

後設資料規劃與應用實務

陳淑君

sophy@sinica.edu.tw

中央研究院計算中心

2002/4/26

本課程預期目標

- 建立對後設資料清晰與具體的概念
- 瞭解後設資料在數位典藏品中扮演的角色與功能
- 瞭解如何開始著手規劃與建置後設資料作業
- 瞭解如何應用與評估後設資料服務

後設資料常見基本問題

- 何謂後設資料?
- 爲什麼需要建立後設資料?
- 誰會需要使用後設資料?
- 爲何會有這麼多種的後設資料標準?
- 該如何撰擇適用的後設資料?
- 可否同時採用不同的後設資料標準?
- 如何開始建置後設資料?
- ...

後設資料的定義探討

名詞界說(1)

- Metadata: 名詞創造於1969年
- Meta-data: Metadata的另一稱法
- Metadatabase: 資料庫的資料庫
- Meta-: 變化，在~之後，超越，在~之間
- Data: 代表任何的資源

後設資料的定義探討 名詞界說(2)

- 後設資料
- 詮釋資料
- 超資料
- 元資料
- 元數據

後設資料的定義探討 定義？

- 資料中的資料 (Data about data)
- 描述資源屬性的資料 (Data describes attributes of resources)
- 有關資料背景與關聯性、資料內涵以及資料控制等相關資訊 (Information about the context of data and the content of data and the control of or over data).

後設資料的定義探討

定義? (續)

- 後設資料是有關一個數位典藏品的資料，通常由典藏品的創作者或提供者來建立，並將數位典藏品串聯或埋置於後設資料中。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。(ALA, 美國圖書館學會)
- 它最簡單的定義是“有關資料的結構性資料”。(DC, 都柏林核心集)

後設資料的定義探討

定義? (續)

- 任何協助辨識、描述與放置網路化電子資源的資料。目前存在許多不同的後設資料格式，有些在描述上很簡單，有些則相當複雜而豐富。
(IFLA, 國際圖書館協會)
- 記載資料元素或屬性（如：名稱、大小、資料類型等）的資料、有關記錄或資料結構（如：長度、欄位、行列等），以及有關資料的資料（如：位置、關聯、擁有者等）。
- 後設資料就是資料 (Metadata is data)
- ...超過30種的定義

後設資料的面貌探討

不同社群的習慣稱呼

- 電腦界：資料字典（Data Dictionary）
- 圖書館界：圖書館目錄（Library Catalogue）
- 博物館界：記錄、登錄、資料標準
（Documentation, Register, Data Standard）
- 檔案館界：檢索工具（Finding Aid）
- ...

後設資料的作用

- 不同的後設資料格式支援不同的目的需求。
 - 簡單 vs. 豐富
 - 一般性 vs. 特定性
- 數位典藏品的擁有者、管理者與提供者希望後設資料能夠協助他們：
 - 建置數位資源控制的機制
 - 管理數位資源
 - 數位資源使用與檢索

後設資料的作用（續）

- 數位典藏品的終端使用者希望借助後設資料協助他們：

- 找尋
- 辨識
- 撰擇
- 獲取使用
- 詮釋

這些數位資源。

後設資料與數位典藏的關係

數位典藏面臨的環境與議題

- 如何找到龐大的數位典藏品? (Discovery)
- 如何確保數位典藏品的長久保存與取得?
(Longevity)
- 如何讓數位典藏品達到互通與共享?
(Interoperability)

後設資料能夠/如何支援上述的議題?

後設資料的類型

以功能為導向

- 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。

後設資料的類型

以組成結構為導向

- 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

後設資料建立的形式

- 理置在網頁中 (Embedded format)
 - 在網頁HTML編碼中，使用META tags放置後設資料。
- 獨立的記錄 (Stand-alone records)
 - 分離後設資料與資源本身
- 放置在資料庫 (Database records)
 - 建置於資料庫，再連結至所描述的資源或數位典藏品。

後設資料的規劃作業

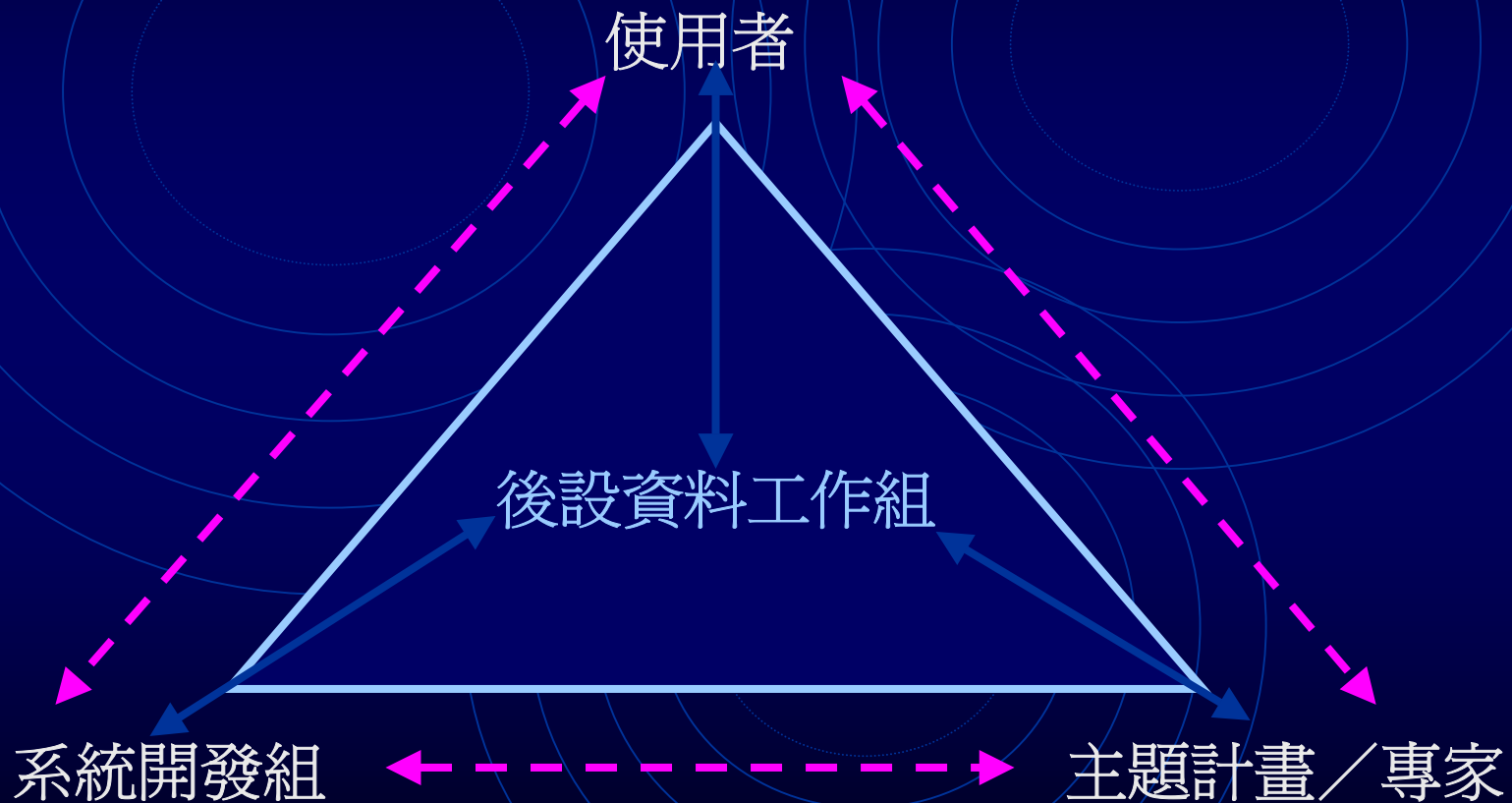
- 後設資料在於數位典藏計畫中扮演的目的, 目標與角色為何?
- 後設資料的建立方式?
- 預期的使用者特質與需求?
- 後設資料預計的範圍與深度?
- 是否已有類似數化計畫的後設資料可以借用?
- 後設資料標準的適用性與評估
- 後設資料架構的建立

後設資料計畫書

- Metadata計畫的範圍, 目的與目標
- 檢視資源 (人力, 設備, 預算)
- 確認可行的方案與重點內容
- 作業時程與項目(Schedule, deadlines)
- 計畫的架構及相關技術

後設資料工作流程

個案：數位典藏國家型科技計畫後設資料工作組
(1)



後設資料工作流程

個案：數位典藏國家型科技計畫後設資料工作組
(2)

- 後設資料需求分析
- 後設資料標準分析
- 後設資料標準建議
- 後設資料國際接軌
- 後設資料需求規格書
- 轉交後設資料系統開發組

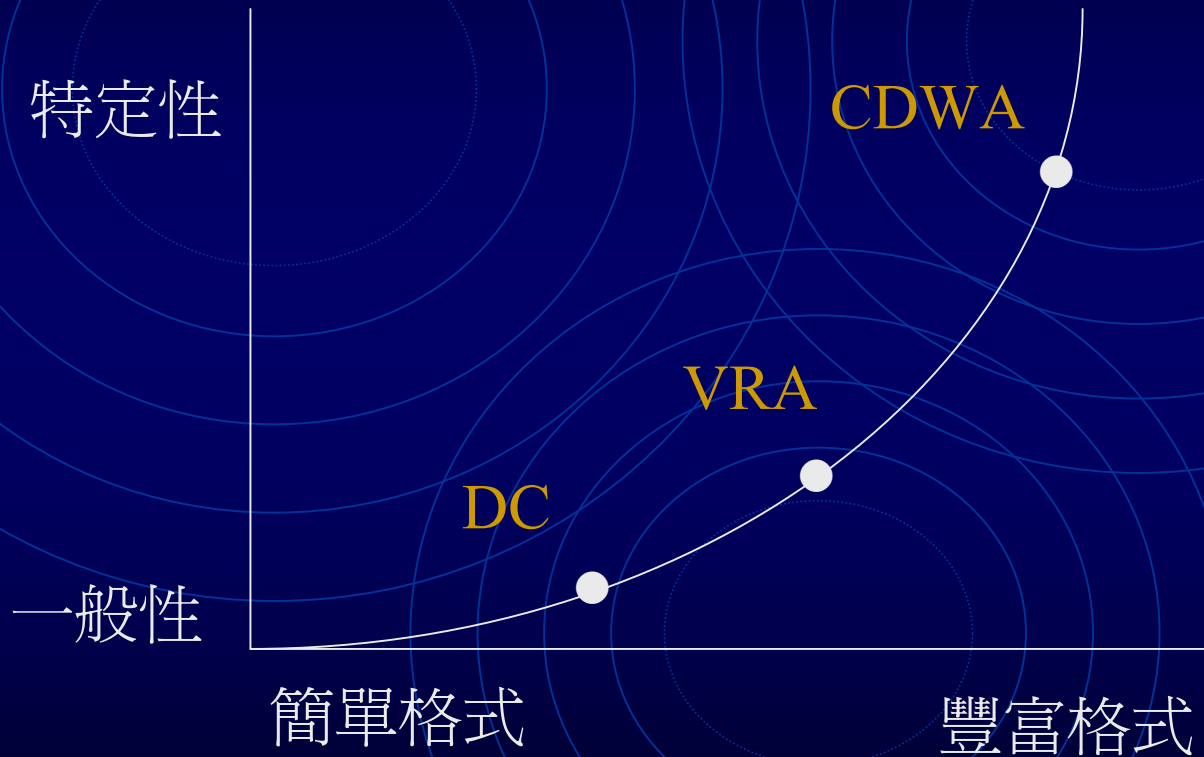
後設資料工作流程

個案：數位典藏國家型科技計畫後設資料工作組
(3)

- 後設資料系統/工具測試與建議
- 後設資料互通機制
- 後設資料使用指引與範例製作
- 後設資料服務
- 後設資料應用評估

後設資料的標準探討

標準的詳簡類型



後設資料的標準探討

一般性後設資料

- 特色：簡單，具延展性，跨學科，跨資料類型。
- 典型代表：都柏林核心集 (Dublin Core)，強調支援數位典藏品的資源探索 (Resource Discovery)。
- 15個元素組成：
 - 題名 (Title)，主要作者 (Creator)，其他作者 (Contributor)，主題 (Subject)，描述 (Description)，出版者 (Publisher)，日期 (Date)，型式 (Format)，類型 (Type)，識別碼 (Identifier)，來源 (Source)，語言 (Language)，關聯性 (Relation)，時空範圍 (Coverage)，著作權 (Rights)

後設資料的標準探討

特定性後設資料

- 特色：著重特定領域資訊的共同需求與著錄標準。
- 典型代表
 - 圖書館社群：MARC
 - 博物館社群：CDWA，CIDOC
 - 檔案館社群：EAD
 - 視覺資源社群：VRA Core Categories
 - 植物標本館社群：HISPID
 - 教育社群：IEEE LTSC LOM

後設資料的標準探討

特定性後設資料(續)

- 藝術品描述類目 (Categories for the Description of Works of Art, CDWA)，包括26項主類目：

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

後設資料的標準探討

標準的選擇方法 (VADS)

- 英國視覺藝術資料服務的建議 (VADS, UK)
 - 檢視是否符合預定的目的 (Fitness for purpose) ?
 - 檢視標準的聲譽 (Reputation) ?
 - 檢視採用採用單位的經驗 (Existing experience) ?

後設資料的標準探討

標準的選擇方法 (UKOLN)

- 英國圖書館與資訊網路辦公室的建議 (UKOLN, UK)
 - 沒有“一套標準適用所有計畫” (one size fits all) 的解決方案，端視後設資料的目的與需求而定。
 - 檢視其他計畫/單位是否已採用？
 - 檢視標準是否提供實踐指引？
 - 檢視標準是否具備某程度的共識，以便利交換與取用？
 - 檢視標準是否已有軟體工具的支援？

後設資料的標準探討

標準的選擇方法 (Kent State Univ.)

- 個案研究：肯特州立大學博物館
- 典藏品：歷史性時尚與服裝典藏品(三度空間)
- 計畫目標：為數位典藏品發展一套目錄
- 方法：分析與評估現存的後設資料格式的適用性

後設資料的標準探討

標準的選擇方法 (Kent State Univ.)(續)

- 候選的後設資料格式：
 - USMARC/AACR
 - DC
 - VRA Core
- 選擇的原則：需考量計畫所需的二大重點
 - 相容性 (Compatibility)
 - 簡易性 (Simplicity)

後設資料的標準探討

標準的選擇方法 (Kent State Univ.)(續)

- 分析與評估方法 (1)
 - 從90件數位典藏品中，選擇15件為樣本，以候選的三種後設資料格式進行著錄。
 - 進行著錄後的適用性分析，包括觀察：記錄內容的豐富性，記錄呈現的結構性，生產力，以及編目品質等要項。
 - 評估結果：採用VRA Core Categories，並以此標準為基礎，增加若干元素。

後設資料的標準探討

標準的選擇方法 (Kent State Univ.)(續)

- 分析與評估方法 (2)
 - 以VRA Core Categories為基礎，發展後設資料需求規格書 (Data Dictionary)與資料輸入範本 (Template)，並將資料輸入。
 - 發展VRA與USMARC，DC之間的對照轉換工具，以備未來與其他單位進行交流之用。

後設資料的作業評估

標準的檢視與評估

- 使用環境
 - 使用者族群特性：如學科、對象
 - 資料類型：如書、視聽媒體、數位檔
 - 符合國際標準的程度與相容性
 - 多語文
 - 應用容易度：包括系統設計者、管理者、終端使用者

後設資料的作業評估

標準的檢視與評估

- 格式
 - 結構
 - 資料交換與對照
 - 文獻互動關係的建立與管理
 - 內容符合比率
- 通訊協定狀況
- 實際應用狀況

後設資料的作業評估

工具/系統的評估

- 交換能力 (Metadata Exchange)
 - 輸入能力 (Import capabilities)
 - 輸出能力 (Export capabilities)
 - 完整性與相容性 (Completeness / Compliance)

後設資料的作業評估

工具/系統的評估 (續)

- 可用性 (Usability)
 - 使用者介面 (User Interface)
 - 提示訊息 (User prompting)
 - 可避免重複輸入 (Minimizing duplication of entry)
 - 整合產生的資料 (Data Generation Integration)
 - 剪貼功能 (Cut and Paste)
 - 容許階段性作業 (Restart-ability)
 - 說明文件 (Documentation)

後設資料的作業評估

工具/系統的評估 (續)

- 管理功能 (Administrative)
 - 使用平臺及安裝 (Platforms / Installation)
 - 獨立運作及維護性 (Stand-alone)
 - 更新能力 (Updates)

後設資料的作業評估

工具/系統的評估 (續)

- 系統可靠性 (Tool Reliability)
 - 強韌性 (Robustness)
 - 復原機制 (Recovery mechanisms)
 - 一致性 (Consistency)

後設資料的互通機制

Metadata Interoperability

- 定義 (Interoperability) :
 - 不同系統的軟硬體能夠彼此溝通。
- 層次
 - 第一層：發現數位典藏資源 (Discovery)
 - 第二層：相似典藏內容間的互通 (Interoperability within or across similar content)
 - 第三層：不同典藏類型間的互通 (Interoperability across diverse content types)
- 方式：Federation，Harvesting
- 協定：Z39.5，OAI，Hybrid Model？

後設資料參考資源與閱讀清單

網站資源

- DIGITAL LIBRARIES: Metadata Resources (IFLA) <http://www.ifla.org/II/metadata.htm>
- UKOLN Metadata <http://www.ukoln.ac.uk/metadata/>
- Metadata and Resource Description <http://www.w3.org/Metadata/>
- CurrentCites “Metadata” Bibliography <http://sunsite.berkeley.edu/CurrentCites/bibondemand.cgi?title=Metadata&query=metadata>

後設資料參考資源與閱讀清單 (續)

- 電子書


- Baca, Murtha, ed. Introduction to Metadata: Pathways to Digital Information. Los Angeles: J. Paul Getty Trust, 1998.

<http://www.getty.edu/research/institute/standards/intro/metadata/>

- 研究報告

- Committee on Cataloging: Description and Access Task Force on Metadata: Final Report, June 2000

<http://www.ala.org/alcts/organization/ccs/ccda/tf-meta6.html>



問題與討論