

圖書自動化

從書目控制到網路資源探索談 Metadata—由 Dublin Core 談起（下）

陳亞寧

（續上期）

現有問題與未來發展方向

即使 Dublin Core 已發展至成熟階段，但經過全世界各式研究計畫的測試實驗，發現 Dublin Core 仍有不足之處。本節茲分為現有問題及未來方向等兩方面，闡述如下：

（一）Dublin Core 現有問題

依循著 Dublin Core 六次會議的召開與討論，Dublin Core 雖然已經獲致某種程度上的成果，但朝向全球使用的世界標準之際，無論就理論與實務兩方面而言，Dublin Core 也有一些既存的問題尚待解決。本節僅就 Dublin Core 的現況發展與歷次會議討論議題中，進行 Dublin Core 問題的剖析與探討。

1.Relation Substructure：在 Dublin Core 元素中的 Relation 是用來標示不同物件間的相互關係，有時也會與 Source 元素配合使用。就現況而言，此一元素仍在發展中，所以無論在定義說明、使用上皆未定案，尤其是在結構上更無定論。就如同 Stuart Weibel、Juha Hakal 兩人發表的文獻中，兩人就提出在使用 Relation 元素時，至少涉及物件本身、相關物件及兩者間的關係類型等三部份實體（entity），但是 Dublin Core 尚未將著錄語法予以明確規定與說明，因而提出相同的質疑【 Weibel, and Hakala, 1998 】。

2.Source Recursive：有關 Source 元素的應用，實際情形也是十分複雜，因為 Source 的應用上會形成類似電腦程式中的迴路（recursive）現象。如果又搭配 Relation 元素的使用與標示，此種情形會更形複雜。但是在 Dublin Core 中，由於沒有正式的使用手冊，因為每個使用者與單位對其內涵與意義的詮釋亦是不同。因而，在使用方式上形成極不一致的現象【 Weibel, and Miller,

1997】。

3. Registry for Schemes and Subelements：由於 Dublin Core 十五個元素的制訂原則上是以通俗、簡單化為主，但是爲了更能精準標示內涵及辨識之用，因而許多實驗計畫或是實際使用單位在應用上，其實大多數皆應用到所謂的附元素或限定詞，或是針對某一元素增加其內容值（value）。然而，爲達統一化、標準化的目標，這些應用值與附元素必須在良好的品質控管下方能達成，所以需要有一處（registry）負責登記、註冊與維護這些附屬元素與應用值【Weibel, Iannella, and Cathro, 1997】。此一問題也隨著 Dublin Core Initiative Homepage（<http://purl.oclc.org/metadata>）的架設，以及相關工作小組的成立而漸趨解決。

4. Versions：在數位資源時代下，往往可以發現同一資訊內涵常以不同形式展現與使用；例如期刊、藝術品。以 Dublin Core 應用原則而言，是採取與圖書館編目規則相同方式—以手上整理的物件爲主，亦即遵循前述的內在固有性的原則。若從實體的觀點切入而言，即可分別物件的性質。以中央研究院 Metadata 工作小組實作分析結果中，將之區分爲原件與複製品等兩種類型。以期刊文章爲例，如果手上取得的是紙本式 JASIS 期刊中的 An Experiment on Node Size in a Hypermedia System 一文，但是網路上亦有電子形式者。基於此種前提下，紙本視爲原件，電子視爲複製品，兩者應用的 Dublin Core 元素就不盡相同，而兩者間亦可經由 Relation 及 Source 元素標引、建立彼此間的相互關係。相同的，同一份資源原本形式爲銅製品（如中央研究院數位圖書館與博物館計畫中子計畫墓葬的銅鏡），原件爲銅鏡，但複製品可能有一張像片、兩個電子檔案。無論是原件與複製品，或是相同形式的電子檔案都有版本判別的問題。

5. Collections vs. Item Level Description and Relation：在數位資源世界中，雖然每個物件看似獨立，但是仍有某種相互依存的關係，而其中之一即是所謂的包含關係，一如叢書與子目。由於 Dublin Core 中的 Resource 與 Relation 的應用與相關係並無明確的說明與規定，因而面對此種數位資源關係的標示，往往是模糊不清，或是由整理單位自行定義使用，未達成共識或一致性規範。所以，對於 Dublin Core 在邁向全球化之際，亦有所減損【Weibel, and Miller, 1997】。

6.Minimalist vs. Structuralist：所謂的簡化（ minimalist ）與結構化（ structuralist ）是相對稱的問題，前者主要源自 Dublin Core 發展制訂的原則－簡單化（ simplicity ）。因而，無論就 Dublin Core 的產生與相關工具的應用（如網路蒐尋引擎）而言，皆十分容易實施，而且可以輕易地達成全球通用的目標。基於資訊精準度的考量， Dublin Core 也允許使用所謂的限定詞或附元素來達成此一要求。在第四屆的 Dublin Core 會議報告內容中， Dublin Core 未將此兩種制訂方向（ approach ）視為衝突【 Weibel, Iannella, and Cathro, 1997 】。從事實應用上而言，其實是兩種 極端不同的發展方式，而地理資訊領域應用 Dublin Core 的 ECAI （ Electronic Cultural Atlas Interactive, <http://www.archaeology.usyd.edu.au/research/ecai/metadata/> ）即是其中一個最明顯的複雜結構化個案。因而，我們可以發現各個實驗計畫只要是應用到所謂的限定詞或附元素的情形下，就應用層次、應用原則與組成方式等方面而言，皆沒有一套共通的遵循規範。所以，前述不同的發展方式仍是一項隱憂，極待解決，否則屆時互通應用時，必須制訂許多不同的轉換表（ mapping table ）。

7.Multilinguality and Charater Set：為了追求全球通用、一致化的前提下，目前 Dublin Core 已有廿八種語文的版本，但是這些版本只限於 Dublin Core 元素的譯本而已，實際上並未將多語文處理的問題加以解決。由於 Dublin Core 強調的是跨語文與學科的特性，因而採取何種字集標準，就顯得十分重要。就現況而言，採用 Unicode 似乎是一項全球共通的趨勢，但是 Unicode 仍有 UTF-8、UTF-16 等問題。因而， Dublin Core 不若預期地容易解決有關字集字碼的問題。

8.Standardization： Dublin Core 雖然已通過 IETF （ Internet Engineering Task Force ）認證而取得 RFC （ Request for Comment ）文件的編號，但是實際上問題叢生。在第六次會議中主要議題之一即是如何應用資料模式（ data model ）與 ISO11179 （ Information Technology – Specification and standardization of data elements ）來規範制訂每一個 Dublin Core 元素，期能達成結構化、制式化、原則化與一致化【 Weibel, 1999 】。

綜合上述問題，其實根源是來自 Dublin Core 本身，亦即與制訂發展原則有著密不可分的關係。若依筆者實作與制訂 Metadata 的

實務經驗中，以及就歷次會議文獻內容探討為基礎，將之區分為五種類型，闡述如下（請參見圖一）：

1. 延展性及彈性（*extensibility and flexibility*）：雖然 *Dublin Core* 本身具備極高的延展與彈性，容許不同使用者進行修改、擴增，相形缺乏一致性規範的規定，形成百家爭鳴的現象，尤以 *Source*、*Relation* 與限定詞 / 附元素的應用最為明顯。

2. 結構化（*structure*）：由於缺乏一致性的著錄與制訂原則前提下，有關限定詞或附元素的應用，各單位可視需求與解釋自行決定。因而，我們可以發現每個 *Dublin Core* 元素若在不敷使用，或是為了追求精確的資訊品質標引時，限定詞或附元素的應用層次深度與結合方式都是不同的。所以基於 *Dublin Core* 高彈性的延展與簡單性兩種原則下，就本質而言，是具備高度的相互衝突、矛盾性的。然而為追求資訊精準度，以及轉換的容易度與成本考慮，必要性的結構化仍是無法避免的；而第六次會議採取應用 *ISO11179* 與資料模式的決定，即可視為重新正視此問題根源，並試圖解決之。

3. 語意著錄獨立性或語文（*syntax-independence or language*）：即使在著錄獨立性的原則下，可以忽略語意表示的問題，似乎可以解決語文的問題。但是基於全球通用的前提下，字集字碼仍與多語文處理的相伴而生，並無法忽略不顧，尤其是在結合應用 *XML* 協定，有更多語意的表示與互通性皆可由 *XML* 來規定、陳述與展現。因而對 *Dublin Core* 而言，多語文的處理仍是必須解決的重要課題之一。

4. 資源內在固有性（*intrinsicity or target*）：無論是紙本或是電子式資源，皆有不同版本與不同資訊載體單位的問題，如何判別出最基本的資訊單位（*basic unit*）仍是根本的問題。若從物件導向（*Object-Orient ed, OO*）觀點而言，亦即最小、最基本的物件單位為何？若從 *IFLA*（*International Federation of Library Associations and Institutions*）的「*Functional Requirements for Bibliographic records*」研究報告內容中，將物件區分為作品（*work*）、表達意念方式（*expression*）、資料形式（*manife station*）與資料實體（*item*）等四種【*International Fe deration of Library Associations and*

Institutions, 1998, pp. 12-23】。也許綜合上述兩種觀點的發展方式，是一項頗為可行的解決途徑。

5. 國際化 (international)：在邁向國際化與全球化之際，Dublin Core 應特別留意一致性與規範性，無論是元素結構、多語文、限定詞 / 附元素的使用等皆是。

(二) Dublin Core 未來發展方向：

即使 Dublin Core 具有上述缺點，但在眾多 Metadata 中，仍是最被看好之一。美國匹茲堡大學資訊科學學院圖書館資訊學系 (Department of Library and Information Science, School of Information Sciences, University of Pittsburgh) 服務的 Harold Thiele，以文獻探討方式針對 Dublin Core 作一深入瞭解、分析後，試從使用者行爲 (behavioral side)、技術層面 (technical side) 與社會議題 (sociological issues) 三大領域提出建議【 Thiele, 1998】：

a. 使用者行爲：

- Dublin Core 應與其他 Metadata 比較，是否較其他 Metadata 更有效地滿足使用者需求；如查詢。

- 在龐大的網路領域中，Dublin Core 這些替代性物件描述實體 (surrogate description) 能否對精確度 (precision) 有所提高或增益？

b. 技術層面：

- 在查詢過程中，對於網路存取記憶機制 (cache) 的運作效能及網路頻寬 (bandwidth) 是否有助益？

- Dublin Core 比較適合於集中式或分散式的網路資源蒐尋引擎的應用？

c. 社會議題：替代性的物件描述實體的模式有傳統學術團體 (traditional academic & research paradigm) 與非傳統學術的個人 (individual) 等兩種，前者通常使用這些描述實體 (如 Dublin Core)，後者則不使用。而使用與否，則取決於網際網路資源是否需要應用這類替代性描述實體作為品質的認證與審核

之機制（ authenticating or validating mechanisms ）。

另類思考 Metadata 的再定位

由 Dublin Core 的探討，對於 Metadata 的制訂、產生與應用也有初步的瞭解。本節以 Metadata 的定義、功能及索引層次與深度等三個方向，重新思索 Metadata 的定位、需求與未來發展方向。

1. 定義：

一般而言，有關 Metadata 的定義十分眾多，而且看法也有所不同，但仍以「資料的資料」（data about data）為多數人接受，且較無爭議。除此之外，重要者列舉如下：

- 描述資源屬性的資料（Data which describes attributes of resources.）【 Dempsey, Heery, Hamilton, Hiom, Knight, Koch, Peereboom, and Powell, 1997 】

- 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.）【 Chilvers, and Feather, 1998 】

- Metadata 是一種有關全球資訊網資源或其他的機讀資訊（Metadata is machine understandable information about web resources or other things.）【 Bernes-Lee, 1997 】

- Metadata 就是資料（Metadata is data.）【 Bernes-Lee, 1997 】

- Metadata 是包括有關資料特性的資訊，但是 Metadata 必須對資料加以描述、詮釋，層次上應該涵蓋：人、物、時、地、原因及途徑等方向（Metadata consist of information that characterizes data. Metadata are used to provide documentation for data products. In essence, Metadata answer who, what, when, where, why, and how about every facet of the data that are being documented.）【” Tools for creation of formal metadata, ” 1998 】

至於真正的 Metadata 內涵究竟為何呢？其實圖書館長久以來建立、維護與提供的線上公共目錄的書目資訊即是其中一個典型範例。但是 Metadata 範疇只限於此？若依據中央研究院 Metadata 工作小組研究與實作心得，已證實 Meta data 除了在於描述、詮釋資料與書目資訊相同外，但在層次與深度方面與傳統書目資訊是有所不同的。主要差異有四個方面：一為，任何物件皆是 Metadata 涵蓋的範圍，諸如圖書館界所熟悉的書、期刊、文章等，乃至器物、人等皆是。二則，範圍從實體典藏的物件擴大至虛擬典藏的各類物件。三是，詮釋的深度遠比以往更為深入，並不僅限於內容主題的分析標引而已，尚包括了物件的彼此互動關係，包括人、時、地、物、主題 / 事件 (events) 等五大主軸間的互動、牽引；換言之，從資料、資訊的整理提昇至知識內涵的建構。四為，物件的辨識、保存、展示、取用、篩選與評估、服務及系統管理等方面皆是環環相扣，密不可分，並不能只從單一觀點視之或處理。

2.功能

Metadata 目的是為了有效整理網路源，進而解決資源探索與資訊篩選的問題。究竟 Metadata 是否只需具備描述資源的功能即可以達成前述要求？若依 Lor can Dempsey 與 Rachel Heery 等人提出的「Specification for resource description methods. Part

1. A review of metadata: a survey of current resource description formats」報告內容中，他們列舉 Meta data 應該具備資源的位置標引 (location)、探索 (discovery)、記載 (documentation)、評估 (evaluation) 與選擇 (selection) 等五大功能。在此報告內容中，也提到 Metadata 還有兩項基礎功能：一為，是一種有效利用資源的一種基礎 (be fundamental to effective use of found resources)；二則，必須跨學科領域的互通應用 (be interoperable across protocol domains)

【 Dempsey, et al., 1997 】。由本文探討 Dublin Core 中，亦可發現 Dublin Core 的確符合如此的要求，也正朝這些方向進行中。

3.索引層次與深度

既然 Metadata 是為了解資源探索而生的，自然與檢索脫離不了關係。依 Stuart Weibel, Renato Iannella 與 Warwick Cathro 等

人在第四屆 Dublin Core 會議報告文獻內容中，他們將 Dublin Core 發展方向歸納為簡單化與結構化（minimalists and structuralist）等兩種，同時進而引發、歸納資源探索的五種方式：全文式索引（full-text searching）、非欄位化的替代性描述物件（unfielded surrogates）、簡化式欄位的替代性描述物件（minimally fielded surrogates）、制式化替代性描述物件（qualified surrogates）與高度結構化替代性描述物件（richly-structured surrogates）。一般而言，愈簡單化時，資訊建立所需的成本與時間是愈低，但是資源精確度則趨向愈低；反之，精確度愈高，所需的成本與時間則愈高。前者如 Dublin Core，後者如 FGDC（Federal Geographic Data Committed）Metadata。除此之外，在選用、制訂一套 Metadata 要留意索引的層次與深度外，更重要的是與不同 Metadata 間的轉換，以及展現方式的詳簡層次等方面亦有十分密切的關係。

從圖書館書目控制的觀點反思 Metadata 的應用與發展

究竟 Metadata 何去何從，其最終結果尚待時間與應用上等雙重考核後才有定論。在未有任何明確結果或定論前，無妨從荷蘭 DESIRE（Development of a European Service for Information on Research and Education）Project 預計要求 Metadata 達成目的為例，以作為 Metadata 的一種預測，如下：

- 是否足以支援有效的使用者檢索及資源篩選？
- 能否利用現存在網際網路的各項協定來檢索、查詢？
- Metadata 是否容易產生而不論使用者背景為何？
- 不同 Metadata 轉換時的資料漏失程度與比率；即精確度為何？【Heery, 1996】

在圖書館亟思改變採用或制訂 Metadata，而放棄既有的書目控制、分類與編目，以及 MARC 前，不如反覆思索或反問自己下列問題：

- 為何不用 MARC 既有的資料欄位同時來描述現有的非電子與電子式的資源？原因為何？有何差異及優缺點？

- Metadata 可以應用於非電子形式的資源？原因為何？
- 就實質層面而言，Metadata 與 MARC 的差異為何？
- 圖書館行之多年的書目控制與分編理論、實務真得完全不再適用？或是仍可適用？原因為何？如何應用（修正與詮釋）？
- Metadata 應用層面只限於描述與資訊檢索而已？是否包括典藏（archive）、詮釋？方式與書目控制、分編及 MARC 有何不同？
- Metadata 僅局限於書目資訊而已？
- 原有圖書館書目控制強調的聚合功能（syndetic），在 Metadata 應用與展現時，是否仍具價值？
- Metadata 如何與現有的圖書館書目資訊結合為一？方法為何？效能如何？
- Metadata 是由學科專家、一般使用者，或是圖書館學專家產生的、使用的？

參考書目

- Chilvers, A., and Feather, J. (1998). The Management of digital data: a metadata approach. *The Electronic Library* 16(6), 365-372.
- Dempsey, L, Heery, R., Hamilton, M., Hiom, D., Knight, J., Koch, T., Peereboom, M., and Powell, A. (1997). A Review of metadata: a survey of current resource description formats, executive summary section (Version 1.0.) [Online, Access Date : 24 May, 1997]. Available : http://1.0..ukoln.ac.uk/metadata/DESIRE/overview/rev_pre.htm
- Heery, R. (1996). Resource description: initial recommendations for metadata formats (Ver. 1). [Online,

Access Date : 24 August, 1998]. Available :
<http://www.ukoln.ac.uk/metadata/desire/recommendations>

- International Federation of Library Associations and Institutions. (1998). IFLA Study Group on the functional requirements for bibliographic records. Munchen: Saur. [Online, Access Date : 24 December, 1998]. Available :
<http://www.ifla.org/VII/s13/frbr/frbr.pdf>

- Thiele, H. (1997). The Dublin Core and warwick framework. D-Lib Magazine. [Online, Access Date : 22 January, 1998]. Available :
<http://www.dlib.org/dlib/january98/01thiele.html>

- Tools for creation of formal metadata: frequently-asked questions on FGDC metadata. (1998). [Online, Access Date : 29 August, 1998]. Available :
<http://geology.usgs.gov/tools/metadata/tools/doc/faq.html>

- Weibel, S., and Hakala, J. (1998). DC-5: the Helsinki metadata workshop: a report on the workshop and subsequent developments. D-Lib Magazine. [Online, Access Date : 20 August, 1998]. Available :
<http://www.dlib.org/dlib/february98/02weibel.html>

- Weibel, S., Iannella, R., and Cathro, W. (1997). The 4th Dublin Core metadata workshop report. D-Lib Magazine. [Online, Access Date : 19 June, 1997] Available :
<http://www.dlib.org/dlib/june97/metadata/06weibel.html>

- Weibel, S., and Miller, E. (1997). Image description on the Internet: a summary of the CNI/OCLC Image Metadata workshop. D-Lib Magazine. [Online, Access Date : 20 August, 1998]. Available :
<http://www.dlib.org/dlib/january97/oclc/01weibel.html>