

後設資料品質指標的發展與應用

陳亞寧 鍾豐謙
後設資料工作組 (MAAT)

30 Jul., 2009

1

聯合目錄資料範例：1974玫瑰花¹

1974 玫瑰花



後設資料 授權與引用資訊 連結到原始資料網站

資料類型：

原件；美術

著作者：

李石樵

主題與關鍵字：

玫瑰花／靜物／花瓶

寫實靜物

西方媒材／油彩

描述：

說明：在回返寫實後，李氏不再刻意追求形的變化，而以自然形象所涵蘊之結構為主，但有時並不明顯顯出經營痕跡。此作雖描繪簡單的靜物，但極寫實的表象卻奠基於幾何結構上，玫瑰花錯落有致，但自然形成一個圈，此外綠圓圈又為葉子所形成之內線所呼應，花瓶之渾圓就不言而喻了。背景之方塊嵌後關係蓄意為之，因此方塊與畫面左右不相連屬。桌面自然隱示方塊，而其上的方塊圖案則與桌面和背景之方塊呼應，在簡單中見複雜。

材質：畫布／油彩

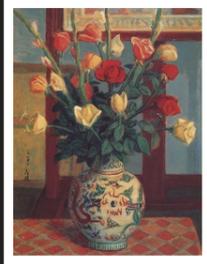
出版者：

藝術家出版社／台北縣立文化中心

(續下頁)

聯合目錄資料範例：1974玫瑰花²

1974 玫瑰花



日期：

創作日期：1974

格式：

典藏實體：1件；72.5 × 60.5 cm

關聯：

李石樵著，黃明政主編，林宗興、千音、黃玉淇攝影，《李石樵畫集》，1996，台北：台北縣立文化中心／李石樵繪；王德育著，《李石樵》，臺灣美術全集（8），1993，臺北市：藝術家出版社

範圍：

創作日期：1974

內容地點：台北

管理權：

財團法人李石樵美術館基金會

3

聯合目錄資料範例：魚型燈心袋

中文品名：魚型燈心袋

後設資料 授權與引用資訊 連結到原始資料網站

資料識別：

登錄號：88-00097

資料類型：

編織

格式：

縱：78

橫：22

寬：3

關聯：

朝代：清末民初

4

緣起

- 從90年起，在國科會的經費贊助下，有關數位化計畫的發展與成果已有一定的基礎
- 從數位典藏國家型科技計畫（NDAP）至數位典藏與數位學習國家型科技計畫（TELDAP）發展過程中，也必須從數量目標的追求轉為品質的提升，以利於未來數位內容產業的加值與再利用

目的

- 主要對象為後設資料記錄，執行記錄內容的品質確保與改善，主要的預計目的有：
 - 提供主題計畫進行自我評鑑
 - 符合國際後設資料標準的要求及其接軌，以利於資源發掘與交換共享
 - TELDAP年度成果評鑑的參考指標
 - 促進資料的再利用與商業加值



後設資料品質指標的依據

- 預期符合性 (conformance to expectations)
 - 來源性 (provenance)
 - 完整性 (completeness)
 - 正確性 (accuracy)
 - 取用性 (accessibility)
 - 一致性 (logical consistency and coherence)
 - 即時性 (timeliness)：包括新穎性 (currency) 與時間延遲性 (lag)
- Bruce, T.R. and Hillmann, D.I. (2004). The continuum of metadata quality: Defining, expressing and exploiting. In *Metadata in Practice* (pp. 238-256), edited by D.I. Hillmann and E.L. Westbrook (Chicago: American Library Association).

7

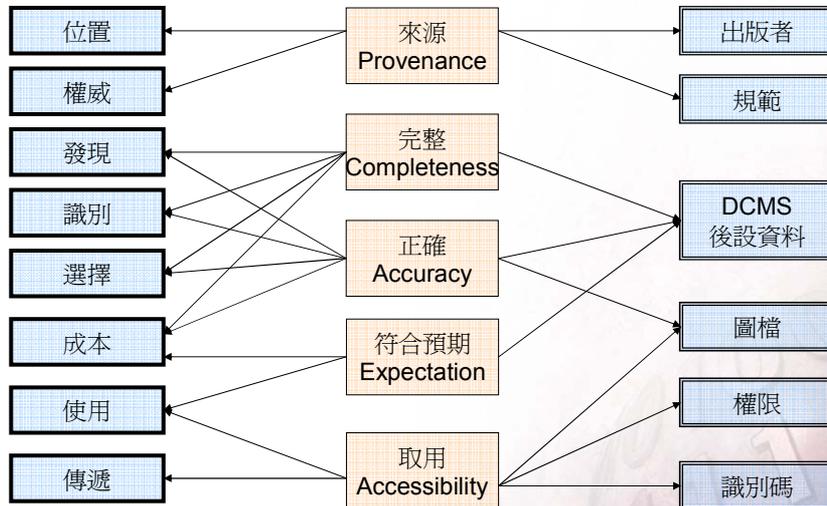


TELDAP計畫的後設資料品質指標

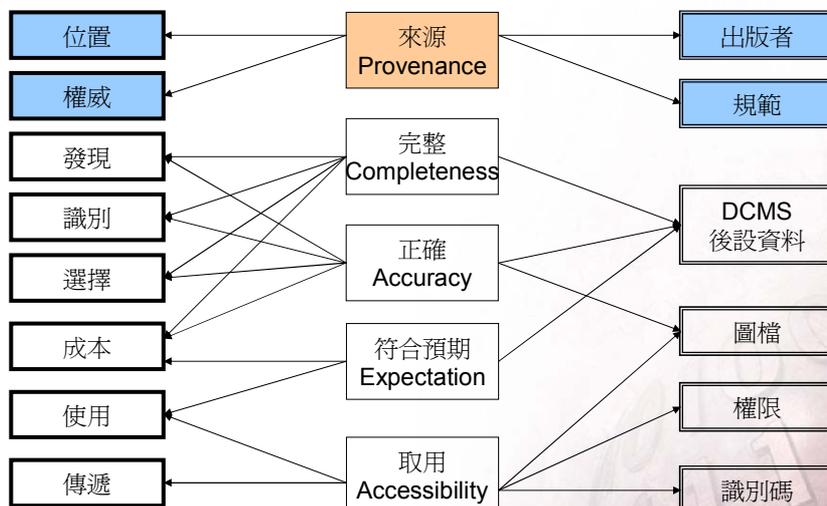
- 依據上述品質指標，進行在地化導入，主要差異有：
 - 進一步區分為資料欄位與結構、資料內容與值兩大類型，再配合Bruce and Hillmann (2004) 提出的品質指標，形成一份後設資料品質的指標矩陣
 - 實際前測後，以資料欄位與結構為主要標的，建議採用預期符合性、來源性、完整性、正確性與取用性等五大指標

8

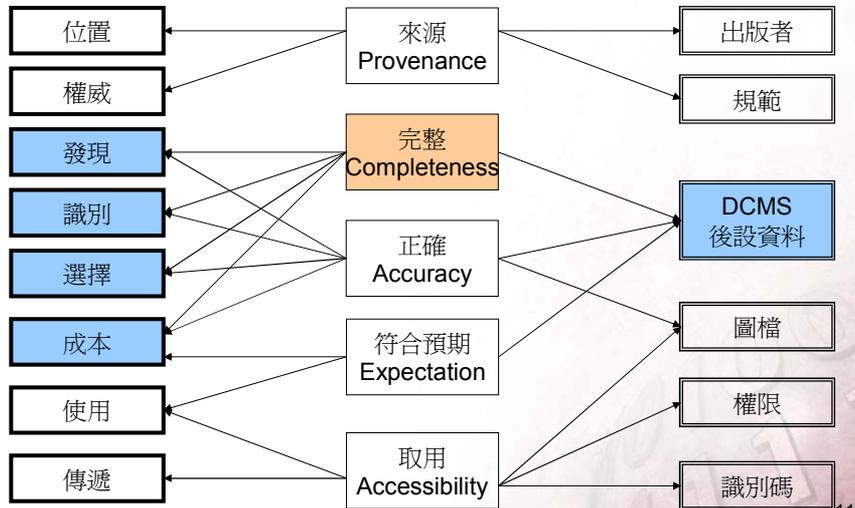
後設資料品質指標及其後設資料元素與功能關係圖



後設資料品質指標及其後設資料元素與功能關係圖

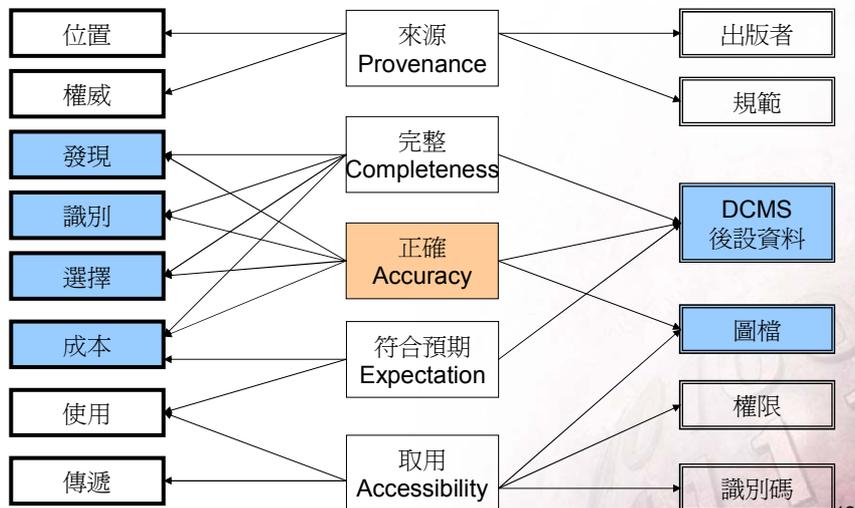


後設資料品質指標及其後設資料元素與功能關係圖



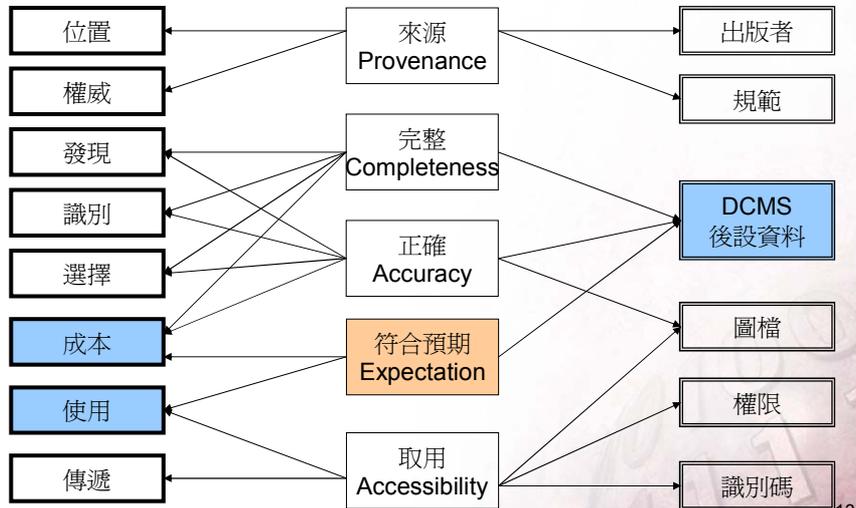
11

後設資料品質指標及其後設資料元素與功能關係圖



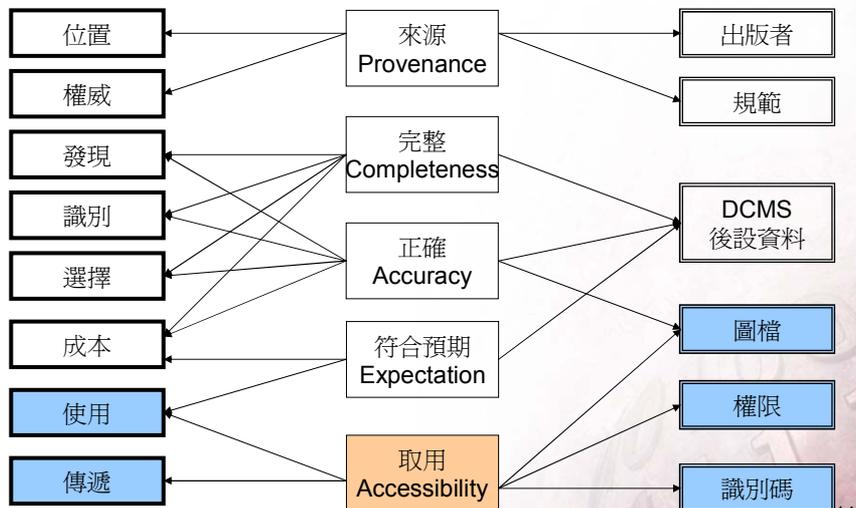
12

後設資料品質指標及其後設資料元素與功能關係圖



13

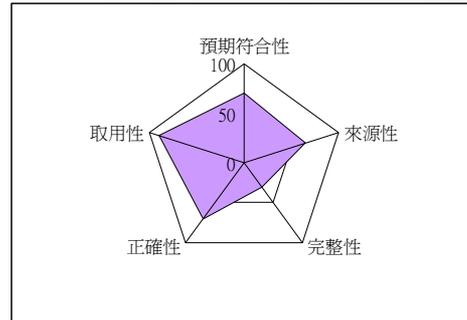
後設資料品質指標及其後設資料元素與功能關係圖



14

後設資料品質指標計算與結果

- 以權重、平均等方式計算各之指標之得分，並以雷達圖方式呈現
- 分數並非絕對，也非評鑑之唯一方式，僅呈現客觀的觀察結果，並提供未來改善之建議



後設資料的品質指標¹

指標	定義	算式	權重正規化
O.I.1. 藏品數量	計畫「當年度」承諾新增多少筆藏品XML	計畫當年度承諾匯入之藏品XML record總數	O類指標為概括性品質指標，為雷達圖第一部份。先以年度新增record總數與計畫書承諾達成的record數相除求取提交率（O.I.2.）得O部分的分數
O.I.2. 藏品匯入比例	計畫「當年度」確實新增匯入筆數佔承諾匯入總數之比例	年度確實新增record總數/計畫當年度承諾匯入之藏品XML record總數	

後設資料的品質指標²

指標	定義	算式	權重正規化
P.I.1. 提供單位資訊	Publisher或Rights欄位是否填值？	Publisher或Rights欄位非空白的紀錄總數 / 紀錄總數	P類指標為產生來源性品質指標，為雷達圖第二部份。四項指標（P.I.1, P.I.2, P.II.1, P.II.2）各佔25%，加總得P部分的分數
P.I.2. 元素著錄規範的採用	是否有採用著錄規範進行元素的著錄？如：CCO、AACR/RDA、ISAD（G）。或內部自行發展的著錄規範？	人工判斷	
P.II.1. 記錄元素資料值之著錄規範	是否採用著錄規範進行記錄元素資料值的著錄？如：AACR、DACS、CCO等。或內部自行發展的著錄規範？	人工判斷	
P.II.2. 控制詞彙、索引典或權威檔的採用	是否有使用控制詞彙、索引典或權威檔等現有標準，以著錄特定元素的內容？如：LCSH、LCNAF、AAT、TGN、ULAN等。或內部自行發展的控制詞彙、索引典或權威檔，乃至於代碼表？	人工判斷	

後設資料的品質指標³

指標	定義	算式	權重正規化
C.I.1. 元素使用完整性	Title = 計畫有使用Title的藏品record總數 / 計畫藏品record總數	DC十五個元素都計算一次	C類指標為資料完整性品質指標，為雷達圖第三部份。以C.I.1為基礎計算百分比得C部分的分數
A.I.1. 元素使用正確性	元素是否正確使用？亦即元素的實際使用情形是否符合每一個Element的定義？如：Format & Type、Subject & Description、Source & Relation等。	人工判斷	A類指標為資料正確性品質指標，為雷達圖第四部份。以A.I.1為基礎計算百分比得A部分的分數

後設資料的品質指標⁴

- 經由後設資料記錄的Identifier為依據，連回原系統的記錄，包括：

指標	定義	算式	權重
Ac.1.1. a. identifier存在比例	identifier總數÷記錄總數	計畫原始資料庫連結總數 / 計畫藏品總數	(後述)
b. identifier有效比例	identifier有效總數÷記錄總數	計畫原始資料庫有效連結總數 / 計畫藏品總數	
c. identifier相異比例	Identifier不重複總數 (扣除連結至同一頁面)÷記錄總數	(計畫原始資料庫有效連結總數 - 重複連結總數) / 計畫藏品總數	

後設資料的品質指標⁵

指標	定義	算式	權重
Ac.1.2. a. 數位檔存在比例	包含數位檔的紀錄總數÷記錄總數	計畫有提供數位檔連結的XML紀錄總數 / 計畫XML紀錄總數 (ex: 有提供圖檔連結的紀錄70筆/ 總共100筆)	
b. 數位檔有效比例	提供有效數位檔紀錄總數÷包含數位檔的紀錄總數	提供有效數位檔連結之紀錄總數 / 有提供數位檔連結的紀錄總數 (ex: 70筆共有65筆圖檔有效 / 70筆有提供圖檔URL)	
c. 數位檔相異比例	扣除多筆資料使用同一個數位檔連結： (有效數位檔紀錄總數 - 有效數位檔重複的筆數) / 有效數位檔紀錄總數	(計畫有效數位檔紀錄總數 - 有效數位檔連結重複總數) / 計畫數位檔有效連結總數 (ex: 65筆120個圖檔URL不重複 / 65筆130個圖檔URL有效)	
d. 有效數位檔紀錄數	包含有效數位檔的紀錄數	計畫有效數位檔的XML總數 (ex: 70筆記錄有140圖)	

後設資料的品質指標⁵

指標	定義	算式	權重
Ac.1.2' 數位檔顯示 a. 數位檔存在總數	數位檔連結總數	計畫提供的數位檔連結總數 (ex: 有提供70筆圖檔連結)	
b. 數位檔有效比例	數位檔連結有效的比例	數位檔連結有效總數 / 數位檔連結總數 (ex: 70筆圖檔共有65筆圖檔有效)	
c. 數位檔相異比例	扣除多筆連結其實指向同一個數位檔連結者	(有效數位檔連結總數 - 有效數位檔重複的連結總數) / 有效數位檔連結總數 (ex: 60筆圖檔URL不重複 / 65筆圖檔URL有效)	

21

後設資料的品質指標⁶

- **Ac**類指標為存取性品質指標，為雷達圖第五部份，由 identifier 與圖檔存取性兩類指標組成。
- Identifier 部分的指標計算方式為：先以 identifier 存在比例 (**Ac.1.1 a**) 乘以 identifier 有效比例 (**Ac.1.1 b**)，再乘以 identifier 相異比例 (**Ac.1.1 c**) 得有效又相異 identifier 之比例。
- 圖檔部分的指標計算方式為：先以圖檔存在比例 (**Ac.1.2 a**) 乘以圖檔有效比例 (**Ac.1.2 b**)，再乘以圖檔相異比例 (**Ac.1.2 c**) 得有效又相異圖檔之比例。可再視需要以有效圖檔紀錄數 (**Ac.1.2 d**) 結果對得分微調之。
- **Ac**類指標得分由 identifier 及圖檔兩部分之得分平均求得。

22

配套措施-指引

- 好的後設資料
 - 符合至少一套國際標準
 - 符合元素的定義
 - 著錄格式一致
 - 必備或基本欄位的著錄要求
 - 反映物件的意義
 - 註明提供單位
 - 提供正確的圖示與URL
 - 資料元素的設計與使用，必須達成資源的描述、發掘與利用

配套措施-對照服務

- 對照機制的建立與對照表的調整（MAAT與聯合目錄）
 - MAAT小組可提供有關標準的應用及對照等規範指引及其訓練推廣課程
 - 各計畫在匯入後設資料至聯合目錄前，由MAAT小組同仁以品質指標進行事前品質控制
 - 後設資料匯入後由聯合目錄團隊進行資料匯入的檢查及整體效益評估
- Registry and template的開發（MAAT）
 - 作為計畫相關文件的上載與保存，利於評鑑及未來再利用與加值應用的依據
http://metadata.teldap.tw/design/mapping/service_apply_mapping.php

配套措施-其他

- 權利項目的宣告與規範（拓展臺灣數位典藏計畫及相關團隊）
 - 標示現有數位圖檔的**使用限制條件**，同時也間接影響了使用者是否進一步點選**URL鏈結**功能連回主題計畫的原始資料庫或網站之意願，以及商業加值
- 匯入的必備元素與值（拓展臺灣數位典藏計畫）
 - 共包括**Title**、**Publisher**、**Subject**、**Format**等四個元素及記錄**ID**
- 數據蒐集與統計的軟體開發（聯合目錄）
 - 包括資料元素及資料內容與值等兩種類型

25

問題與討論

26