

後設資料發展趨勢與數位典藏應用之比較研究

A Comparative Study on Metadata Applications between the Global Trends and Developments of the National Digital Archives Program in Taiwan

陳 亞 寧

Ya-Ning Chen

中央研究院計算中心系統分析師兼組長

Systems Analyst

Computing Centre, Academia Sinica

E-mail : arthur@gate.sinica.edu.tw

【摘要 Abstract】

隨著數位圖書館計畫的推展與實施，後設資料 (Metadata) 已成為不可或缺的基礎研究與工作項目之一。本文採取比較法，以後設資料的全球發展趨勢為依據，探索數位典藏國家型科技計畫後設資料的發展與應用現況。本文發現主要有四項：第一，應用情形朝多元化方向發展，範圍同時兼具深廣度，並以結構性後設資料標準為主。第二，應用方式則是採取多種標準與平行策略。第三，在描述層面方面，以物 (即典藏品)、人與地理空間為主。第四，以學科與社群為導向的後設資料核心集有逐漸增加的現象。

Metadata has become one of the fundamentals for digital library projects worldwide in recent years. With reference to the global trends of metadata development, this paper adopts comparison as the research methodology to explore the current metadata application and progress attained by the National Digital Archives Program (NDAP) in Taiwan. Through this comparative study, four findings are in focus. Firstly, metadata applications by NDAP involve a rich diversity, including various communities, themes, data types, and data functions. In addition, the selected metadata standards by NDAP projects are primarily structured ones. Secondly, based on the strategy of parallel employment, multiple metadata standards are customized and interwoven for the applications of NDAP projects. Thirdly, metadata descriptions in NDAP projects are mainly focused on thing, person, and geo-spatial objects. Lastly, several common core sets of metadata, either on the thematic or on the community basis, have been gradually developed for metadata interoperability.

關鍵詞 Keyword

後設資料 數位圖書館 數位典藏國家型科技計畫

Metadata ; Digital library ; National Digital Archives Program in Taiwan (NDAP)

壹、前言

隨著全球各式數位圖書館計畫 (Digital library projects) 的進行，後設資料 (Metadata) 也成為上述計畫眾多重要的研究與工作項目之一，同時也在圖書館、博物館、檔案館與電腦科技各界引起重視與廣泛討論。本文旨在探討後設資料的現況應用與發展趨勢，同時檢視國內「數位典藏國家型科技計畫」(National Digital Archives Program in Taiwan, NDAP) 應用後設資料的實際情形。本文共分為五個部份，包括全球後設資料發展現況、數位典藏國家型科技計畫後設資料應用現況之分析、研究結果與討論、研究發現，以及結語與建議等。

貳、研究範圍與方法

本文主要採取比較法為主要的研究方法，以分析、比較國際上與 NDAP 在後設資料方面的應用發展。首先，全球後設資料發展現況方面，本文選取四份最具代表性的文獻進行研究，以探討與分析國際上後設資料的現況發展——英文文獻 3 篇，中文文獻 1 篇。在英文文獻方面，涵蓋英、美、加三國。其中，第 1 篇文獻係由英國圖書館網路辦公室 (UK Office for Library Networking, UKOLN) 提出，由於 UKOLN 共有 3 篇文獻屬同一主題，且先後發表，本文選取最近 1 篇作為研究對象；其餘 2 篇相關資訊如下：

- Heery, R. (1996). Review of metadata formats. Program, 30(4), pp. 345-373.
- Dempsey, L., Heery, R., Hamilton, M., Hiom, D., Knight, J., Koch, T., Peereboom, M., & Powell, A. (1997). Specification for resource description methods: Part 1: A review of metadata: A survey of current resource description formats. (URL: <http://www.ukoln.ac.uk/metadata/DESIRE/overview>

directory)

另外，本文也選取了美國學者 Gilliland-Swetland 與國內陳雪華教授所發表有關後設資料綜論性質的文獻，以及選擇加拿大 MetaMap 計畫 (MetaMap Project) 所彙整的後設資料知識地圖 (MetaMap)，以更完整與新穎的方式探索後設資料的全貌。這四篇文獻如下：

- Dempsey, L., & Heery, R. (1998). Metadata: A current view of practice and issues. Journal of Documentation, 54(2), 145-172.
- 陳雪華。(民 86)。網路資源組織與 Metadata 之發展。圖書館學刊，12，頁 19-37。
- Gilliland-Swetland, A. (2000). Introduction to metadata: Setting the stage. (URL:http://www.getty.edu/research/conducting_research/standards/intrometadata/pdf/swetland.pdf)
- Turner, J.M., Moal, V., & Desnoyers, J. (2004). MetaMap: Acronym index (Ver. 1.2). (URL: <http://mapageweb.umontreal.ca/turner/meta/nglish/metamap.html>)

次則，本文以中央研究院後設資料工作組 (Metadata Architecture and Application Team, MAAT) 所支援的 NDAP 主題計畫為範圍，並以完成後設資料系統規格書 (Metadata system specification) 或後設資料功能需求書 (Metadata functional requirements specification) 為限，如果同一計畫因需求或典藏品不同所產出的規格書，仍視為不同的規格書，但版本不同則取最新一版為主。另外，由於 MAAT 已為檔案主題小組完成核心元素集，雖無規格書，也一併列入，共計選取了 10 個機構 25 項主題計畫 54 份規格書，如表一所示。最後，本文以前述文獻與 NDAP 計畫的分析資料為基礎，進行交叉比對，以提出研究結果與發現。

表一：25 項 NDAP 主題計畫

機構名稱	計畫名稱	規格書份數
中央研究院	民族學研究所－臺灣原住民數位典藏計畫	5
	近代歷史研究所－檔案館外交經濟重要檔案數位典藏計畫	1
	歷史語言研究所－珍藏歷史文物數位典藏計畫	9
	動物學研究所－臺灣動物相典藏之研究：魚類與貝類	2
	植物學研究所－臺灣本土植物數位典藏計畫	1
臺灣省諮議會、中央研究院臺灣史研究所	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	2
故宮博物院	故宮文物數位典藏系統之開發與建立	4
歷史博物館	國家歷史文物數位典藏計畫	7
國史館	國史館典藏國家檔案與總統文物數位化中程計畫	4
國史館臺灣文獻館	典藏日據時期與光復初期史料數位化計畫	5
臺北藝術大學、中央研究院臺灣史研究所	臺灣視覺記憶數位典藏計畫（前身：臺灣老照片數位博物館計畫）	1
臺北藝術大學	臺灣社會人文電子影音數位博物館	1
	電子影音資料庫加值應用之研究	1
真理大學	「馬偕與牛津學堂」數位典藏計畫	1
臺灣史前文化博物館	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	1
文化大學	近現代珍貴數位典藏系統之研製	1
屏東科技大學	臺灣傳統農耕機具資訊檢索及網路展示設計	1
	數位典藏聯合目錄建置	1
	動物主題小組	1
	植物主題小組	1
	動植物聯合主題小組	1
	檔案主題小組	1
	書畫主題小組	1
	善本古籍主題小組	1

參、全球後設資料發展現況：文獻分析

一、Dempsey, L., & Heery, R. (1998). Metadata: A current view of practice and issues. *Journal of Documentation*, 54(2), 145-172.

此文獻係英國 UKOLN 執行一項研究計畫的調查結果為基礎，將其研究成果與發現予以彙整。此份文獻將現有後設資料依下列六種觀點進行調查、分析與歸納，分述如下：

- (一) 使用環境 (Environment of use)：後設資料所描述的服務 (Services)。
- (二) 功能需求 (Function of record)：如找尋 (Location)、選擇 (Selection)、評鑑

(Evaluation)、分析 (Analysis) 等。

- (三) 產生方式 (Creation method)：如人工或自動。
- (四) 結構 (Complexity of designation)：欄位或元素的複雜度，如是否有次欄位 (Sub-fields) 或限定詞 (Qualifiers)。
- (五) 相關的查詢協定 (Associated search protocols)：是否有搭配其他相關的查詢檢索協定，如 Z39.50。
- (六) 狀態 (Status)：是否為專屬 (Proprietary) 或特定領域 (Domain)。

依上述六種觀點，Dempsey 與 Heery 兩位學者將後設資料區分為三種類型 (Band)，如表二所示：

表二：Dempsey 與 Heery 兩位學者所歸納的三種後設資料類型

	類 型 一	類 型 二	類 型 三
使用環境	全球網際網路蒐尋服務、全球資訊網索引服務	選擇式的網際網路蒐尋服務、指引 (Directory) 服務	學術資料及重要典藏庫的描述
功能需求	找尋	發掘、找尋與選擇	找尋、選擇、評鑑與分析
產生方式	程式自動抓取	程式自動抓取暨人工輸入	人工方式，通常需要具備專業知識，以及相關的資訊人員
結構	無結構化	有限的結構	結構化，乃至標誌 (Mark up)
相關的查詢協定	界面協定 (如：HTTP 與 CGI)	界面協定 (如：HTTP 與 CGI)、指引服務協定 (如：WHOIS++, LDAP) 與查詢 (如：Common Indexing Protocol)	Z39.50、SGML 瀏覽器與查詢
狀態	專屬	網際網路標準	特殊領域標準

(資料來源：Dempsey and Heery, 1998, p.157)

如果進一步深究上述三種類型，依據 Dempsey et al. (1997) 等人的歸納，第一種類型後設資料是一種簡易格式 (Simple formats)，通常進行全文式索引 (Full text indexing)，如 Lycos、Altavista 與 Yahoo 等。第二種則是結構化格式 (Structured formats)，屬於一種最近興起的標準 (Emerging standards)，具備欄位結構 (Field structure)，如 DC、IAFA Templates、RFC 1807、SOIF 與 LDIF 等。第三種則是結構豐富的格式 (Rich formats)，屬於一種現有的國際標準，通常格式內容進行非常詳盡的標記 (Elaborate tagging)，如 ICPSR、CIMI、EAD、TEI 與 MARC 等。此外，Dempsey et al. (1997) 等人也將後設資料的功能劃分為：找尋、發掘、記載 (Documentation)、評鑑 (Evaluation) 與選擇 (Selection) 等五項。

二、陳雪華。(民 86)。網路資源組織與 Metadata 之發展。圖書館學刊，12，頁 19-37。

陳教授以前述「Specification for resource description methods: Part 1: A review of metadata: A survey of current resource description formats.」為基礎進行分析研究，將後設資料區分為早已普遍使用的 metadata 格式、描述科技文獻之 metadata 格式、描述人文及社會科學資源之 metadata 格式、描述政府資訊之 metadata 格式、描述地理空間性資源之 metadata 格式、描述博物館藏品與檔案特藏之 metadata 格式、描述大量網路資源之 metadata 格式與其他 (請參見表三)。此外，陳教授也特別指出上述這份報告未將 CMARC、Handle System 與 US Environmental Protection Agency (EPA): Scientific metadata Standards Project 等列入。

表三：陳雪華教授所歸納的後設資料種類

Metadata 種類	Metadata 格式
早已普遍使用的 metadata 格式	MARC(包括 USMARC, UKMARC 與 UNIMARC), PICA
描述科技文獻之 metadata 格式	BibTex, EELS, EEVL, RFC 1807
描述人文及社會科學資源之 metadata 格式	ICPSR SGML Codebook Initiative, TEI Headers
描述政府資訊之 metadata 格式	GILS
描述地理空間性資源之 metadata 格式	CSDGM/FGDC
描述博物館藏品與檔案特藏之 metadata 格式	CDWA, CIMI, EAD
描述大量網路資源之 metadata 格式	DC, IAFA/WHOIS++ Templates, LDIF, SOIF, URCS
其他	Warwick Framework

(資料來源：陳雪華，民 86，頁 23-28)

三、Gilliland-Swetland, A. (2000). Introduction to metadata: Setting the stage.

此份文獻雖是導論性質的文章，然而對後設資料卻有極為宏觀的論述。此份文獻試從圖書館、檔

案館、博物館與電腦科技界等不同領域去闡述後設資料，其中內容將後設資料類型與功能 (Type and function)，以及屬性與特質 (Attributes and Characteristics) 加以區分。在後設資料類型方面，Gilliland-Swetland 劃分為管理 (Administrative)、描

述 (Descriptive)、長期保存 (Preservation)、技術 (Technical) 與使用 (Use) 等五種 (詳表四：後設資料的類型與功能)。在後設資料屬性與特質方面，則有資料來源 (Source)、產生方式 (Method of metadata creation)、性質 (Nature)、狀態 (Status)、結構 (Structure)、語意 (Semantic) 與層次 (Level) (詳表五：後設資料的屬性與特質)。此外，Gilliland-Swetland (2000, p. 12) 也以數位資訊物

件的生命週期 (Life cycle of objects contained in a digital information) 為依據，提出後設資料在不同層面 (Layers) 所扮演的功能；這些層面包括產生 (Creation and multi-versioning)、組織 (Organization)、查詢與檢索 (Searching and retrieval)、使用 (Utilization)、長期保存與淘汰 (Preservation and disposition)。

表四：後設資料的類型與功能

類型	定義	範例
管理型	應用在資訊資源的管理與行政方面	<input type="checkbox"/> 徵集資訊 <input type="checkbox"/> 智財權與複製的追蹤 <input type="checkbox"/> 法律取用方面的記載需求 <input type="checkbox"/> 所在位置的資訊 <input type="checkbox"/> 數位化的選擇要點 <input checked="" type="checkbox"/> 版本控制及差異 <input checked="" type="checkbox"/> 文書保存系統的稽查
描述型	應用在資訊資源的描述與辨識	<input type="checkbox"/> 編目記錄 <input type="checkbox"/> 查詢指引 <input type="checkbox"/> 特殊索引 <input type="checkbox"/> 資源的鏈結關係 <input type="checkbox"/> 使用者的註解 <input type="checkbox"/> 文書保存系統的後設資料
保存型	有關資訊資源的保存管理	<input type="checkbox"/> 實體資源情形的記載 <input type="checkbox"/> 採取保存實體與數位資源的行動記載，例如：儲存媒體的更換、電腦軟硬體的轉置
技術型	有關於系統功能如何運作與後設資料的運作	<input type="checkbox"/> 有關於電腦軟硬體方面的記載 <input type="checkbox"/> 數位化相關資訊，例如：檔案格式、壓縮比例、資料排程 (scaling routines) <input type="checkbox"/> 系統反應時間的追查 <input type="checkbox"/> 資料的認證與安全性，例如：加密、密碼
使用型	有關於資訊資源的使用層次與類型	<input type="checkbox"/> 展覽記錄 <input type="checkbox"/> 使用與使用者追蹤 <input type="checkbox"/> 內容再利用與多重版本的資訊

(資料來源：Gilliland-Swetland, 2000, p. 12)

表五：後設資料的屬性與特質

屬性	特質	範例
後設資料的來源	內部後設資料，通常由產生代理人在第一次創造或數位化時所產生。	<input type="checkbox"/> 檔案名稱與表頭資訊 <input type="checkbox"/> 目錄架構 <input type="checkbox"/> 檔案系統與壓縮綱要
	外部後設資料，通常不是由原創者提供，而是經由一段時間後，由相關人員或單位提供	<input type="checkbox"/> 註冊與編目記錄 <input type="checkbox"/> 智財權及其他法律相關資訊
後設資料的產生方式	由電腦自動產生	<input type="checkbox"/> 關鍵詞索引 <input type="checkbox"/> 使用者異動檔
	由人工產生	<input type="checkbox"/> 描述性的替代品，例如：編目記錄、都柏林核心集後設資料
後設資料的性質	後設資料是由資訊物件原創者所產生，而不是由學科或資訊專家產生。	<input type="checkbox"/> 個人網頁上的 Metatags <input type="checkbox"/> 個人檔案系統
	後設資料是由學科或資訊專家所產生，並非由資訊物件原創者產生。	<input type="checkbox"/> 專門的學科標題 <input type="checkbox"/> 機讀編目格式記錄 <input type="checkbox"/> 檔案查詢清單
狀態	靜態式後設資料，一旦產生後，即不再變動。	<input type="checkbox"/> 有關資訊物件的題名、來源與創造日期。
	變動式後設資料，會依使用與處理的狀況而改變。	<input type="checkbox"/> 目錄結構 <input type="checkbox"/> 使用者異動檔 <input type="checkbox"/> 影像解析度
	長期性後設資料，基本上是確保資訊物件可以持續取用與使用。	<input type="checkbox"/> 技術性檔案格式及其處理資訊 <input type="checkbox"/> 智財權資訊 <input type="checkbox"/> 保存管理方面的記載
	短期性後設資料，主要是一種異動的性質。	
結構	結構化後設資料，通常符合某一套標準化或未標準化的結構。	<input type="checkbox"/> 機讀編目格式 <input type="checkbox"/> 文獻編碼(TEI)與檔案描述編碼格式(EAD) <input type="checkbox"/> 自行發展的資料庫格式

(續下表)

(接上表)

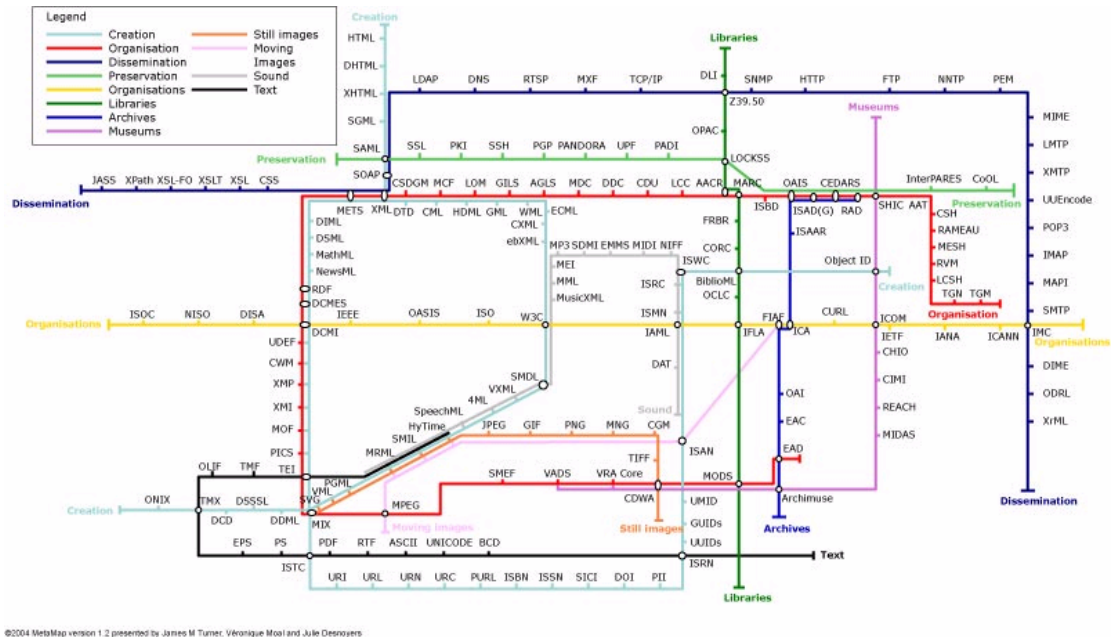
屬性	特質	範例
	未結構化後設資料	<input type="checkbox"/> 未結構化的備註欄位與註解
語意	控制型後設資料，通常遵循某一標準化詞彙或規範標目。	<input type="checkbox"/> 藝術暨建築索引典(AAT) <input type="checkbox"/> 藝術家姓名聯合清單(ULAN) <input type="checkbox"/> 英美編目規則第二版(AACR2) <input type="checkbox"/> 文字型的備註 <input type="checkbox"/> HTML Metatags
	非控制型後設資料，通常不遵循某一標準化詞彙或規範標目。	
層次	館藏合集式後設資料，有關於資訊物件形成的館藏。	<input type="checkbox"/> 館藏合集層級的記錄，例如：機讀編目格式的記錄或查詢清單。 <input type="checkbox"/> 專門性的索引
	物件後設資料，有關單一資訊物件，通常屬於包含在館藏合集之內。	<input type="checkbox"/> 抄錄的影像標題與日期 <input type="checkbox"/> 檔案格式資訊

(資料來源：Gilliland-Swetland, 2000, pp. 6-7)

四、Turner, J.M., Moal, V., & Desnoyers, J. (2004). MetaMap: Acronym index (Ver. 1.2).

MetaMap 計畫嘗試蒐集全球有關後設資料方面的標準、格式與開創活動 (Metadata standards, sets and initiatives, MSSIs)，並以圖示方式 (詳圖一：MetaMap 知識圖) 展現後設資料空間 (Metaspace)。整體而言，MetaMap 涵蓋三種後設資料的相關資訊：包括資訊管理的程序、(Processes of information management)、機構組織 (Institutions with expertise in managing information) 與資訊檔案類型 (Types of information files) (MetaMap, 2003)。如果以另外一種觀點切入，分析 MetaMap 知識圖所涵蓋的內容，約可區分為三個主軸，說明如下：

- (一) 功能：產生 (Creation)、組織 (Organisation)、傳佈 (Dissemination) 與長期保存 (Preservation) 等，即原來 MetaMap 的資訊管理的程序，以平行線表示。
- (二) 檔案類型：如圖片 (Still images)、動畫 (Moving images)、聲音 (Sound)、文字 (Text)，即原來 MetaMap 的資訊檔案類型。
- (三) 機構社群：包括前述 MSSIs 的專業機構，橫跨圖書館、博物館、檔案館與電腦科技等社群，如：ISOC、NISO、DISA、DCMI、IEEE、OASIS、ISO、W3C、IAML、IFLA、FIAF、ICA、CURL、ICOM、IANA、ICANN、IMC 等；即原來 MetaMap 的機構組織，以直線表示。



圖一：MetaMap 知識圖

(資料來源：URL: <http://mapageweb.umontreal.ca/turner/meta/english/printmapjpeg.html>)

經由上述探討，本文分析、歸納出目前後設資料的全球發展趨勢共有 12 項的重點方向，如下：

- (一) 功能 (Functionality)：如資源發掘 (Resource discovery)、描述 (Description)、管理、技術、使用、長期保存與著作權權限管理 (如 ODRL, XrML)。
- (二) 領域 (Domain)：除了數位典藏 (Digital library) 外，數位學習 (如 LOM)、電子化政府 (如 GILS)、科學 (如 MathML) 與電子商務 (如 ECML, ebXML) 等領域的應用與發展正逐漸興起。
- (三) 社群 (Community)：以圖書館、博物館、檔案館、電腦科技與網路上虛擬自治社群 (如 W3C) 為主。
- (四) 產生方式：人工、自動與半自動。
- (五) 描述深度 (Level)：除了原有的資料結構 (Data structure, 如 CDWA、CSDGM、EAD、METS、

- MODS) 外，有往向下或更深層次的發展傾向，包括資料內容 (data content, 如 AACR、CCO、ISAD、ISBD、RAD)、資料值 (data value, 如 AAT、DDC、LCSH、MESH)。
- (六) 描述層次：除了有單一館藏物件 (Item) 外，還有以館藏合集與全文為主的方式 (如 TEI、NewsML) 等。
- (七) 資源識別碼 (Identifier)：如 DOI、GUID、ISAN、ISBN、ISSN、MIX、PII、PURL、SICI、UMID、URC、URI、URL、URN、UUID 等。
- (八) 檔案格式：包括靜態與動態兩種，如：GIF、HyTime、JPEG、MNG、MPEG、PNG、SMIL、SVG、TIFF、VML 等。
- (九) 概念模式 (Conceptual model)：利用概念模式來增益現有後設資料標準的應用層面與深度，包括了圖書館界的 FRBR、博物館界的 CIDOC CRM、檔案館界的 Records Continuum

Model、數位資訊長期保存方面的 OAIS Reference Model。

- (十) 資源交換與互通 (Interchange and interoperation)：後設資料除了注重在資源的發掘與描述外，如何應用著錄規則、控制詞彙與 RDF，達成資料品質的一致性、深度化語意描述與標引，乃至於如何在網際網路 (Internet) 上宣告所屬領域空間 (Name space)，促成數位資源網路化、互通化 (Interoperable)、共享化 (Sharing) 與結構化 (Structured)，則是近年來後設資料發展重點逐漸從描述與發掘朝向語意化邁進。無可諱言的是，XML 已成爲資源交換的語言，後設資料如經由 XML 的包裝與規範化，極易達成資源後設資料結構的明確定義與互通有無。
- (十一) 檢索與界面協定 (Protocol)：鑑於各式後設資源標準描述深度的詳簡不一，在 OAI 的推動下，OAI-PMH 協定逐漸成爲數位資源的共通查詢協定，以利網路上的檢索與傳佈；其他還有 SOAP、LDAP 等，以利各系統資料庫間的互通。另外在界面協定方面，以規範全球資訊網網頁的呈現方式，如：CSS、XSL、XSLT。
- (十二) 框架 (Framework)：在 Warwick Framework 概念推出後，W3C 也受到影響，並制訂出 RDF，作爲容納與共享各式後設資料標準的基礎。

肆、數位典藏國家型科技計畫後設資料應用現況之分析

本文以 MAAT 協助 NDAP 主題計畫爲範圍，並以完成後設資料系統規格書或功能需求書初稿爲限制，共計橫跨 10 個機構的 25 項主題計畫 54 份規格書符合研究要求 (詳表六：NDAP 主題計畫所採用後設資料標準一覽)，而列爲研究對象。以

這些主題計畫爲基礎，進行後設資料的應用現況分析，可以發現下列重點：

一、採用的標準類型

約略可以歸納爲十種類型，包括藝術博物館類 (如：CDWA、CIDOC Ethno Core Data Standards、VRA Core Categories 等)、檔案館類 (如：EAD)、圖書館類 (如：MARC21)、生物類 (如：HISPID、Species 2000、Darwin Core)、空間資訊類 (如：CSDGM)、電子商務 (如：ECML、XrML)。然而，也有些計畫採用導入概念模式 (如：FRBR-based Metadata，即 ECHO)，以深化後設資料標準的應用。此外，因爲數位典藏聯合目錄建置的要求，各主題計畫必須同步採用 OAI 與 DC 兩項標準，以達成互通有無的彙整目的；只是各主題計畫實施程度與時程有所不同而已。最後，爲了因應不同資料庫系統間的資料交換，也採取 XML DTD 的方式進行，以利後設資料記錄的滙入與滙出 (Import and export)。

二、權威控制

除了針對藏品進行描述外，可以發現某些計畫致力於權威控制，並且以人名居多，所採用標準爲 MARC21-Authority Format，計畫範圍包括圖書館、博物館與檔案館。次則，也有博物館類計畫進行藝術類的詞彙控制，而嘗試採用 AAT 爲基礎加以規範。三是，在生物類方面，則採用了 Species 2000 作爲物種名錄的標引與控制。最後，由於有一項計畫 (即：時空座標-歷史地圖運用的推動計畫) 致力於地理與空間資訊方面的研發，所以在地名方面由其集中發展，而各主題計畫則是經由地理空間系統 (Geo-spatial information system, GIS) 的結合，以利用、展現相關的地名與地理空間資訊，所採用標準則是 CSDGM 與 ADL GCS。

三、選擇與應用

就 25 項主題計畫而言，除了配合聯目錄必須採用的 DC 與 OAI 外，在資料結構方面的標準，以採用 CDWA、EAD 為最多；在資料值（即控制詞彙）方面，則以 MARC21-Authority Format 與 Species 2000 居多。從表六的內容而言，大多數

NDAP 計畫是採用前述文獻分析一節中所提及的結構化標準為主，因典藏品內容具備豐富的結構，因而訴求重點也集中在描述與組織。此外，NDAP 主題計畫在選擇標準時，焦點聚集在學科、社群與資料類型三大主軸。整體而言，NDAP 各主題計畫在選用標準時，首先著重在深化的學科知識組織，同時兼顧社群與資料類型等多重取向的描述。

表六：NDAP 主題計畫所採用後設資料標準一覽

機構名稱	計畫名稱	採用的後設資料標準
中央研究院	民族學研究所－臺灣原住民數位典藏計畫	CIDOC Ethno Core Data Standards、CDWA、MARC21-Bibliographic and Authority Format
	近代歷史研究所－檔案館外交經濟重要檔案數位典藏計畫	EAD
	歷史語言研究所－珍藏歷史文物數位典藏計畫	CDWA、EAD、MARC21-Authority Format
	動物學研究所－臺灣動物相典藏之研究：魚類與貝類	SPECIES 2000
	植物學研究所－臺灣本土植物數位典藏計畫	HISPID
臺灣省諮議會、中央研究院臺灣史研究所	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	EAD
故宮博物院	故宮文物數位典藏系統之開發與建立	AAT、CDWA、MARC21-Authority Format
歷史博物館	國家歷史文物數位典藏計畫	CDWA
國史館	國史館典藏國家檔案與總統文物數位化中程計畫	EAD、MARC21-Authority Format
國史館臺灣文獻館	典藏日據時期與光復初期史料數位化計畫	EAD、MARC21-Authority Format
臺北藝術大學、中央研究院臺灣史研究所	臺灣視覺記憶數位典藏計畫（前身：臺灣老照片數位博物	VRA Core Categories、FRBR

（續下表）

(接上表)

機 構 名 稱	計 畫 名 稱	採用的後設資料標準
	館計畫)	
臺北藝術大學	臺灣社會人文電子影音數位博物館	ECHO/FRBR
	電子影音資料庫加值應用之研究	ECML、XrML
真理大學	「馬偕與牛津學堂」數位典藏計畫	MARC21-Authority Format
臺灣史前文化博物館	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	CDWA
文化大學	近現代珍貴數位典藏系統之研製	CDWA
屏東科技大學	臺灣傳統農耕機具資訊檢索及網路展示設計	CDWA
	數位典藏聯合目錄建置	DC
	動物主題小組	SPECIES 2000
	植物主題小組	HISPID
	動植物聯合主題小組	Darwin Core
	檔案主題小組	EAD
	書畫主題小組	CDWA
	善本古籍主題小組	MARC21-Bibliographic Format

伍、研究結果與討論

一、功能訴求：結構化描述與組織為主

由於 NDAP 主題計畫是朝深化的學科知識組織及社群與資料的描述為主，因而訴求重點是朝結構化描述的方向發展，進而也可以從表六所列的後設資料標準得知。然而，隨著聯合目錄建置計畫的推動，NDAP 主題計畫也必須配合提供符合 DC 與 OAI 要求的功能需求。此外，從 25 項主題計畫的規格書內容中，也發現具備其他類型的後設資料需

求，包括管理（如：徵集資訊）、技術（如：電腦軟硬體環境、檔案格式）與使用（如：展覽記錄、後設資料記錄的使用與追蹤）等。整體而言，NDAP 主題計畫是以深化描述與通用式發掘並重，將後設資料標準的主要功能定位在數位資訊物件生命週期中的組織、查詢與檢索及使用等三個層面。

二、領域與社群：朝多元化發展

以 NDAP 主題計畫屬性而言，泛屬在數位典藏大領域之下，實際上包括數位圖書館、數位博物

館、數位檔案館與數位標本館等四大領域，及其所屬的社群。就學科而言，則是涵蓋了藝術、人文、社會與生物多樣性等。進而以此為基礎，擴展至電子商務，並未涉及數位學習、電子化政府。

三、數位檔案與產生方式：以人工方式為主

目前 NDAP 各項主題計畫已邁入自動化數位掃描作業與程序的階段，但在後設資料的建立方面則停留在人工輸入的作業階段。其實，國際上在進行典藏品數位掃描及其檔案時，已有相關的標準與規範；如：MPEG、JPEG、GIF 等。如何導入適當的電腦科技，從數位檔案中直接萃取基本的資料（如：前述的技術型與使用型後設資料）至後設資料系統中，加速原有作業速度，並節省更多的人力，以利投入至典藏品的內容詮釋與知識組織，則是未來可以考慮列為重要發展的一項核心技術。

四、描述深度：以資料結構與資料值為主

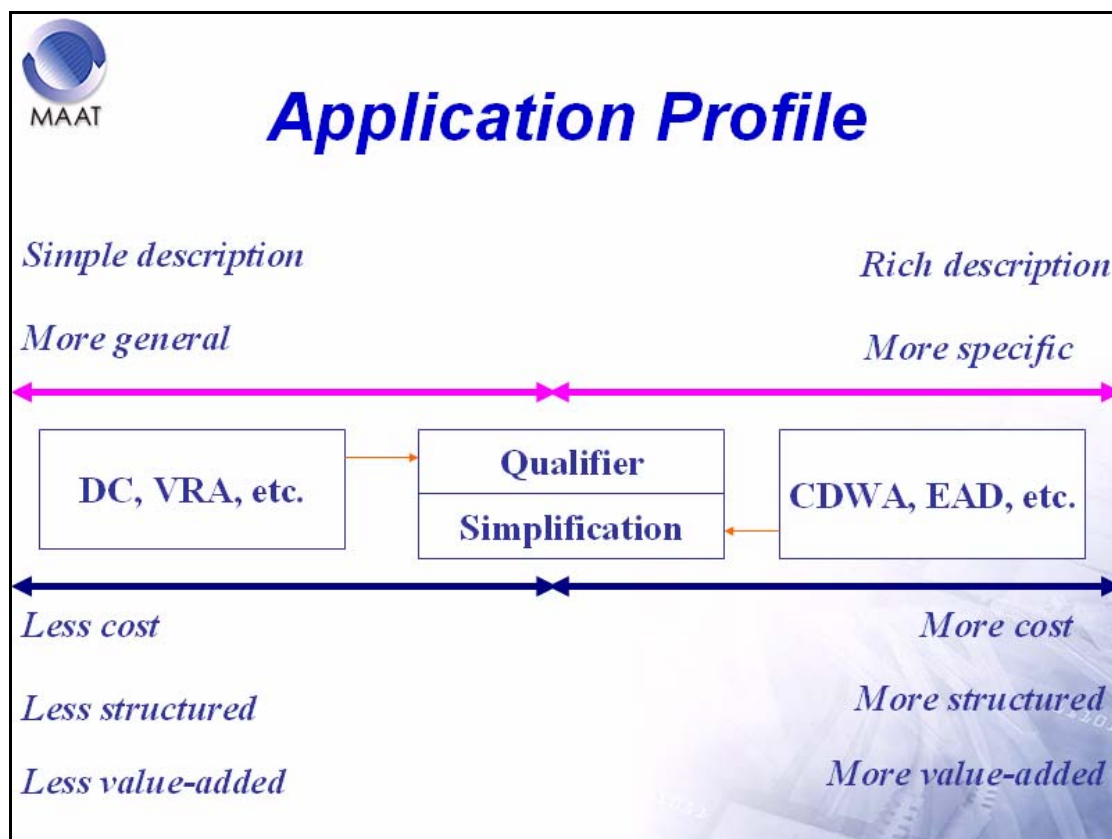
以本文選取 25 項 NDAP 主題計畫的發展現況而言，在描述深度方面集中在資料結構，同時涵蓋一些資料值的標準（如 AAT、MARC21-Authority Format 與 Species 2000 等），並未涉及資料內容方面的標準。以圖書館界而言，由於發展較早，從資料結構、資料內容與資料值等三方面皆已制訂出相關的標準，同時也累積一定的理論、規範與實務經驗；如：MARC21、AACR2、LC Authority File 等。如果在 NDAP 的博物館與檔案館類計畫方面，乃至於標本館類計畫，也許可以自圖書館界汲取相關經驗，以節省相關的人物力成本。換言之，未來的 NDAP 主題計畫可以朝資料內容與資料值的方面發展，以深化資料內涵及達到品質一致與互通等多重目的。

五、描述層次：以典藏品為主

在典藏品層次方面，除了針對某一典藏品（Item）及其內容（Full text）進行標誌外，自 1999 年起，英國 UKOLN 進行一項有關館藏合集層次的調查研究，並完成「A review of existing practice: An eLib supporting study」（URL: <http://www.ukoln.ac.uk/metadata/cld/study/>）研究報告，自此相關後設資料機構組織也致力研究館藏合集層次的後設資料標準。換言之，以描述層次觀點而言，包括了館藏合集、典藏品、典藏品內容全文等三個層次。就 NDAP 各主題計畫現況而言，大多數集中在典藏品此一層次，少數計畫致力於內容全文此一層次，並以數位方式掃描為數位檔案（如前述的 GIF）方式達成，尚未使用 TEI 此一標準進行內容標誌；在館藏合集描述方面則尚未進行。

六、其他：XML Schema、應用特徵檔與 MDRs

為了兼顧國際化與本土化的雙向不同需求，以及避免發展一套全新的後設資料標準，適當客製化（Customization）是必要的。因應此種需求，在 2000 年英國 UKOLN 提出應用特徵檔（Application profile）的觀念，以利於現有各式後設資料標準的應用（Heery and Patel, 2000）。換言之，將簡單格式複雜化，或是將複雜格式予以簡化（詳圖二：應用特徵檔的應用圖解）。此外，為能促進資訊共享與互通，XML 的發展也由 DTD 轉為 Schema；而後設資料註冊中心（Metadata registries, MDRs）也成為各式後設資料標準的註冊與登記集中地，以加速後設資料標準的應用與互通。就 NDAP 而言，基本上前述 25 項主題計畫皆採取應用特徵檔的方式應用現有的各式後設資料標準，同時也正在進行 MDRs 系統雛型，而 XML 則採用 DTD 的方式。



圖二：應用特徵檔的應用圖解

依據本文前述的探討，將全球與 25 項 NDAP 主題 計畫有關後設資料的發展與應用整理成下列表格。

表七：全球與 NDAP 後設資料應用發展之比較

	全 球	NDAP
功能	發掘、描述、管理、技術、使用、長期保存與著作權管理	以發掘與描述為主，其餘為輔
領域	數位典藏、數位學習、電子化政府、科學、電子商務	數位典藏為主，另有一項計畫已跨入電子商務
社群	圖書館、博物館、檔案館、電腦科技、網路虛擬自治社群	圖書館、博物館、檔案館與標本館
產生方式	自動、人工與半自動	人工
描述深度	館藏合集、典藏品、內容全文	典藏品與內容全文
描述層次	資料結構、資料內容、資料值	以資料結構與資料值為主

(續下表)

(接上表)

資源識別碼	數位與傳統	傳統
檔案格式	靜態與動態	靜態與動態
概念模式	FRBR、CIDOC CRM、Records Continuum Model、OAIS Reference Model、RDF Model 等	FRBR
資源交換與互通	XML DTD & Schema	XML DTD
檢索與界面協定	OAI、SOAP、LDAP；CSS、XSL、XSLT 等	OAI
框架	RDF	無
應用特徵檔	▲	▲
註冊與登記	▲	▲

陸、研究發現

如果將全球與 NDAP 間之後設資料應用發展作一比較，NDAP 主要特色在於多元化，主要研究發現有四項，如下：

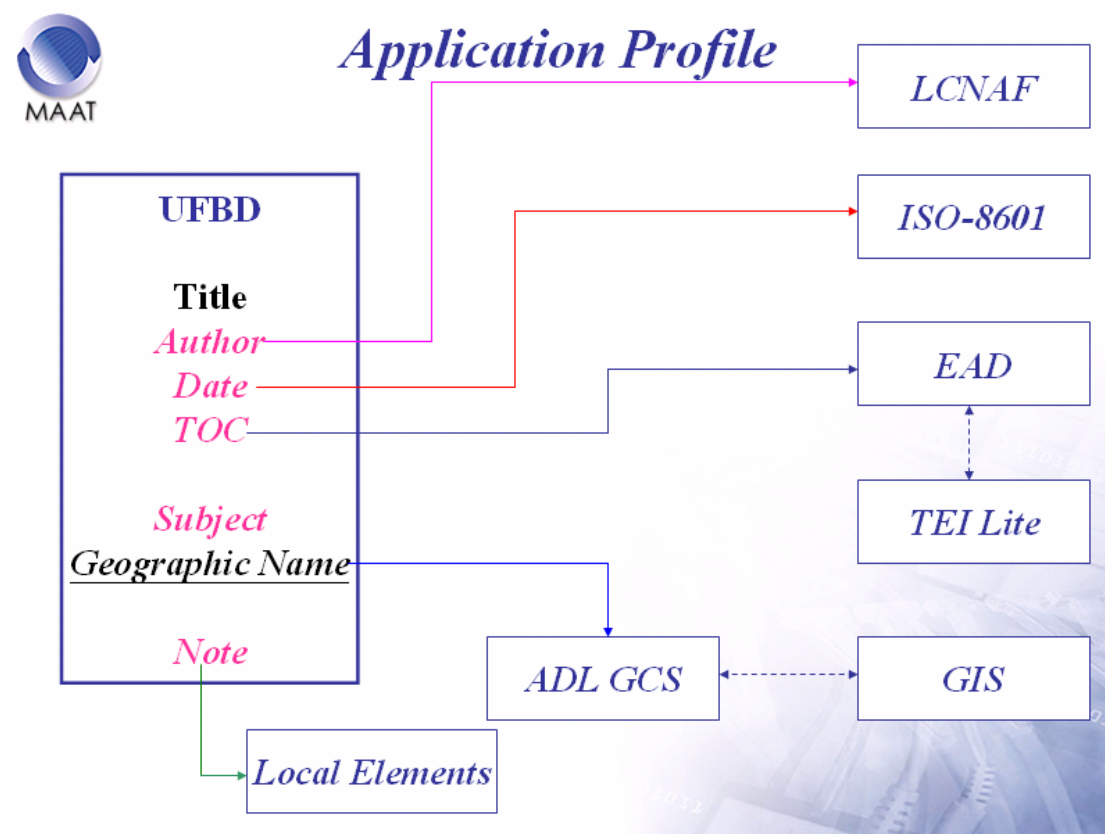
一、應用情形：多元化、精緻化，並以結構性後設資料為主

整體而言，本文研究的 NDAP 主題計畫內容反映出多元化的特質，涵蓋不同的社群（如：圖書館、博物館、檔案館與標本館）、學科（包括藝術、人文、社會與生物多樣性等）、資料類型（如：拓片、善本、照片、影音、書畫、田野調查、古文契書等）與功能（包括發掘、描述、技術、管理、使用等），傾向朝精緻與深化的學科知識組織的方向發展，因而在後設資料的應用與發展上，則是以採用結構性後設資料，以展現內涵與關係豐富的結構為主。然而，隨著資訊與通訊技術（Information and Communication Technology, ICT）的精進，對於數位資訊的使用與長久保存方面的議題尚未引起廣

泛的注意，則是未來可以加以研究與發展的重點之一。

二、應用方式：多種標準與平行策略

由於應用特徵檔觀念（參見圖三：應用特徵檔的應用範例圖示）的導入與應用，有些 NDAP 主題計畫並不僅採用一套標準而已，而是兩套以上的標準，以因應不同的需求，類型包括：兩套以上的資料結構（如：ECML 與 XrML 之於臺北藝術大學－電子影音資料庫加值應用之研究）、資料結構與資料值各一套（如：CIDOC Ethno Core Data Standards、CDWA 與 MARC21 - Bibliographic and Authority Format 之於民族學研究所－臺灣原住民數位典藏計畫），以及兩套以上的資料值（如：AAT 與 MARC21 - Authority Format 之於故宮文物數位典藏系統之開發與建立）。此外，為了達成整合與大眾使用的目的，NDAP 主題計畫除了採用某一社群或學科的後設資料標準外，也一致使用 DC、OAI 作為資源發掘與聯合目錄之用，以達成資源發掘與互通。



圖三：應用特徵檔的應用範例圖示

三、描述層面：以物（即典藏品）、人與地理空間為主

如果將 NDAP 主題計畫所採用後設資料的資料結構與資料值加以彙整，並以人、事、時、地、物五大主軸進行分析後，發現 NDAP 主題計畫是以物、人、空間與地名為主，在時間與事件方面則尚未進行，尤其在時間方面，因為同一時期對藝術、歷史與考古學而言，劃分方式的原則與涵蓋年份是有所差異的。由於典藏品豐富，在 NDAP 範圍內難有統一、一致的共通標準可以採用與達成深度化的描述。然而，以原有基礎為出發點（即人與地），可以擴展至時間與事件，成為共通性的後設

資料，有利於共享共建及貫穿每個計畫，以展現不同的風貌。

四、以學科與社群為導向的後設資料核心集

由於在 NDAP 內容分項計畫之下，成立 12 個主題小組，並有一些主題小組開始進行所謂的後設資料核心元素集的發展。從發展現況觀點而言，這些小組的發展結果與 NDAP 其他主題計畫的應用結果十分雷同：描述深度以資料結構（如：CDWA、Darwin Core、EAD、HISPID、MARC21）與資料值（如：Species 2000）為主，其中多數集中在資料結構方面。另一方面，除了 DC 外，這些核心元素集傾向於以學科與社群為主要發展方向，前者如

Species 2000 與 Darwin Core，兩者皆是以生物多樣性為主；後者如 CDWA、EAD、HISPID、MARC21，傾向於博物館、檔案館、標本館與圖書館。不過值得注意的是，此種核心元素集的發展策略也有所不同，第一種是以某一標準為基礎進行分析與篩選；如：CDWA（書畫主題小組）、EAD（檔案主題小組）、HISPID（植物主題小組）、MARC21（善本主題小組）、Species 2000（動物主題小組）。第二種則是採用某一學科或社群為導向的既有核心元素集為依歸；如：Darwin Core（動植物聯合主題小組）。最後，也有一項主題計畫直接採用某一學科的既有核心元素為計畫標準；如：VRA Core Categories（臺北藝術大學、中央研究院臺灣史研究所－臺灣視覺記憶數位典藏計畫）。

柒、結語與建議

從上述的探討內容中，可以確知 NDAP 應用現況與全球後設資料發展趨勢兩者間雖然存有差異，也可以作為改善之參考。從另外一觀點而言，就一個國家型計畫已涵蓋多元化的後設資料，包括社群、學科、資料類型、功能，誠屬難能可貴，也無須苛求。最後，本報告從後設資料的研究與實務等觀點切入，建議未來 NDAP 在後設資料的研究與應用發展方面，提出下列八項建議：

- 應用領域：如何將應用與數位學習、電子商務予以銜接，以擴展更大的再利用價值。

- 層面：考慮是否發展時間與事件類的共通性後設資料。
- 功能：長久保存日漸重要，尤其珍貴典藏品可否經由一次數位化作業後，如何妥善保存這些數位資源，以利長久使用，避免第二次原件物品的數位化作業。
- 產生方式：可否經由電腦資訊科技的導入，半自動或全自動產生相關的後設資料，加速後設資料記錄的建立；如技術型與使用型後設資料。
- 描述深度：著錄規範與索引典的導入及其應用，將有助於資訊品質一致性的要求，以及未來聯合目錄的建置。
- 描述層次：除針對單一物件的描述，可考慮同時採取廣且深的雙向策略，選擇計畫進行館藏合集描述與全文標誌。
- 資訊交換（Interchange）：除了現行的 XML DTD 外，可留意發展專屬於 NDAP 的 XML Schema，以利於華文數位資源規範與標準的發展。
- 多種標準的整合與應用：應用兩種標準不是特例，未來應用兩種以上的後設資料標準逐漸增多，屆時相關的資料轉換機制與服務也須一併列入考量。

（收稿日期：2005 年 3 月 9 日）

（本文承蒙中央研究院計算中心沈漢聰先生的協助，得以順利蒐集與確認本文選取的數位典藏國家型科技計畫中，25 項主題計畫 54 份規格書有關後設資料標準方面的內容，在此一併致謝！）

辭彙對照表：

AACR: Anglo-American Cataloguing Rules

AAT: Art & Architecture Thesaurus

CCO: Cataloguing Cultural Objects: A Guide to Describing Cultural Works and Their Images

CDWA: Categories for the Description of Works of Art

CIDOC: The International Committee for Documentation of the International Council of Museums

CIMI: Computer Interchange of Museum Information

CLD: Collection Level Description

CRM: Conceptual Reference Model

CSDGM: Content Standards for Digital Geospatial Metadata

CSS: Cascading Style Sheets

CURL: Consortium of University Research Libraries

DC: Dublin Core

DCMI: Dublin Core Metadata Initiative

DDC: Dewey Decimal Classification

DISA: Data Interchange Standards Association

DOI: Digital Object Identifier

EAD: Encoded Archival Description

ebXML: Electronic Business using eXtensible Markup Language

ECHO: European CHronicles Online

ECML: Electronic Commerce Modeling Language

EELS: Engineering Electronic Library

EEVL: Edinburgh Engineering Virtual Library

FGDC: Federal Geographic Data Committee

FIAF: International Federation of Film Archives

FRBR: Functional Requirements for Bibliographic Records

GCS: Gazetteer Content Standard

GILS: Government Information Locator Service

GUID: Globally Unique Identifier

HISPID: Herbarium Information Standards and Protocols

IAFA: Internet Anonymous FTP Archives

IAML: International Association of Music Libraries

IANA: Internet Assigned Number Authority

ICA: International Council of Archives

ICANN: Internet Corporation for Assigned Names and Numbers
ICOM: International Council of Museums
ICPSR: Inter-university Consortium for Political and Social Research
IFLA: International Federation of Library Associations
IMC: Internet Mail Consortium
ISAD: International Standard for Archival Description General
ISAN: International Standard Audiovisual Number
ISBD: International Standard Bibliographic Description
ISBN: International Standard Book Numbering
ISOC: Internet Society
LCSH: Library of Congress Subject Headings
LDAP: Lightweight Directory Access Protocol
LDIF: LDAP Data Interchange Format
LOM: Learning Object Metadata
MARC: MACHine-Readable Catalogue
MADS: Metadata Authority Description Standard
MDRs: Metadata Registries
MESH: MEdical Subject Headings
MIX: NISO Metadata for Images in XML Schema
MNG: Multi-image for Network Graphics
MODS: Metadata Object Description Standard
OAI-PMH: Open Archives Initiative-Protocol for Metadata Harvesting
OAIS: Open Archival Information System
OASIS: Organization for the Advancement of Structured Information Standards
ODRL: Open Digital Rights Language
OLAC: Open Language Archives Community
PII: Publisher Item Identifier
PNG: Portal Network Graphics
PURL: Persistent URL
RAD: Rules for Archival Description
RDF: Resource Description Framework
RFC: Requests for Comments
SGML: Standard Generalized Markup Language
SICI: Serial Item and Contribution Identifier
SMIL: Synchronized Multimedia Integration Language

SOAP: Simple Object Access Protocol
SOIF: Summary Object Interchange Format
SVG: Scalable Vector Graphics
TEI: Text Encoding Initiative
UFBD: USMARC Format for Bibliographic Data
UMID: Unique MAterial Identifier
URC: Uniform Resource Characteristics
URI: Uniform Resource Identifiers
URN: Uniform Resource Name
VML: Vector Markup Language
VRA: Visual Resources Association
XML: eXtensible Markup Language
XrML: eXtensible Rights Markup Language
XSL: eXtensible Stylesheet Language
XSLT: eXtensible Stylesheet Language Transformations

參考書目：

- 陳雪華。(民 86)。網路資源組織與Metadata之發展。《圖書館學刊》，12，19-37。
- Dempsey, L., & Heery, R. (1998). Metadata: A current view of practice and issues. *Journal of Documentation*, 54(2), 145-172.
- Dempsey, L., Heery, R., Hamilton, M., Hiom, D., Knight, J., Koch, T., Peereboom, M., & Powell, A. (1997). Specification for resource description methods: Part 1: A review of metadata: A survey of current resource description formats. Retrieved May 24, 1997, from <http://www.ukoln.ac.uk/metadata/DESIRE/overview> directory
- Gilliland-Swetland, A. (2000). Introduction to metadata: Setting the stage. Retrieved December 24, 2003, from http://www.getty.edu/research/conducting_research/standards/intrometadata/pdf/swetland.pdf
- Heery, R., & Patel, M. (2000). Application profiles: Mixing and matching metadata shemas. *Ariadne*, 25. Retrieved September 27, 2000, from <http://www.ariadne.ac.uk/issue25/app-profiels/>
- MetaMap. (2003). Why a MetaMap? Retrieved February 23, 2005, from <http://mapageweb.umontreal.ca/turner/meta/english/whatfor.html>
- Turner, J.M., Moal, V., & Desnoyers, J. (2004). *MetaMap: Acronym index (Ver. 1.2)*. Retrieved December 10, 2004, from <http://mapageweb.umontreal.ca/turner/meta/english/indexacroeng.html>