

網路資源組織與Metadata之發展

Organizing Internet Resources and the Development of Metadata

陳 雪 華

Hsueh-Hua Chen

國立臺灣大學圖書館學系暨研究所教授兼系主任、所長

Professor and Chairman
Department and Graduate Institute of Library Science
National Taiwan University

【摘 要】

網際網路中的數位化資訊與傳統圖書館資源差異甚大，在邁入資訊網路社會的現在，為了掌握數位化資源的檢索與利用，圖書館有必要去了解其組織整理的方式。網路資源的使用方式可以透過檢索引擎或是主題指南，前者利用機器自動擷取，較容易且經濟。主題指南則經過一些人工加值工作，比前者費時費力費財。但此兩者均無法提供較詳細的資源描述給使用者，無法滿足使用者的資訊需求。要想有效率地進行資料檢索，metadata的建立是十分必要的。本文說明metadata之定義與功能、metadata的創作者與來源、metadata的種類、metadata的架構與完整性等，最後說明圖書館界的回應。

【ABSTRACT】

There exist lots of differences between information resources on the Internet and those in the traditional libraries. In order to retrieve and utilize digital information effectively in the coming era of information network, libraries have to explore how Internet resources are organized. Using search engines and subject gateway services are two common ways to retrieve and utilize Internet resources.

Search engines are based on robot to extract metadata, which are automatic and are cheap to create. Subject gateway services add value through intellectual effort, and are correspondingly expensive. But neither approach is complete as users are interested in resources at various levels of granularity and aggregation which may not be satisfied by either of these two simplified approaches. In order to use the Internet resources effectively, the establishment of metadata is very important.

This article describes the definitions and functions of metadata, a variety of metadata creators and sources, the different formats of metadata, the level of structure and fullness of metadata, and finally the responses and reactions from people in library field.

關鍵詞 Keywords :

網路資源；詮釋資料

Internet Resources; Metadata

壹、前言

近年來網際網路的風行，以及使用者對網際網路的興趣呈現快速成長，使得電子資源的需求變得殷切，不僅是商業性資源，也包括學術研究性資源。重要的專業資源，包括文字、圖像、以及聲音等各種類型資源，現亦可透過網際網路由不同的地方來檢索，成為重要的學術研究資源。以往，關於資料在何處以及其可獲性通常僅在某一特定團體中靠口耳相傳而已；至於對外界的人而言，要有效地獲得這些資訊就比較困難了。（註1）

有一些設計良好的網路檢索引擎可以自動檢索出網路上相關的網站，並且經常維護以保持其時效性及正確性。這種方法在有特定學科且小範圍的館藏上是最為有效的；但對資料庫很龐大並且內容跨越許多不同學科領域時，其正確性與相關性就較差了。此外，檢索引擎僅描述網頁（web pages），一些較新的趨向是加上伺服器（servers），但卻不描述這些伺服器中所存放的文件資料或這些文件之間的關係。

主題指南（subject information gateways）則彌補此缺點。如：Yahoo，蕃薯藤等。網路中還有許多以人工篩選並依不同的分類法整理而成的主題指南，這種經過加值的網路資源指南，提供一個便於使用者集中利用網路資源之處。此種系統

所提供的是一種邏輯性的主題檢索，可以藉著類號之間的層次隸屬或平行的關係，擴大或縮小檢索主題的範圍。所使用的分類法中主要包括杜威十進分類法（Dewey Decimal Classification，簡稱DDC）、國際十進分類法（Universal Decimal Classification，簡稱UDC）、以及美國國會圖書館分類法（Library of Congress Classification，簡稱LCC）等。以標題法來整理網路資源，最重要者為美國國會圖書館標題表（Library of Congress Subject Headings，簡稱LCSH）、以及醫學標題表（Medical Subject Headings，簡稱MeSH）這兩個標題表，前者所包含的主題為綜合性；後者則僅專以醫學為其範圍。

檢索引擎利用機器來自動擷取非架構式metadata的方式，較容易且經濟。資訊指南經過一些人工加值工作，比前者費時費力費財。但此兩者均無法提供較詳細的資源描述給使用者，無法滿足使用者的資訊需求。因此，如果要這些網路資源能夠很有效率地被找尋到，一定要讓熟知其內容的學科專家來加以描述。Metadata現已被多個與資訊及網際網路相關領域所熱烈討論及引用。在圖書館界存在已久的書目記錄，實質上就是一種metadata，它們對資料本體提供了書目的描述及相關資訊。隨著全球資訊網的快速發展，以及網際網路上存在的眾多資訊，要想有效率地進行資料檢索，metadata的建立是十分必要的。（註2）有一些特定領域的使用者（如：檔案、博物館等）期望對資源的描述資訊能夠再更豐富一些。這些資訊可以充分描述資源本身，並表明資訊之間的關係。這種完整性的metadata對資源的描述最為費時費力費財。這些因素將驅使創造作者自行放入文件中之metadata以及較複雜的自動擷取技術的發展。

貳、Metadata之定義與功能

Metadata是敘述一個資源屬性的資料，通常稱為“Data about data”或是“Data describes other data”（註3）、“Additional information that is necessary for data to be useful”（註4），可以翻譯為詮釋資料（註5）或是資源描述格式，亦有稱為超資料（註6）或是元資料（註7）者。Metadata可以用在許多不同的領域，甚至在日常生活當中，例如：描述個人的資料。在個人資料表中一般包含生辰年月、性別、身高、體重、學歷、經歷等項目，這種資料亦屬於metadata。但此文所稱之metadata則偏重於對網路資源的描述，這些描述可以幫助使用者在未看到網路資源之前即可知道該資源所在之處，並辨識其特質與內涵，讓使用者可以進一步去做評估與選擇。

在網際網路之極大的資源空間中，要有效地管理這無限多的網路資訊，非常有賴於metadata的有效管理。在目前的網路環境中，提供metadata服務的需要已經很顯明。在未來，網路資訊類型與通訊協定愈來愈多樣化時，對metadata服務的需要將愈大。Metadata不僅是資訊探索之鑰，它也是網路資源有效利用的基礎，並且在不同的通訊協定中，它亦可提供跨平台的操作功能。然而，在網路中似乎不會只有單一的metadata。其理由是：網路中存在著許多不同型式的資源，有些資源較為簡單，有些則相當複雜。有些資源只是滿足使用者暫時性的需要，因此，僅簡單描述即可；有些資源具有學術性或商業性的價值，則需有較詳細的描述。學術性資源之完整metadata亦有不同程度的使用，例如：完整的描述資料可以協助文件深入分析與使用，但其簡單的描述資料則可提供一般性的檢索服務。資源的提供者也相當多樣化：從商業性質到學術研究性質，從個人到機構等，其服務的對象亦是包羅萬象。由此之故，在網路中存在許多不同種類的metadata是一件極為自然的事。（註8）

一般而言，metadata可以包含許多方面的資訊，如：題名、作者、出版者、主題內容、品質、架構、歷史、使用權限、與其他作品的關係、適用對象等。Metadata所提供的功能包括：1.定位（location）：如何知道所需資源的儲存位置所在；2.探索（discovery）：如何找到所需的資源；3.文件記錄（documentation）：描述並記錄文件的性質與內涵；4.評估（evaluation）：協助使用者判斷資源對其之價值；5.選擇（selection）：幫助使用者決定是否取用該資源。這些功能不僅由網路資源的使用者來執行，亦可由其他的代理者（包括人或自動化的檢索引擎）所執行。（註9）

參、Metadata的創作者與來源

Metadata的創作者與來源有許多，大致而言，可以分為三大類型：作者（authors）、資源儲存的管理者（repository managers）、與其他第三者的創作者（third party creators）。作者本人所建立的metadata之重要性漸漸顯明。這種類型的metadata未來很可能以使用<meta>標示語言隱藏在HTML文獻中的方式呈現。第二類型的metadata創作者，對所管理的資源有責任，因此也可能對該資源加以描述。第三類的metadata的創作者，包括各類資源的仲介者，他們描述資源的對象可能本身並沒管理或儲存。（註10）此外，資源的描述可隨著不同的使用階段將之不同程度的加值。加值者可以包括：資源儲存的

管理者、提供主題指南服務者（如：蕃薯藤或Yahoo等）、提供檢索引擎服務者，以及許多仲介服務單位，包括圖書館館員或新知服務（Current Awareness）、專題選粹（Selective Dissemination of Information，簡稱SDI）、索摘服務的單位等。

Metadata，一旦被創造即可與他人分享。舉例來說，作者創造隱藏在HTML文件的metadata，可以被檢索引擎檢索。另外，metadata與資源本身可以是結合在一起或是分開的。前者最淺顯的例子是隱藏在HTML文件的<meta>標示符號。但最普遍的情形是一些特定領域的SGML架構（frameworks），如TEI header必須與TEI文獻相配合，但僅是TEI header也可單獨存在，它所描述的文件可以是存放在遠處。

肆、Metadata的種類

Metadata種類繁多，許多領域（如：檔案、地理資訊、博物館、圖書等）在進行同領域使用之metadata之標準化工作，目前這是一個熱門的研究領域。根據Lorcan Dempsey and Rachel Heery之研究報告《Specification for Resource Description Methods. Part I. A Review of Metadata: A Survey of Current Resource Description Formats》（March 19 1997）收錄22種metadata（註11）。茲將之歸納為八個大類，包括：

一、早已普遍使用的metadata格式

（一）MARC（註12）：USMARC（註13）、UKMARC（註14）UNIMARC（註15）

這是圖書館界使用的metadata。MARC是圖書館間為了交換及分享編目記錄所產生的標準化格式，現已為各圖書館及書目中心所廣泛利用。尤其在ISO2709中規定MARC記錄乃由資料欄（變長欄、定長欄）組成，使用不同的tag no.及分欄代碼來定義一些書目資料，此外並需有記錄標示、指引、資料欄分隔符號、記錄分隔符號等。所以即使不同的MARC其外在結構也都很類似，而不同的國家或機構也可依其需求產生適合的MARC，如USMARC、UKMARC等。

（二）PICA+

PICA（The Dutch Center for Library Automation）是一非營利性組織，為荷蘭許多學術及公共圖書館及德國部分圖書館網提供共享編目系統、館際互借、及文件傳遞系統等系統及服務。雖然PICA+ 格式深受MARC格式的影響，且依循ISBD

標準，但因其不符合ISO2709標準，所以並不被視為MARC格式的一種，但PICA+與MARC格式之間可作資訊交換。

二、描述科技文獻之metadata格式

(一) BibTeX (註16)

為描述科技文獻書目資料之格式，是LaTeX的一部份(LaTeX為文書排版軟體)，BibTeX原為其中從書目資料庫抓取文獻書目資料之程式。在科學、工業、學術界中廣泛使用在互相傳送科技期刊文章，格式並不繁複，惟需有使用LaTeX的經驗。

(二) EELS (註17)

EELS (Engineering Electronic Library) 是瑞典技術大學 (The Swedish University of Technology) 圖書館的一個計畫，目的在建立一個促進網路上工程方面資源之取得的資訊系統。計畫中所使用的metadata格式為其自訂，目前正考慮是否要改用較標準的格式，如Dublin Core或IAFA。

(三) EEVL (註18)

EEVL全名為英國愛丁堡工程虛擬圖書館 (The Edinburgh Engineering Virtual Library)，其目的與EELS大致相同。計畫中所用之metadata格式也屬自訂，但在設計上則盡量與IAFA相對照。

(四) RFC 1807 (註19)

為透過電子郵件傳遞技術報告書目記錄之格式，主要應用在美國科技界。其格式非常簡單，並且只需要普通的編輯器 (如Notepad) 即可，不需任何特殊軟體。RFC 1807 目前被五所大學的電腦科系所採用，分別為康乃爾大學、史丹佛大學、加州大學、麻省理工學院、及卡內基-美侖大學。

三、描述人文及社會科學資源之metadata格式

(一) ICPSR SGML Codebook Initiative (註20)

為描述社會科學資料組 (data sets) 的格式，1995年由大學政治與社會研究協會 (Inter-university Consortium for Political and Social Research) 所發展。欄位中

主要記錄資料檔案 (datafile) 或資料組 (data set) 的結構、內容、及形式等資訊。

(二) TEI (Text Encoding Initiative) Headers (註21)

原為人文學及語言學電子文件的描述格式，目的在促進資源的交換，但漸漸地也應用到其他學科範疇的電子文件上。基本上TEI包括了一組標示 (tags)，為SGML眾多DTD (Document type definition) 中的一種，換句話說，是SGML的一種應用。使用的難易程度則視所根據的外在規則而定，例如：如果根據AACR2或想要著錄與MARC格式相同的內容，著錄者就可能須具備編目的專業知識了。CIMI 的CHIO系統就採用TEI Lite (簡單版的TEI) (註22) 並融入一些CDWA的欄位來描述博物館的全文資料，例如展覽目錄 (Exhibition Catalog)。

四、描述政府資訊之metadata格式

(一) GILS (Government Information Locator Service) (註23)

政府資訊指引服務為美國聯邦政府所建立，目的在方便一般大眾及政府職員取得政府機構相關資訊。因其metadata涵蓋範圍廣，且訂定時深受MARC及Z39.50的影響，所以格式較為繁複，欄位中除了描述文獻式資料 (document-like-object) 的metadata之外，另外還描述簡單的地理空間性及時間性資料。目前已有與USMARC的對照表。美國聯邦政府大部分的機關已採用GILS，一些其他國家的政府，例如：加拿大及澳洲，也考慮運用GILS。

五、描述地理空間性資源之metadata格式

(一) CSDGM (Content Standards for Digital Geospatial Metadata) 又稱為FGDC (Federal Geographic Data Committee) (註24)

為地理空間性資料之描述格式，由美國聯邦地理資料委員會 (Federal Geographic Data Committee) 於1992年開始訂定，格式複雜，包含了超過三百個欄位。目前聯邦地理資料委員會正致力修改FGDC格式使成為一國際性的地理資訊描述標準。FGDC與DIF、GILS、USMARC、與Dublin Core等其他metadata格式的對照表已產生。美國六大電子圖書館計畫中之Alexandria Digital Library (加州大學聖塔芭芭拉分校) 即採用FGDC描述其地理空間性資料。

六、描述博物館藏品與檔案特藏之metadata格式

(一) CDWA (Categories for the Description of Works of Art) (註25)

1996年由AITF (Art Information Task Force) 提出，為描述藝術品或其影像之格式，適合博物館藏品及圖書館特藏。欄位共分二十六類，大類下又有子類，包括主題、記述、管理等項目，是頗複雜的描述格式。由於才剛被提出不久，未來發展如何仍待觀察，但美國博物館資訊交換協會 (Consortium for Computer Interchange of Museum Information, 簡稱CIMI) 在其CHIO (Culture Heritage Information Online) 計畫中訂定檢索點時，即採用了CDWA中許多主題描述性的欄位項目。

(二) CIMI (Computer Interchange of Museum Information) (註26)

為博物館藏品、圖書館特藏等之描述格式。其特點為採用了SGML DTD，將描述項目以標示 (tags) 的方式，直接在全文資料的內容中將metadata標示出來，優點在於可將冗長的全文資料根據其內容及結構有組織的區隔成小段，提供較傳統全文檢索精確的蒐尋及檢索。由於其格式牽涉到SGML (Standard General Markup Language)，在使用上相對地比較複雜。CIMI展示系統計畫“CHIO”目前已接近完成階段，該計畫的目標是要對CIMI結合SGML及Z39.50二大國際標準，以促進博物館資源的組織與檢索的理念進行試驗。

(三) EAD (Encoding Archival Description) (註27)

整個EAD包含了SGML DTD、Tag library、使用指南、及實際範例，主要對象為檔案、手稿特藏的描述。目的在提供檔案、手稿資料一個標準的機讀描述格式，並促進網路上手稿、檔案資料的取得。EAD與MARC可以並行使用，MARC格式描述資源的簡略書目性資料，EAD則提供階層式的、更詳細的描述，因此在使用上需具備檔案及手稿管理之專業知識。目前美國國會圖書館 (Library of Congress Network/MARC Standards Office) 為EAD標準的維護機構，而美國檔案管理員協會 (The Society of American Archivists) 則負責此標準的繼續發展，許多學術機構也有相關計畫進行中，例如：耶魯大學、加州柏克萊大學、哈佛大學、史丹佛大學等，因此EAD目前雖不是國際標準，仍有其發展潛力。

七、描述大量網路資源之metadata格式

(一) Dublin Core (Dublin Metadata Core Element Set) (註28)

為描述網路上的電子資源的簡單格式，1995年3月在 OCLC/NCSA Metadata Workshop 中被提出。由於其目的是要讓資源的創造者或出版者自行建立資源的 metadata，格式簡單容易使用為其一大特點。Dublin Core 的格式雖然簡單，但為了符合使用者的不同需要，仍然維持了相當的延展性 (extensibility) 及可變性 (modifiability)。在訂定之初，為了顧及其彈性，對於格式語法 (syntax) 並未明確規定，另外，qualifiers 的運用也一直是個爭議。

由於初期 Dublin Core 的設計特別是針對文獻式 (document-like-object) 的網路資源，目前已廣泛的使用在 HTML 格式的文件上；但在第三次 Dublin Core Workshop (1996) 中提出以 Dublin Core 描述影像的討論之後，Dublin Core 已被用在影像的描述上。

(二) IAFA/WHOIS++ Templates (註29)

為 IETF (Internet Engineering Task Force) 下的 IAFA (Internet Anonymous FTP Archive) 工作小組所設計，用來描述 FTP 檔案 (包括影像、文件、及聲音)、電子郵件論壇 (Mailing Lists) 及資料庫服務、電子郵件論壇檔案 (Mailing List Archives)、Usenet Archives、資料組 (datasets)、及軟體 (software packages) 等型式的資源。除了 Whois++ White Pages Directory System 之外，ROADS (Resource Organisation and Discovery in Subject-based Services) 亦採用 IAFA 作為資源描述格式，而目前 SOSIG (Social Science Information Gateway)、OMNI (Medical Information Gateway)、及 ADAM (Art, Design, Architecture and Media) 三個計畫即採用 ROADS 軟體。

(三) LDIF (LDAP Data Interchange Format) (註30)

LDAP 的全名為“輕量指南擷取協定” (Lightweight Directory Access Protocol)。值得一提之處為 Netscape 公司宣稱他們將會把 LDAP 與著名的 Netscape Navigator WWW 瀏覽器結合。LDAP 目前大多應用在電話簿型 (White Pages) 服務上。

(四) SOIF (Summary Object Interchange Format) (註31)

SOIF 為科羅拉多大學 Boulder 分校之 Harvest Architecture 計畫的一部分。SOIF

records描述了Harvest收集器（gatherer）所找到的資源之摘要，目前大部分的記錄是由程式自動產生，但因其格式簡單，也可以由人工記錄產生。目前SOIF僅在Harvest內部使用，無意成為一正式標準。

（五）URCs（Uniform Resource Characteristics/Citations）（註32）

在1992年，Internet Engineering Task Force（IETF）設立一致性資源辨識（Uniform Resource Identifiers，簡稱URI）工作小組，討論並發展網路資源有關於命名、描述與位置的標準。URI工作小組提出兩個草案：一致性名稱（Uniform Resource Names，簡稱URN）與一致性特徵（Uniform Resource Characteristics，簡稱URC）。URN是有關於網路資源的辨識，URC則為網路資源的metadata，換言之，URC的標準將提供描述網路資源的語法。經過五年多的努力，一套趨近一致性的URN標準已近完成，但是URC則仍然處於很模糊的狀況。雖然在IETF內部，URC的進展不多，但是類似於URC所扮演的功能與目的卻在Dublin Core的系列研討會，以及其他metadata計劃中有相當的成就與進展。（註33）

八、其他

（一）Warwick Framework（註34）

是不同的metadata之間的涵蓋架構（container architecture），作為交換之用。沒有任何一個metadata可以單獨滿足所有族群的需要，不同的族群需要有不同的描述格式，因此這個架構必須要集整各種不同的描述格式的功能。

Lorcan Dempsey and Rachel Heery之研究報告並未完整列出所有的metadata，尚有許多metadata並沒有列入在此報告之內，例如：CMARC、Handle System（Corporation for Networking Research Initiates, CNRI）、US Environmental Protection Agency（EPA）：Scientific Metadata Standards Project等。

伍、Metadata的架構與完整性

Metadata架構的複雜程度及所需付上的心智程度與資源價值、使用的環境息息相關。Paul Miller將metadata形式的粗略分為兩類：專家層次（expert approach）與檢索引擎層次（search engine approach）。前者包括MARC、TEI header等較複雜的資料描述架構；後者則在HTML文件中隱藏<META>的語法，使之可以被檢索引擎來

檢索（註35）。另外，Lorcan Dempsey與Rachel Heery依metadata的架構與完整性，將metadata分為以下三個大類（註36）：

表1：Metadata形式的種類

類別 項目	第一類	第二類	第三類
記錄	簡單格式	結構化格式	複雜格式
特質	專屬的	漸形成標準	國際標準
種類	Lycos Alta Vista Yahoo · ·	Dublin Core IAFA templates RFC 1807 SOIF LDIF	ICPSR CIMI EAD TEI MARC

第一類的metadata包括相當多無結構化的資料，典型的例子是透過搜索引擎自動擷取的資料。這種資料在語意上並沒有描述得很清楚，無法依領域來檢索。如果使用者找一個已知的項目，這種方式可能較有效，因其包括的範圍可遍及全球。但如果是屬於瀏覽性的搜尋則較無效率。因使用者可能找到一大堆資料，但可能多非所需，而有些潛在的有價值資料又因索引不到而無法找到。另外，這種方式找到的資料也往往因為 metadata資料的不完整，無法提供使用者正確的判斷。目前在第一類的metadata並無廣泛使用的標準，雖然有在發展一個交換的共享格式，也許是以SOIF做為發展的基礎。

第二類的metadata包括足夠的資源描述資訊，允許使用者在檢索之後，可以判斷是否需要取全文資料。這種metadata相當結構化，可以支援特定領域的檢索。一般而言，這種metadata記錄並不會太複雜，可以由文件作者來產生，或者可以由不具學科背景（discipline-specific knowledge）的第三者來產生。這種描述工作通常是以人工方式來進行，或是自動萃取描述工作，但需人工查檢與增強。這種metadata的資源可以直接載入資源指南以供瀏覽檢索。此類的服務包含OCLC的NetFirst以及Dublin Core，UK Electronic Libraries Programme Subject-based Information Gateways。這種服務往往會對網路資源加以選擇過濾，並多少有一選擇評鑑的標準。這種資料的產生較為昂貴，因此引起由作者本人或出版者自行產生描述資料，或自動萃取技術的興趣。例如：Essence試驗的Harvest軟體。

在第二類的metadata用於檢索服務方面則傾向於用簡單的架構，不包括詳細描述內部結構，因此不太容易表現資源之間的關係。在簡單性（simplicity）與表達性

(expressiveness) 兩者之間需做取舍。第二類metadata檢索服務通常透過檢索引擎或是資源指南，著名的例子如：Whois++、LDAP、Dienst。有些也用Z39.50，如GILS。

第三類metadata包含非常完整的資源描述架構。不僅包含描述資料，並包含所描述資料本身。典型的例子即TEI所設計的metadata。在社會科學、博物館、檔案和地理方面的領域也正在進行建立其DTD，致力於發展非常完整的屬性以描述資源。第三類的metadata包括非常完整的資源描述資訊，可以提供定位(location)與探索(discovery)的功能；此外，它尚扮演一個記錄網路資源的角色。它們通常與研究或學術活動有密切的關係，需要專門的知識背景才能產生並維護這種資源描述的資料。此資料可滿足特定領域學家的需求，如ICPSR描述社會科學類的資源；EAD描述檔案資源；CSDGM描述地理空間性的資源；CIMI描述博物館的資源。第三類metadata資訊的檢索協定的發展，目前仍在初期發展階段，其中Z39.50受到相當多的注目，許多學者專家在討論應用它的可能性。

以上的分類並不十分完美，尤其實施方式有很大的差異。例如，GILS與CIMI對資源的描述亦可視為第二類的層次。第三類之metadata與第二類的區別可以從：1.完整性(較詳細)；2.架構(包含較豐富的項目)；3.專門性(可能屬於某一特殊領域)來判別。很顯然，學者專家層次的使用者，可能會喜歡直接來查詢此類的資料，但為了顧及一般人查詢的需求，應該將此類metadata另以較簡略的方式輸出。

陸、Metadata實施應用情形

標準化的資源檔案服務的發展，目前仍處於早期發展的階段。有許多metadata格式的發展仍在進行中，或者有些雖已發展完成，但尚未廣為應用。這種情形在第三類metadata尤然。

第二類的metadata，IAFA/Whois++ templates應用在一些檢索服務的計劃，Dublin Core目前也應用在一些先導計劃中，RFC-1807用於NCSTRL計劃，SOIF則廣泛使用於Harvest的內部格式，但尚未有一致性的內容定義。LDIF與SOIF的情形類似，缺乏一致性資源描述的架構。但兩者都吸引許多的注意，因Netscape決定要利用LDIF建立其directory server產品在LDAP上。利用SOIF建立catalog server產品在Harvest上。(註37)

在所有metadata應用中，MARC格式的應用是最成熟的，具有悠久的歷史。目前，全世界已有數以億計的MARC記錄，其架構詳細，技術成熟，可以很容易產生

記錄並彼此分享。在圖書館界，長久以來利用MARC描述印刷式的資源，但最近也致力於用它來描述電子資源。有些單位也希望用它來描述網路資源，以求取與現成系統的一致性。MARC的記錄僅在某些層次上是一致性的。因各國均有其自己的MARC（如CMARC、USMARC、UKMARC等）。

目前，已存在的Z39.50的應用大多用於檢索MARC格式的資訊。但當其他metadata漸漸被使用之後，這種情形可能會改變。

柒、圖書館界的回應

最近數年，網路中電子文獻數量快速增長，網路資源的組織與整理受到許多相關領域人士的重視而紛紛發展各種metadata。圖書館界一向都非常重視資訊的組織與整理，因此對metadata的發展亦賦予極高的重視並以實際的行動來回應。圖書館館藏與技術服務協會（Association for Library Collections & Technical Services，簡稱ALCTS）在1995年成立一個工作小組（ALCTS Task Force on Meta Access），其主要的責任為：界定電子環境中書目控制與檢索的機制，以及此機制與使用者之間的溝通事宜。該工作小組將領導發展電子資訊之書目控制的工具，並將與美國圖書館協會資訊技術政策部（ALA Office for Information Technology Policy，簡稱OITP）密切合作，與其他圖書館或資訊界充分溝通。（註38）對工作小組而言，其重點即在了解究竟圖書館可以在電子環境中提供何種的書目檢索訊息。在此，工作小組認為首先要先界定出對圖書館界重要的問題；再來建議實際的行動讓圖書館界可以影響、參與、和回應，以期共同來解決問題。

該工作小組於1996年1月第一次會面，訂下工作目標與工作項目，並在1997年1月時完成其工作，在1997年4月時提出其工作報告：《ALCTS Taskforce on Meta Access: Final Report》。該報告提出未來發展的假設，並以此假設作為背景資料，陳述在此環境中所面臨的問題，並說明其建議的行動方案，可以作為圖書資訊學界未來發展的參考。以下即列出該報告之核心部份：（註39）

一、假設

1. 圖書館應繼續進行選擇、徵集、組織、提供檢索及傳遞資訊的服務。相信藉此不僅可儲存原始資料並發揮最大的效用，也能長久保存電子數位資料。
2. 國家層級的權威檔資料庫已提供圖書館及使用者相當好的服務。

3. 科技環境以及電子資源的持續改變是可以預期的。
4. 在描述資料、分析內容、組織資訊、書目查詢方面，有已通過的原則可採用，以達有效及有效率的資訊檢索。
5. 電子資源的特性與紙本式是不同的：如電子資源以超文字連結格式（hypertextual form）使其呈現的格式相當多樣化。
6. 在電子資源的檢索方面，圖書館會像提供其它形式的資源檢索一樣，如透過圖書館目錄、索引、資料庫等方式，以提供讀者利用。

二、問題陳述與建議行動方案：

（一）命名系統

在電子環境中，要確保每一筆資料的名字具有永久性與唯一性是很重要的；因此建議應該要採用通用性資源名稱（Universal Resources Names，簡稱URNs）固定且一致的命名方式。目前有兩個系統正在採用此命名方式，包括PURLS和Handles，此外尚其他的系統仍在醞釀階段中。ALCTS對此問題的建議行動方案為：

1. 將召開會議，了解實際存在的問題，並找出合宜的解決方案以維護電子文件的檢索。會議成員將包括美國國會圖書館、澳洲國家圖書館、挪威國家圖書館、芬蘭國家圖書館、國家數位化圖書館聯盟（National Digital Library Federation）、研究圖書館組織（Research Libraries Group）、OCLC、網路研究發展組織（Corporation for Networking Research Initiatives，簡稱CNRI）等。
2. 鼓勵資訊提供者去使用唯一且永久的命名方式，並提供為何使用及如何使用的相關資訊與說明。

（二）資料描述/metadata

將有多樣化的metadata系統來進行資料的描述、搜尋、以及檢索，而ALCTS的角色就是要從圖書館的角度來建立起對metadata有系統連貫性的觀點。其對此問題的建議行動方案為：

1. 派代表出席1997年7月於加州大學柏克萊分校所舉行的研討會：Workshop on Metadata Registries。
2. 在美國圖書館協會（American Library Association，簡稱ALA）所屬的兩個委員會（CC:DA與MARBI）中組織一個共同工作團體，其目的除了注意、參與圖書館界以外其他領域有關metadata的活動，也藉此讓其他領域

- 的人士了解圖書館界的意見與看法。
3. 對目前正在進行研擬出一套適於描述網際網路資源的標準，以利日後資源共享的工作，圖書館界應該要相當支持。這些工作包括ALA所屬的委員會（CC:DA與MARBI）。ALCTS支持Dublin Core做為資源描述的工作起點。
 4. 鼓勵圖書館界成為提升使用metadata標準的機構，特別是大學圖書館、政府機關圖書館、以及其它電子文件資訊提供的單位。
 5. 鼓勵各單位採用現有的分類系統，如杜威十進分類法、美國國會圖書館分類法、美國國家醫學圖書館分類法等。
 6. ALCTS注重電子資料的主題分析與分類。除了請研究者與學科專家提供意見外，也致力於與其它相關單位，如Internet Resources Task Force Working Group on Internet Resource Description的合作與協調。

（三）語法/格式

Metadata的種類有許多並以多種格式出現，來自不同metadata的記錄通常可能包括單一系統或是跨平台所使用的metadata。因此，各館的目錄系統應有更大的彈性，以便繼續支援MARC的使用，並能容納其他類型資源。工作小組目前正在支持由美國國會圖書館與其他單位共同合作MARC與SGML彼此對映（mapping）工作的發展。此工作的成果將能讓各圖書館在新的書目資料描述方式上有一致性。

在這方面的建議行動方案包括由ALA來召開一個國際高峰會議，召集不同國家的專家來研討在原先不同資源描述的架構組織中，資源如何彼此對映，希望產生metadata註冊的必備項目，特別是在資源描述方面。參與此會議的組織將應該包括ISO/NISO、LC、OCLC、RLG、ALCTS/LITA/RUSA MARBI、National Digital Library Federation（NDLF）等單位。

（四）記錄的管理

在1996年4月由英國圖書館與資訊網路部門（UK Office for Library and Information Networking，簡稱UKOLN）及OCLC所共同召開的The Warwick Metadata Workshop中，討論了許多有關metadata應如何由各圖書館的線上系統來管理的問題。各館的系統中包括了多種不同格式/語法的資料，因此，如何有系統地儲存、檢索、展示、及維護就成為重要的課題。

ALCTS在這方面的建議行動方案包括：將要召集系統管理者、編目館員、系統代理商、國家圖書館及資料庫的代表來開會，針對現有的書目控制（國家層次與國際

層次)架構中,書目資料的管理、館藏、metadata、書目記錄交換、權威檔的保存與維護等議題來進行討論。此外,尚需將未來國家書目控制系統的目標及要項界定出來。

捌、結語

描述資訊、整理資訊以便資訊較易被檢索,此工作在圖書館界中早已行之有年。然而在網路環境中,資訊的多樣性與其數量的龐大,使得資訊整理的工作面臨極大的挑戰。

Metadata是用來描述並記錄網路上的資源,並常以提供索引或目錄方式協助使用者(人或機器)探索到並使用這些存在的資源。所以metadata可以增加網路資源被檢索的機會,提供使用者對某主題相關資源做較全面性的認識,並提昇使用者在類似資源中的區辨能力。(註40)由於網路資源的組織與整理攸關其資訊檢索的效率與品質,metadata的發展將持續成為注目的焦點,不僅圖書館界,還有許多其他專業領域人士亦極為重視,例如美國電機電子工程師學會(Institute of Electrical and Electronic Engineers,簡稱IEEE)在1996年舉辦第一屆國際性metadata研討會,1997年9月又再舉辦第二屆。

未來網路資源的組織與整理,以及metadata要朝何種方向發展,目前而言,尚未有一個明確的方向。在過去兩年間,圖書資訊界與其他專業領域人士對以下所列的問題做許多探討:(註41)

1. 哪些資源值得提供metadata?
2. 由誰來提供這些metadata?
3. 不同metadata之間,如何彼此相容?
4. 單一metadata可否用來描述異質性的資源?
5. Metadata在何種繁簡層次與架構層次上須標準化?在何時、由何人來負責標準化的工作?
6. 當資源被描述之後,如何確保在資源有生之年可以被檢索到?
7. 如何處理許多具動態性質的網路資源?

身為圖書資訊界的一份子,我們必須參與探討以上的問題,並密切注意metadata的發展。在此網路資訊時代,扮演適當的角色並發揮我們的長處,對網路資源的組織與整理有所貢獻。

註釋

- 註 1 : Stuart Weibel, "Metadata: The Foundations of Resource Description," *D-Lib Magazine*, (July 1995)
(URL: <http://www.dlib.org/dlib/July95/07weibel.html>)
- 註 2 : "Metadata, Dublin Core and USMARC: a Review of Current Efforts", Discussion Paper No.99(Jan. 21, 1997)
(URL: <gopher://marvel.loc.gov:70/00/.listarch/usmarc/dp99.doc>)
- 註 3 : Terry Kuny, <Terry.Kuny@xist.com> "Metadata: What is It?" 24 Apr. 1997, <DIGLIB@INFOSEVNLC-BNC.CA> (24 Apr. 1997)
- 註 4 : Terry Kuny, <Terry.Kuny@xist.com> "Second IEEE Metadata Conference," 2 Dec. 1996, <DIGLIB@INFOSEVNLC-BNC.CA> (2 Dec. 1996)
- 註 5 : 陳雪華, *圖書館與網路資源* (台北市:文華, 民國85年), 頁207.
- 註 6 : 陳昭珍, 「電子圖書館與相關標準」, 在 *中國圖書館學會八十六年電子圖書館與資訊檢索專題研習班研習手冊*, 台大圖書館學系編 (台北市:該系, 民國86年7月), 頁9-9。
- 註 7 : 吳政叡, 「三個元資料格式的比較分析」, *中國圖書館學會會報* 57 (民85年12月), 頁35-45。
- 註 8 : Lorcan Dempsey and Rachel Heery, *Specification for resource description methods. Part I. A Review of Metadata : A Survey of Current Resource Description Formats* (March 19 1997.)
(URL: <http://www.ukoln.ac.uk/metadata/DESIRE/overview/>)
- 註 9 : 同註8。
- 註10 : 同註8。
- 註11 : 同註8。
- 註12 : MARC
(URL: <http://www.loc.gov/marc/marc.html>)
- 註13 : USMARC
(URL: <http://www.loc.gov/marc/marc.html>)
- 註14 : UKMARC
(URL: <http://minos.bl.uk/services/bsds/nbs/marc/>)
- 註15 : UNIMARC
(URL: <http://www.nlc-bnc.ca/ifla/VI/3/ubcim.htm>)

- 註16 : BibTeX
(URL: <ftp://ftp.shsu.edu/tex-archive/biblio/bibtex/distrib/doc/>)
- 註17 : EELS
(URL: <http://www.ub2.lu.se/eel/eelhome.html/>)
- 註18 : EEVL
(URL: <http://www.eevl.ac.uk/pub3.html>)
- 註19 : RFC 1807
(URL: <http://www.cis.ohio-state.edu/htbin/rfc/rfc1807.html>)
- 註20 : ICPSR SGML Codebook Initiative
(URL: <http://www.lib.umich.edu/codebook.html>)
- 註21 : TEI (Text Encoding Initiative) Headers
(URL: <ftp://info.ox.ac.uk/pub/ota/TEI/doc/teij31.sgml>)
- 註22 : TEI Lite
(URL: <ftp://info.ox.ac.uk:80/~archive/teij31/WHAT.html>)
- 註23 : GILS (Government Information Locator Service)
(URL: http://www.usgs.gov/public/gils/prof_v2.html)
- 註24 : FGDC (Federal Geographic Data Committee)
(URL: <http://www.fgdc.gov/Metadata/metahome.html>)
- 註25 : CDWA (Categories for the Description of Works of Art)
(URL: <http://www.ahip.getty.edu/gii/cdwa/>)
- 註26 : CIMI (Computer Interchange of Museum Information)
(URL: <http://www.cni.org/pub/CIMI/www/framework.html>)
- 註27 : EAD (Encoding Archival Description)
(URL: <ftp://library.berkeley.edu/pub/sgml/ead/>)
- 註28 : Dublin Core (Dublin Metadata Core Element Set)
(URL: http://www.oclc.org:5046/conferences/metadata/dublin_core_report.html)
- 註29 : IAFA/WHOIS++ Templates
(URL: <http://info.webcrawler.com/mak/projects/iafa/iafa.txt>)
- 註30 : LDIF (LDAP Data Interchange Format)
(URL: <http://www.umich.edu/~rsug/>)
- 註31 : SOIF (Summary Object Interchange Format)
(URL: <http://harvest.transarc.com/afs/transarc.com/public/trg/Harvest/user-manual/node151.html>)

- 註32 : URCS (Uniform Resource Characteristics/Citations)
(URL: <http://www.acl.lanl.gov/URI/>或<http://www.gatech.edu/iiir/iiir.html>)
- 註33 : Ray Schwartz, "Uniform Resource Identifiers and the Effort to Bring 'Bibliographic Control' to the Web: An Overview of Current Progress". *Bulletin of the American Society for Information Science* 24(1)(1997): 12-14.
- 註34 : Warwick Framework
(URL: <http://cstr.cs.cornell.edu:80/Dienst/Repository/2.0/Body/ncstrl.cornell%2fTR96-1593/html>)
- 註35 : Paul Miller, 「Metadata for the Masses」
(URL: <http://www.ukoln.ac.uk/ariadne/issue5/metadata-masses>)
- 註36 : 同註8.
- 註37 : 同註8.
- 註38 : "ALCTS Task Force to Address Digital Cataloging" *ALCTS Network News* 10:11 (OCT.19,1995)
- 註39 : Jennifer A. Younger. *ALCTS Taskforce on Meta Access: Final Report*. (Last modified: 30 Apr., 1997)
(URL: <http://www.lib.virginia.edu/alcts/about/final.html>)
- 註40 : Efthimis N. Efthimiadis and Allyson Carlyle, "Organizing Internet Resources: Metadata and the Web," *Bulletin of the American Society for Information Science* 24(1)(1997):4-5.
- 註41 : 同註40.