>

---

- 地理空間資訊
  - Content Standards for Digital Geospatial Metadata (CSDGM)
- 音樂性資訊
  - Standard Music Description Language (SMDL)
- 數據性資訊
  - ICPSR Data Documentation Initiative
  - Standard for Survey Design and Statistical Methodology Metadata (SDSM)

# 後設資料標準種類與清單列舉<sup>4</sup>

---

- 合集層次的描述
  - EAD (Encoded Archival Description)
- 電子商務
  - The INDEXS project
- 智財權管理
  - Rights Metadata
  - DOI (Digital Object Identifier)
- 影音媒體
  - MPEG-7 (Moving Picture Experts Group)
  - UMID (Unique Material Identifier)

# 後設資料標準種類與清單列舉<sup>5</sup>

---

- 教育
  - IEEE Draft Standard for Learning Object Metadata
  - DC-Ed (Dublin Core Education Working Group)
- 保存典藏
  - Metadata Framework to Support the Preservation of Digital Objects
  - CEDARS Project: CEDARS Preservation Metadata Elements

# 標準與應用探討：Dublin Core 1

## ■ 背景

- 1994, Chicago, 2nd Intl World Wide Web Conference, *semantics & the Web*
- 1995, Dublin, Ohio, OCLC/NCSA Metadata Workshop, *metadata*  
*semantics*: 若能為網路資源發展一套核心組，將非常有助於網路分類，並進而讓其檢索更加容易。“Dublin Core metadata”由此定名。
- 2002/10, 將於義大利召開第10次工作會議，以及第2屆大會。

# 標準與應用探討：Dublin Core 2

- 目標與定位：更容易找到資源
  - 發展出能夠探索(**discovery**)跨領域(**across domains**)資源的後設資料標準。
  - 訂定出能夠讓不同套後設資料間互通的架構(**framework**)
  - 發展出與特定領域/社群(**discipline-specific**)後設資料標準間，能夠跨領域探索資源與後設資料互通(**interoperability**)的架構。



# 標準與應用探討：Dublin Core 3

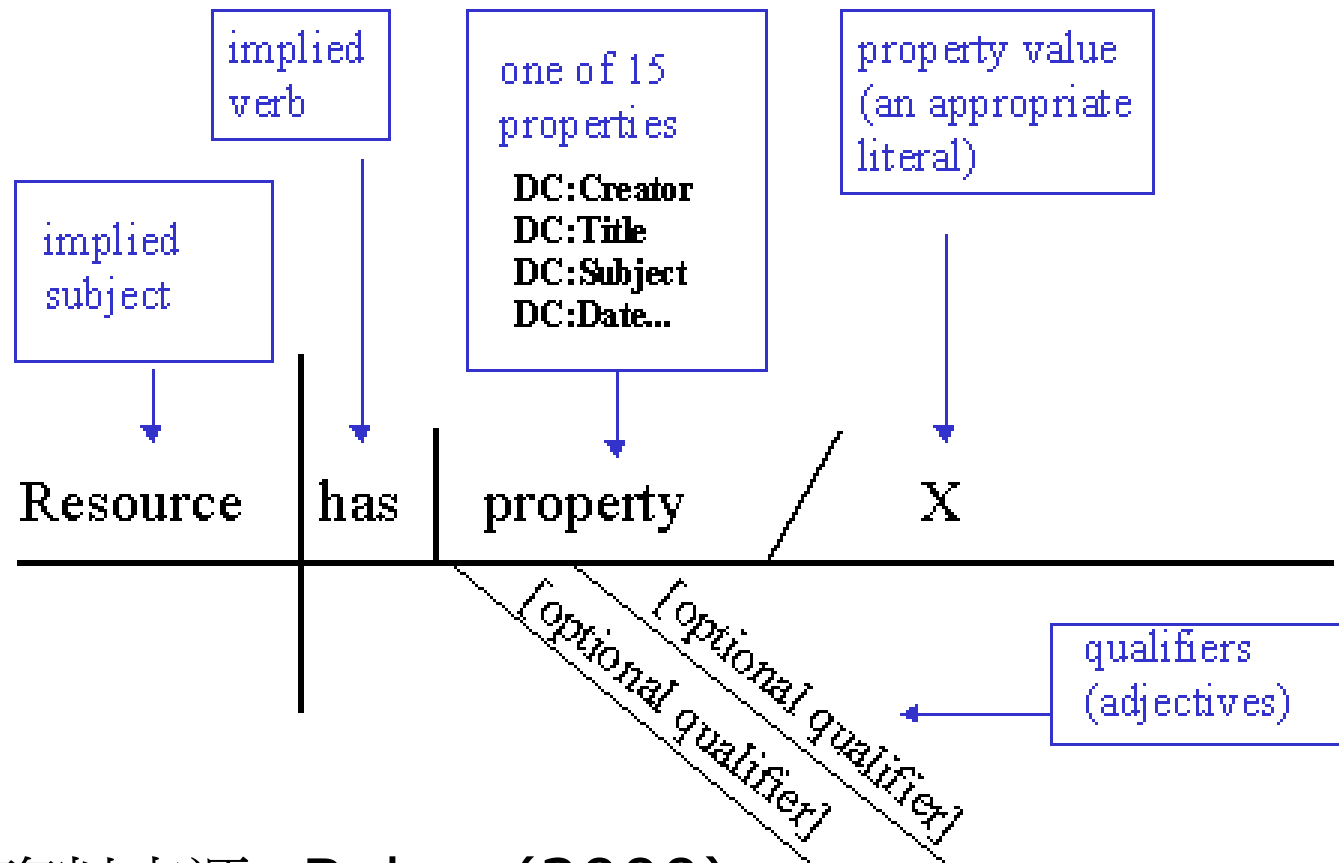
---

- 基本結構
  - DCMES (Dublin Core Metadata Element Set) Version 1.1
  - *Elements*, Element *Refinements*, Element *Encoding Schemes*
  - 請詳講義：附件一(pp.4-5)

# 標準與應用探討：Dublin Core 4

- 重要的應用原則：A grammar of DC
  - 簡單化
    - 二大層次的語言：elements, qualifiers(名，形)
    - 為數位遊客設計的洋涇濱語(A Pidgin for Digital Tourists)
  - 元素與修飾語
    - 15個元素：有助於跨資料庫與跨領域的檢索
    - Encoding schemes: 指標到相關資訊，解釋資料值
    - Element refinements: 使元素更特定與明確

# 標準與應用探討：Dublin Core 5



資料來源：Baker (2000)

# 標準與應用探討：Dublin Core 4

---

- 修飾語的的應用原則：Dumb-Down Principle, refine vs. extend
- 元素值的應用原則：是否有助於資源探索？是正確應用的核心觀念。通常，元素值會有預期出現的字串型式。
- 評估應用的良莠

# 標準與應用探討：Dublin Core 4

Resource	has	DC:Date	/"2000-06-13"
----------	-----	---------	---------------

*DCQ:Revised*  
*DCQ:ISO8601*

Resource	has	DC:Subject	/"Languages -- Grammar"
----------	-----	------------	-------------------------

*DCQ:LCSH*

資料來源：Baker (2000)

# 標準與應用探討：Dublin Core 5

---

- 工具

- 詳細的工具清單

<http://dublincore.org/tools/>

- 練習與試用 <http://www.lub.lu.se/cgi-bin/nmdc.pl>

# 標準與應用探討：Dublin Core 6

---

- 語言：洋涇濱語(pidgin)
- 字彙：elements, qualifiers
- 句子：xml
- 文法：名詞，形容詞
- 段落與文章：application profile
- 字典/庫：registry

# 標準與應用探討：CDWA 1

---

## ■ 背景與定位

- 由Getty研究機構的藝術資訊專案小組(Art Information Task Force, AITF)所發展
- 描述物件(object)與影像(image)的後設資料
- 提供一個藝術資料庫的詳盡的結構指引
- 應用單位可以視實際個案，調整之



# 標準與應用探討：CDWA 2

---

## ■ 背景與定位

- 原以學術研究的需求與觀點發展而成，並進而延申至提供博物館，展覽與保存的需求
- 使用群包括：博物館，鑒賞家，海關，執法機購，保險公司...等
- 提供保護藝術作品最基本所需的資訊

# 標準與應用探討：CDWA 3

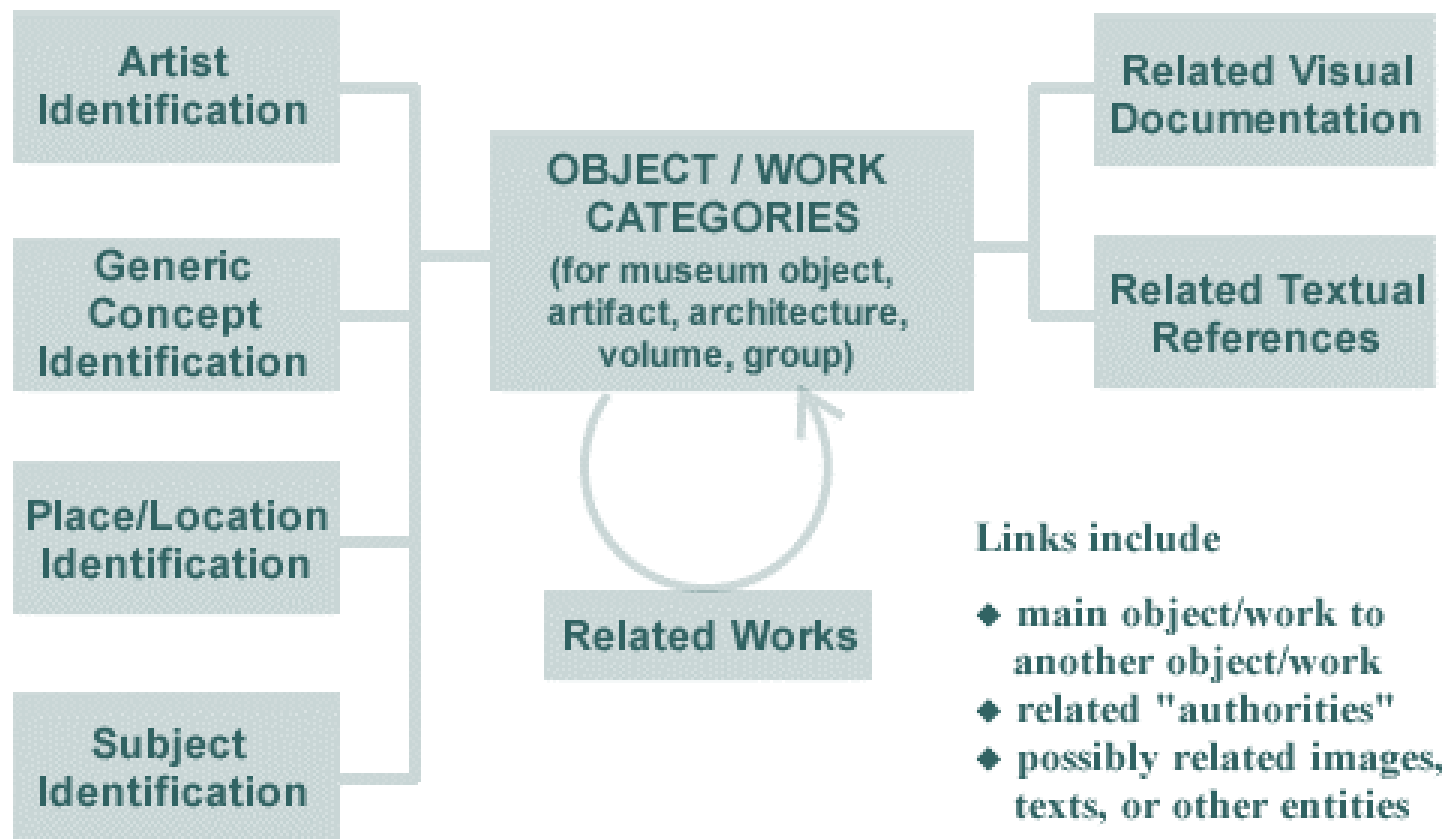
## ■ 結構：27個類目

物件／作品，題名，階段，版本，創作，現藏地點，分類，主題，背景，風格／時期，測量材質與技術，手法，形式描述，銘文／標記，現況／鑑定，保存／處理歷史，方位／佈局，展覽／借展歷史，擁有權／收藏歷史，版權／限制，編目歷史，相關作品，相關視覺檔案，相關參考文獻，批評性回應，描述註記

## ■ 請詳講義：附件二 (pp.6-36)

# 標準與應用探討：CDWA 4

## ■ 重要的應用原則：連結關係圖



資料來源：Getty & College Art Association (2000)

# 後設資料個案分析

---

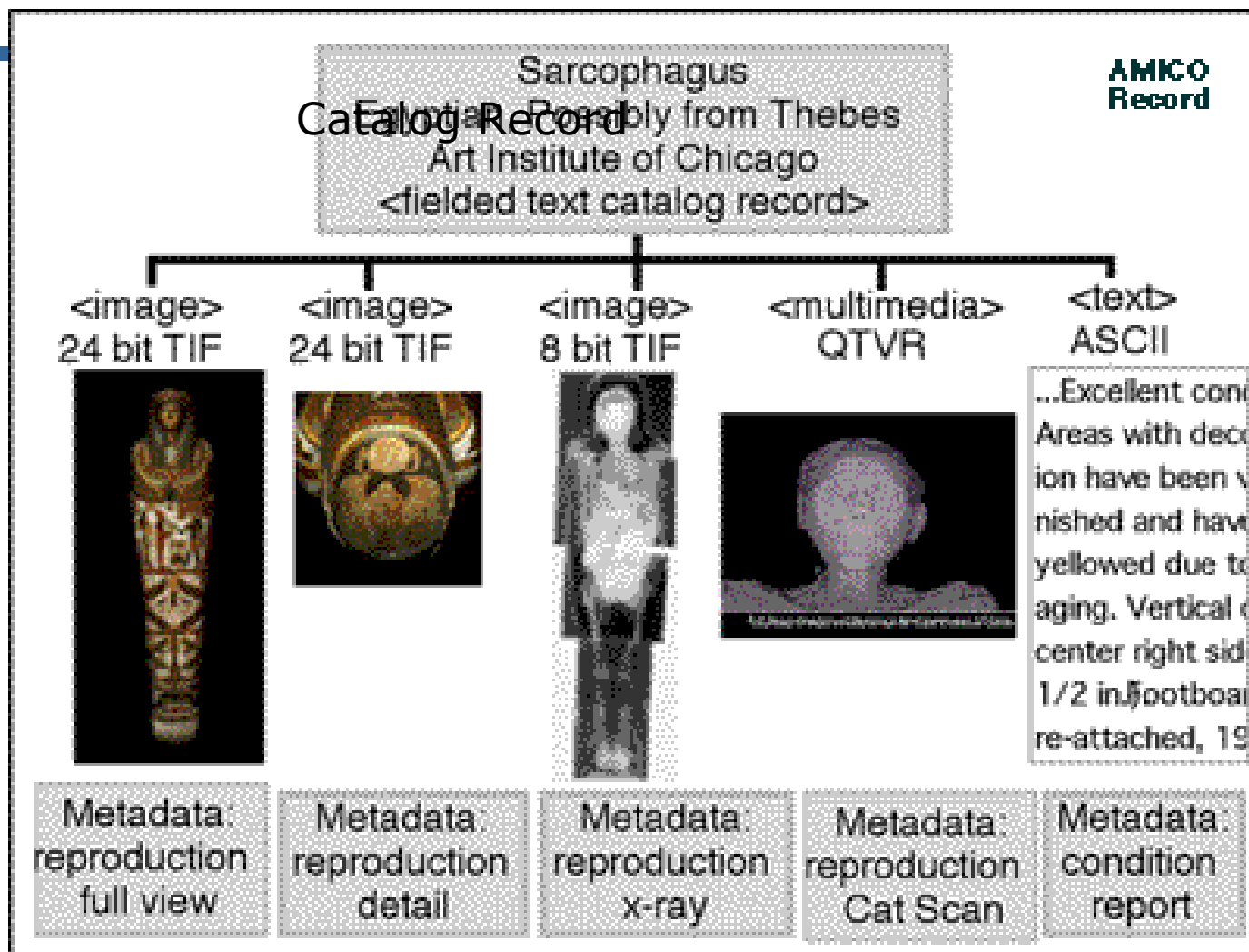
- I : California State Library Metadata Standard
- II : AMICO (the Art Museum Image Consortium)
- III : Kent State Univ. Museum

# 後設資料個案分析：AMICO

Catalog Record

Associated Digital Files

Metadata records for each file



# 後設資料的工作流程

- 確認典藏計畫的目標
- 釐訂後設資料在數位典藏計畫的功能
- 進行典藏品屬性分析
- 評估與選擇適用的後設資料標準；若必要，給予若干修改，以符合計畫所需。
- 完成需求規格書，轉交系統開發者，或找尋廠商開發。
- 進行系統測試與評估，並撰寫使用規範與指引
- 進行後設資料效益評估
- 請詳講義：附件三~十 (pp. 52-1430)

# 著錄實務

---


- 請詳附件四，五

# 參考書目

---

- Baker, Thomas. (2000). A Grammar of Dublin Core.  
<http://www.dlib.org/dlib/october00/baker/10baker.html>
- Getty & College Art Association. (2000). Entity Relationship Diagram.  
[http://www.getty.edu/research/institute/standards/cdwa/6\\_entity/index.html](http://www.getty.edu/research/institute/standards/cdwa/6_entity/index.html)





---

# 問題與討論