

## 三個元資料格式的比較分析

吳政叡

### 中文摘要

在本文中，我們介紹了三個跟 metadata 相關的標準，它們分別是 FGDC 的 Digital Geospatial Metadata、Dublin Core、和 URC。雖然它們各有自己的設計目標和特質，但都是假設其操作環境為類似網際網路的環境。FGDC 的 Digital Geospatial Metadata 是設計來專門處理地理性資料，由於它有聯邦行政命令的支持，可說是已成為美國在地理方面的資料著錄國家標準。Dublin Core 則比較像是 USMARC 的網路節縮版，使非專業人士也能在短時間內熟悉和使用此格式來著錄收藏資料，但在現階段祇針對類似傳統印刷品的電子文件。由 IETE 的 URI 工作小組所負責的 URC，其原始的設計目的雖是用來連結 URL 和 URN，但為因應電子圖書館時代的要求，其內含逐漸擴大，雖然尚在發展中，但由於有 IETE 的支持，未來成為網際網路上通用標準的可能性極大。在此文中，我們也從幾個不同角度，分析和比較這三個 metadata 格式的異同和優缺點。

關鍵字：元資料，metadata，URC，Dublin Core，Digital Geospatial Metadata

### Comparison of Three Metadata Related Standards

#### Abstract

In this work, we introduce three metadata related standards. The first standard is Digital Geospatial Metadata which was created by FDGC on June 8, 1994 under the orders of US Executive Order 12906. The second one, the Dublin Core, is the product of OCLC/NCSA Metadata Workshop held in 1995. The last one is URC which is still under construction and supervised by the URI Working Group of IETE. Although these three different metadata standards have their own objectives and characteristics, they all emphasize Internet-like working environment. The first two aim at more specific types of data; The Digital Geospatial Metadata is designed for geospatial data while the Dublin Core is like an abbreviated network version of USMARC to handle document-like objects. On the other hand, URC is the most comprehensive metadata format among them and can be applied to all kinds of documents on Internet.

#### 一、前言

元資料 (metadata) 一詞最先出現於 NASA 的 Directory Interchange Format (DIF) 手冊中 [註 1]，最常見的英文定義是 "data about data"，[註 2, 註 3] 可直譯為描述資料的資料，就其本義而言，跟目錄 (Catalogue) 所扮演的角色並無太大的差別。[註 2] 編製目錄的目的，也在描述收藏資料的內容或特色，進而達成協助資料檢索的目的。

在電腦尚未興起前，資訊檢索的效率幾乎全依賴人工製作的卡片目錄的品質。在電腦興起後，圖書館自動化逐漸盛行，於是有機讀目錄 (Machine Readable Cataloguing, MARC) 的產生，[註 4] 來利用電腦提昇編目效率和結合資料庫來改善檢索效率。[註 5] 不過 MARC 的整體架構可說是承襲卡片目錄，換言之，是卡片目錄的數位化版本，後來隨著網路的發展，線上公用目錄 (Online Public

Access Cataloging, OPAC) 也躍上舞台, [註 6] 使得資訊檢索更為便利。但就其設計理念和使用者的檢索方式來看, 可說是MARC的網路化, 也就是卡片目錄 (數位 + 網路) 的化身。當一個讀者使用 OPAC 時, 基本上所有的輸出入都是以文字為主, 檢索方式也是主要透過書名、作者和關鍵字等項目來著手, 我們可以將此種作業方式, 想像成一個讀者叫一個動作很快的電子僕人, 進入指定的圖書館, 再利用卡片目錄去尋找所要的資料。

然而由於科技的進步, 網路頻寬 (bandwidth) 的不斷增加, 促成多媒體 (multimedia) 和 World Wide Web (WWW) 的盛行, 再加上國家資訊基礎建設 (National Information Infrastructure, NII) [註 7] 的推波助瀾, 催生了電子圖書館 (Digital Library) [註 8, 註 9], 文字媒體不再是單一主角, 聲音、圖片、影像、動畫等和文字成為舞台上的共同主角, 而文件 (document) 也取代書籍 (book) 成為 WWW、NII、和 Digital Library 文獻上關於資料的主要代表名詞, 而文件一詞所涵蓋的資料類型可說是包羅萬象, 包含如私人信件、商業文書、電腦檔案 (含程式) 等, 它同時也可以是包含文字、聲音和圖像等的混合媒體。同時在資訊檢索的方法上也起很大的變化, 我們現在要求這位電子僕人自己去利用檢索引擎 (search engine), 如 Infoseek [註 10], 找出資料來源 (我們不再須要自行指定 IP 位址), 自動處理資料傳遞 (即使你退出系統, 系統仍會繼續運作, 並在下次你登入系統時, 提醒你有那些新資料, 其作業方式很類似現在的電子郵件, 可說是你個人的電子秘書), 協助你過濾資料, 並將資料按重要性和關連性加以排序, 以利讀者做更進一步查核。為了要達成這種種新增的服務需求, 在對資料內容和特性的描述上, 我們須要增加一些新的項目, 來滿足電子圖書館時代的資訊檢索需求。

本文擬從下列三個方向來介紹現在 metadata 的發展現況和未來趨勢：

首先是介紹美國聯邦地理資料委員會 (Federal Geographic Data Committee, FDGC) 的地理電子元資料 (Digital Geospatial Metadata) 標準, 這個 metadata 的出現, 似乎在某種程度上反映了 USMARC 在文字媒體以外的領域, 不能滿足某些專業人士的需求。其次是介紹 1995 年 OCLC/NCSA Metadata Workshop 的成果 -- 都柏林核心集 (Dublin Core), 這個 metadata 格式反映出圖書館界或 MARC 為求適應新潮流, 所嘗試邁出的一步。最後是涉及範圍最廣的資源特性統一描述格式 (Uniform Resource Characteristics, URC), 這是由於 WWW 盛行後, 所暴露出的問題, 在電腦和網路界所逐漸發展出的一個 metadata 格式, 做為對現存 WWW 各種衍生問題的一個解決方案。在結論中, 我將對這三種 metadata 做進一步的比較分析。

## 二、地理電子元資料標準

FGDC 的 Digital Geospatial Metadata [註 11] 是根據 1994 年美國總統行政命令 (Executive Order 12906) [註 12] 而來, 此行政命令的大概內容是說鑒於地理資訊對於經濟發展、改善自然資源管理、和環境保護的重要性, 命令 FGDC 負責制定一個關於地理資料收藏、著錄和交換的標準, 用來統合政府和民間所收藏地理資料的散布和傳播。此命令中最重要的部分, 是強制所有聯邦機構, 自此命令公布的 9 個月後 (即 1995 年 1 月起), 必須遵循此一標準, 包含所有聯邦機構參與的活動 (如經費贊助)。在此行政命令的規範下, FGDC 的 Digital Geospatial Metadata 可說是成了美國地理資料著錄的標準。

事實上在現階段的文獻上, metadata 一詞幾乎與 FDGC 的 metadata 標準劃上等號, 如在 Internet 上以 metadata 一詞來做搜尋時, 十之八九指向與 FDGC 的 metadata 標準相關的站址。FDGC 將 metadata 定義為 [註 11]

"data about the content, quality, condition, and other characteristics of data",

另一個類似的定義為 [註 13]

"descriptive information about a digital spatial data set that includes text, numeric, date, and geographic elements stored as addressable fields"

概括而言，metadata 的定義是

對電子地理資料其內容、製作品質、狀況和其他特性的描述，用來協助網路上的資訊檢索。

至於地理資料的定義：[註 12]

地理資料是用來標示地球上地理位置和自然或人造特色的資訊，這資訊可能源自遙測或其他調查製作技術，經收集機構判定後的統計資料亦包含在內。

由以上定義可知，此標準涵蓋的資料類型極廣，包含地圖、衛星照片等，甚至包括地理方面的統計資料，可說是祇要與地球上的地理有任何關連，都是其所欲處理的對象。

根據 1994 年 6 月 8 日 FDGC 公布的標準，共有約 300 個資料項，分成十個節（section），其中有七個主要節（main section），三個支持性節（support section）。因為篇幅的限制，在此不列出該標準的詳細內容，請自行參照[註 11]，但為了使讀者對其整體架構有概括的認識，在此將第一層和第二層的項目列出。為讀者閱讀便利，先將下面表示法用到的符號加以說明：

1. 一對大括號 "{}" 左右兩邊的數字，分別表示被大括號所包含資料項的最少和最多重覆次數，如 "0{}1" -- 表示資料項最少出現零次，最多出現壹次；"1{}n" -- 表示資料項最少出現壹次，最多出現的次數則無限制，可按實際需要自由重覆此資料項。
2. 一對中括號 "[]" 內含許多資料項，並以垂直線隔開，如 "[ | | ]"，表示在中括號所含的許多資料項中，擇一適當的資料項來應用。
3. 一對小括號 "()" 表示被其所包含的資料項可省略。

Metadata =

Identification\_Information + 0{Data\_Quality\_Information}1 +  
0{Spatial\_Data\_Organization\_Information}1 + 0{Spatial\_Reference\_Information}1 +  
0{Entity\_and\_Attribute\_Information}1 + 0{Distribution\_Information}1 +  
Metadata\_Reference\_Information

第一節. 識別資訊: 資料集的基本資料。

Identification\_Information =

Citation + Description + Time\_Period\_of\_Content + Status + Spatial\_Domain + Keywords +  
Access\_Constraints + Use\_Constraints + (Point\_of\_Contact) + (1{Browse\_Graphic}n) +  
(Data\_Set\_Credit) + (Security\_Information) + (Native\_Data\_Set\_Environment) +  
(1{Cross\_Reference}n)

第二節. 資料品質資訊: 資料集製作品質的一般性陳述。

Data\_Quality\_Information =

0{Attribute\_Accuracy}1 + Logical\_Consistency\_Report + Completeness\_Report +  
0{Positional\_Accuracy}1 + Lineage + (Cloud\_Cover)

第三節. 資料製作方式資訊: 呈現資料的機制。

Spatial\_Data\_Organization\_Information =

0{Indirect\_Spatial\_Reference}1 + 0{Direct\_Spatial\_Reference\_Method +  
([Point\_and\_Vector\_Object\_Information | Raster\_Object\_Information])}1

第四節. 地理參考資訊: 有關投射座標和解析度的資訊。

Spatial\_Reference\_Information =

0{Horizontal\_Coordinate\_System\_Definition}1 + 0{Vertical\_Coordinate\_System\_Definition}1

第五節. 實體和屬性資訊: 資料所代表對象的資訊。

Entity\_and\_Attribute\_Information =

[Detailed\_Description | Overview\_Description]

第六節. 傳布資訊: 傳布機構資訊和獲取資料的方法。

Distribution\_Information =

1{Distributor} + 0{Resource\_Description}1 + Distribution\_Liability + 0{Standard\_Order\_Process}n +  
0{Custom\_Order\_Process}1 + (Technical\_Prerequisites) + (Available\_Time\_Period) }n

第七節. metadata 本身相關的參考資訊: metadata 本身和製作者相關的機構資訊。

Metadata\_Reference\_Information =

Metadata\_Date + ( Metadata\_Review\_Date + (Metadata\_Future\_Review\_Date) ) +  
Metadata\_Contact + Metadata\_Standard\_Name + Metadata\_Standard\_Version +  
0{Metadata\_Time\_Convention}1 + (Metadata\_Access\_Constraints) + (Metadata\_Use\_Constraints)  
+ (Metadata\_Security\_Information)

第八節. 引用資訊: 資料集在被引用時，建議使用的資訊。

Citation\_Information =

1{Originator}n + Publication\_Date + (Publication\_Time) + Title + 0{Edition}1 +  
0{Geospatial\_Data\_Presentation\_Form}1 + 0{Series\_Information}1 + 0{Publication\_Information}1 +  
0{Other\_Citation\_Details}1 + (1{Online\_Linkage}n) + 0{Larger\_Work\_Citation}1

第九節. 涵蓋時期資訊: 有關此地理資料時間方面的資訊。

Time\_Period\_Information =

[Single\_Date/Time | Multiple\_Dates/Times | Range\_of\_Dates/Times ]

第十節. 聯絡資訊: 接洽人或組織的相關資訊。

Contact\_Information =

[Contact\_Person\_Primary | Contact\_Organization\_Primary] + (Contact\_Position) + 1{Contact\_Address}n + 1{Contact\_Voice\_Telephone}n + (1{Contact\_TDD/TTY\_Telephone}n) + (1{Contact\_Facsimile\_Telephone}n) + (1{Contact\_Electronic\_Mail\_Address}n) + (Hours\_of\_Service) + (Contact\_Instructions)

以上所列項目可分為三類：必須（mandatory）項目、適用則須有（mandatory if applicable element）項目、可省略（optional）項目，其中必須項目不到三分之一。第一至七節為主要節，而第八至十節為支持性節。

就其設計理念而言，是基於類似 NII 的國家空間資料基礎建設（National Spatial Data Infrastructure, NSDI）架構，在資料的交換、查詢、散布上，以電腦網路為其假想作業環境，因此在 US Executive Order 12906 中，明令要設置全國性地理資料的交換所（national geospatial data clearinghouse）-- 即一個連結地理資料生產者、管理者和使用者的分散式電子網路，[註 12] 換言之，交換所（clearinghouse）是機構或組織用來協助地理資料的搜尋、評價和傳布的一套軟體。[註 13] 交換所（clearinghouse）活動雖是由 FDGC 贊助，但它基本上是一個分散式的網路系統，由散布在網際網路（Internet）上的眾多伺服器（server）組成，利用 ANSI Z39.50 協定作為主從式架構下，主/從站間資料查詢和結果回傳的協定。此外 FDGC 也製作了一個 metadata 的格式規範（profile）-- GEO [註 14] 作為在 Z39.50 架構下，操縱 metadata 資料項的依據。

為了確保著錄資料的語法（syntax）正確，在網際網路上已有此 metadata 的語法查核軟體 [註 15]，來協助維持著錄的品質。此外也有一個教導初學者如何寫作的指引文章。[註 16] 另外也如 HTML 一般，有多種協助寫作的文書軟體。[註 13, 註 17] 同時也可在 WWW 上運作。整體來看，FDGC 的 metadata 已自成格局，有著錄內容的標準，有資料交換的架構，有語法查核器，整個架構已大致完備，事實上在美國已在網際網路上實際運作了。對 FDGC 有興趣的讀者，可參考以下二個站址：

1. <http://www.dwr.ehnr.state.nc.us/metadata/notes.htm>
2. <http://www.blm.gov/gis/nsdi.html>

## 二. 都柏林核心集

Dublin Core 這個 metadata 格式，[註 18] 是 1995 年 3 月由 Online Computer Library Center（OCLC）和 National Center for Supercomputing Applications（NCSA）聯合贊助的研討會，在邀請五十二位來自圖書館、電腦和網路方面的學者和專家，共同研討下的產物。目的是希望建立一套描述網路上電子文件特色的方法，來協助資訊檢索。因此在研討會的報告中，將 metadata 定義為 "resource description"。

很明顯的由於 WWW 的盛行，網際網路的無遠弗屆，NII 的提倡，和 DL 的興起，使得資訊檢索已逐漸走向以網路作業環境為主的操作。這時檢索的主體對象是電子檔案（file）或文件（document），而電腦檔案的類型非常複雜且善變。另一方面，參與描述資料特色（即製作目錄）的人，也擴大到各式專業和非專業的人士，但是圖書館所使用的 MARC，顯然對非專業的圖書館人士是過於複雜。長此以往，即使 MARC 自豪於其著錄品質的嚴謹，也將由於曲高和寡而逐漸走向沒落之途。因此與會

者認為須發展一套較簡易的描述資料特性方法，來揭示電子文件的內容或其他特質，以支援資訊檢索。所以研討會的中心問題是 [註 18]

"How can a simple metadata record be defined that sufficiently describes a wide range of electronic objects?"

根據研討會的報告，Dublin Core 這個 metadata 格式，其處理對象將祇限於類文件物件 (document-like objects, DLO)，這是與會者一方面覺得有必要先暫時縮小問題所涉及的層次，一方面認為文字型態的電子檔案，仍是網路文件上的大宗。那何謂 DLO 呢？簡言之，是可用類似描述傳統印刷文字媒體方式，加以描述的電子檔案。同時為進一步縮小問題的範圍，許多電子文件在商業應用上的重要課題，雖然在電子圖書館的文獻中已被熱烈討論和界定為必備的功能，如計費功能，也不列入此 metadata 範疇。由這種種的限制看來，似乎可以說，Dublin Core 所扮演的角色是 USMARC 的網路節縮版。另外報告中也特別將地理資料排除在外，已隱然承認 FDGC 的 Digital Geospatial Metadata 在地理資料方面的地位。

基於與會者認為沒有任何單一的 metadata 格式，足以適用於任何作業環境的認知，他們主張先建立一套描述資料的最小核心資料項。因此 Dublin Core 的設計原理是使此 metadata 的資料項，同時擁有意義明確、彈性和最小規模三種特色。在設計上所秉持的原則是：內在本質原則、易擴展原則、語法獨立原則、無必須項原則、可重覆原則、和可修飾原則。以下是它們的簡要敘述：[註 18]

1. 內在本質原則 (Intrinsicality)：祇描述跟作品內容和實體相關的特質，例如主題 (subject) 屬於作品的內在本質。但是，收費和存取規定則屬於作品的外在特質，原則上不屬於核心資料項，將透過其他機制來加以處理。
2. 易擴展原則 (Extensibility)：應允許地區性資料以特定規範的方式出現，也應保持 metadata 日後易擴充的特性，以及保有向後相容的能力。
3. 語法獨立原則 (Syntax-Independence)：在此 metadata 成熟前，將盡量避免制定特定語法。
4. 無必須項原則 (Optionality)：所有資料項都是可有可無，以保持彈性和鼓勵非專業人士參與製作。
5. 可重覆原則 (Repeatability)：所有資料項均可重覆。
6. 可修飾原則 (Modifiability)：資料項可用附加限定語 (qualifier) 來進一步修飾其意義。

正如在報告中反覆多次強調的，研討會的目標是發展一個簡單有彈性，且非專業人員也可輕易了解和使用的資料描述格式，所以 Dublin Core 祇規範那些在大多數情況下，必須提及的資料特性，總共有 13 個資料項，在報告中對這些資料項有詳細的列表、解釋和範例說明。[註 18] 在此我們以扼要方式列表，使讀者對其整體架構有清楚的認識。

資料項一. 主題 (Subject)：作品所屬的學術領域。

例子：Subject = Digital Geospatial Metadata

資料項二. 題名 (Title)：作品名稱。

例子：Title = Geospatial Support Staff Metadata Tutorial

資料項三. 著者 (Author) : 作品的創作者或組織。

例子 : Author = Abeyta, Carolyn

資料項四. 出版者 (Publisher) : 負責發行作品的組織。

資料項五. 其他參與者 (OtherAgent) : 對作品創作有貢獻的相關人或組織。

資料項六. 出版日期 (Date) : 作品公開的日期。

例子 : Date=1995/05

資料項七. 資料類型 (ObjectType) : 作品的類型或所屬抽象範疇, 可用來幫助資料檢索。

例子 : ObjectType = tutorial

資料項八. 資料格式 (Form) : 告知檢索者在使用此作品時, 所須的電腦軟體和硬體設備。

例子 : Form=html

資料項九. 識別代號 (Identifier) : 字串或號碼可用來唯一標示此作品。

例子 : Identifier (scheme = URL) =http://www.blm.gov/gis/meta/barney/tut\_met1.html

資料項十. 關連 (Relation) : 與其他作品 (不同內容範疇) 的關連, 或所屬的系列和檔案庫。

例子 : Relation (type = ContainedIn) (identifier=URL) = http://www.blm.gov/

資料項十一. 來源 (Source) : 作品從何處衍生而來 (同內容範疇)。

資料項十二. 語言 (Language) : 作品所使用的語言。

例子 : Language = English

資料項十三. 涵蓋時空 (Coverage) : 作品所涵蓋的時期和地理區域。

檢視上述架構, 可清楚看到某些資料項, 是針對電腦作業環境而設計的, 如資料項八 (資料格式, Form), 其他如資料項七 (資料類型, ObjectType)、資料項十 (關連, Relation) 和資料項十一 (來源, Source), 也和網路或電子作業環境有密切關係。同時此資料描述格式可說是非常簡單和易使用, 幾乎所有資料項都有自我解釋的功能, 大部份人在短時間內就知如何使用, 比起 MARC 的深奧難懂, 此 metadata 在鼓勵非專業人士的自行著錄收藏資料, 以及縮短製作時間和成本上, 可說是成功的。

由於 Dublin Core 祇是一套最小核心資料項, 須要有一個機制來與現存的其他較完整描述格式, 如 USMARC 和 FDGC 的 Digital Geospatial Metadata 等, 來做資料的對照和轉換。使 Dublin Core 的資料, 能在最小成本下, 轉換成更完備的描述格式。在這方面, 有關 USMARC 和 Dublin Core 的連結已在發展中, 有興趣的讀者請參考 [註 19], 在此文中, 對如何將 Dublin Core 的 13 個資料項, 對映到 USMARC 的相關欄位, 如 Dublin Core 的 Subject 資料項應對映到 USMARC 的欄位 653 或欄位 650 等, 有詳盡的分析和解說。

### 三. 資源特性統一描述格式

Uniform Resource Characteristics (URC) 相較前面所介紹的 FDGC 的 Digital Geospatial Metadata 和 OCLC/NCSA 的 Dublin Core 二個 metadata 格式，URC 可說是還未成形和正在發展中的一個 metadata 標準，其原始的功能，主要是用來銜接現在 WWW 上的主角 -- 資源位置統一描述器 (Uniform Resource Locator, URL) 和未來的主角 -- 資源名稱統一描述格式 (Uniform Resource Name, URN)。但由於其涵蓋範圍最廣，將來對資訊檢索的影響也最大。以下我們將依序介紹 URL 的結構及其缺點、URN 的構造、和 URC 的發展情況。

WWW 是起源於在 CERN 中的一個增進高能物理學者間互動的實驗計畫，[註 20] 但 WWW 藉著網際網路的無遠弗屆，易寫作的 HTML 格式和親善的使用介面，在短時間內形成一股風潮席捲全球，可是在急遽擴張後，許多 WWW 的問題逐漸浮現，尤其是在應用上佔重要地位的超連結 (hyperlink)，對資料管理者和使用者均造成極大的困擾，那就是常見的破裂連結 (broken link)，這是因為超連結是建基於 URL。那何謂 URL 呢？URL 最常見的格式是由以下四個部份所組成：(完整的 URL 格式請參考 [註 21])

1. 前置詞 (PrePrefix) : "URL:"。
2. 操作模式 (Scheme) : 表明所使用的通訊傳輸模式，同前支援大部份網際網路使用的傳輸協定，如 ftp、http、gopher、news、wais 等，並以 ":" 與其後部份分隔，如 "http:"。
3. 網路站址 (Internet Protocol Parts) : 以 "/" 開始，以 "/" 結束，如 "http://www.w3.org/"。
4. 路徑 (Path) : 檔案所在目錄。如 "http://www.w3.org/hypertext/WWW/Glossary"。

因此一個完整的 URL 如 "URL:http://www.w3.org/hypertext/WWW/Glossary"。URL 的缺點可概述如下：[註 22]

1. 檔案不能隨意移動所在位置 (目錄)，否則文件 (或檔案) 間的連結即斷裂。
2. 檔案所在主機，必須隨時保持在可運作狀態。
3. 易使資料索求集中於少數著名的伺服器，對其系統和網路造成太大負擔。

URL 的這些缺點，主要是來自它直接標明檔案所在位置 (主機站址 + 路徑)，使用這種直接指引的方式，造成其應用上的缺乏彈性。

為了解決這種直接位置連結表示法，對維持一個穩定的全球資訊系統所帶來的困擾，網際網路工程任務小組 (Internet Engineering Task Force, IETF) 成立了一個資源識別統一描述器 (Uniform Resource Identifier, URI) 工作小組，並決定以間接指引的 URN [註 23] 來取代直接指引的 URL，做為超連結的表示法。那何謂 URN 呢？URN 的整個結構如下：[註 24]

<URN : Encoding\_Scheme : Authority\_Id : Opaque\_string>

1. 前置詞 (PrePrefix) : "URN:"。
2. 字元集編碼模式 (Encoding\_Scheme) : 如 ASCII。
3. 權威機構代碼 (Authority\_Id) : 如 ISBN 和 IANA (Internet Assigned Number Authority)。



4. 識別字串 (Opaque\_string) : "merit.edu : 1929642"

因此一個完整的 URN 如 "URN:ASCII:IANA:merit.edu:1929642"。

URN 設計的基本目的，在確保世界上沒有二個不同文件（或檔案）會有相同的 URN 碼，因此如圖書館的登錄（流水）號一樣，具有唯一性的特質。第二個特質是 URN 碼的形式，看起來似乎跟檔案位置沒有關連，所以顯然是屬於間接指引的。因此若要使用 URN 來指示位置，必須有一個中介東西來將 URN 和檔案位置（或 URL）連結對映起來。在 IETE 的 URI 工作小組設計中，URC（資源特性統一描述格式）即是這個中介東西，而連結對映 URN 到 URL 所須的資訊，即存在 URC 中。

雖然 URC 的原始目的是用來連結 URN 和 URL，但人們很快便發覺 URC 也可用在其他用途，例如若將基本的書目資料，如題名、著者、出版者等，也放入 URC 中，則在網際網路和 WWW 上，也能做書目查詢。另外若將電子簽名（digital signature）的資訊，加入與收藏資料相對映的 URC 中，則系統可提供安全認證（authentication）的功能，因為資料接收者可利用電子簽名，來查證所接收的資料，是否的確來自指定機構或指定人。另一個可能的應用是評價認證（Seals of Approval, SOAP），[註 25] 這是將資料所屬領域專家對此資料的評價資訊放入 URC 中，而此種專業的評價資訊，在檢索時可作為過濾資料的一個標準。簡言之，由於 NII 的推動和電子圖書館的興起，我們如今可以在全球網路上，提供快速和多樣性的服務給資訊檢索者，但這些服務需參考一些關於資料特性的資訊，而這些資訊不論是傳統的基本書目資訊，或是新創造的，都可放在 URC 中，用來協助資訊檢索。

由於 URC 所扮演角色的多樣性，正如 1995 OCLA/NCSA metadata workshop 中與會者的認知，幾乎沒有單一的一組屬性（或資料項），可適用於任何情況，[註 26] 再加上 URC 還是一個正在發展中的資料描述格式，因此在此階段，學者的研究焦點大都集中在釐清 URC 的需求，為了使讀者對未來 URC 可能的面貌有一大略的了解，以下我們參考 [註 27] 一文，將 URC 相關的需求整理如下。在有關 URC 的內容和格式方面：

1. 在不須修改下，能透過現行通用於網際網路上的傳輸標準，如 TCP、FTP 等來傳輸 URC 資料。
2. 須有一個適合與人互動的表現格式，以方便使用者來閱讀或輸入 URC 資料。
3. 格式應簡單，非專業人士亦可輕易自行著錄。
4. URC 中各資料項彼此間，無一定位置次序關係。
5. 能隨時加入新資料項，而不會影響到使用舊版本的應用程式，即能以逐漸增加的方式來擴大其功能和應用範圍。
6. URC 中若有隱藏式物件（object），則這些物件必須能自成一單位，即有自我解釋的功能或結構，使其他應用程式處理時，不會產生問題。
7. 應用程式碰到不認識的資料項，能加以忽略，繼續處理其他資料項。
8. 可利用已存在 URC 格式中部份資料項，來形成新的 URC 格式。
9. 隱含 URC 版本資訊，使應用程式能以適當方式，來同時處理不同版本的 URC。
10. 在可能情況下，能包含其他現存的 metadata 標準。

11. 允許個別的 URC 紀錄有各自的存活時間標記。
12. 允許多個電子簽名和評價認證，同時存在一個 URC 紀錄中，因此一個文件可有多個評價。

在 URC 伺服器 and 相關服務程式的功能和需求方面：[註 27]

1. URC 服務程式須能利用 URN 或 URL 來找出相對映的 URC。
2. URC 伺服器須能與不同能力的瀏覽器工作。
3. URC 伺服器能利用如版本、電子簽名、最後修改日期等資訊，來過濾所要的紀錄。
4. 能在兼顧安全標準的情況下，不影響到自由公開部份資訊的流通效率。
5. 能維持 URN--URC--URL 一貫的認證鏈，以避免安全漏洞的產生。
6. 至少能提供如 UNIX 系統的讀寫存取控制。
7. 具有可維護性。
8. 使用者能同時透過多個 URC 伺服器來找尋資料。
9. 無自動查詢轉送服務，以免產生大量回應，造成系統的負荷和使用者的負擔。
10. 有一定管道或方式，可讓使用者聯絡系統的管理者。

#### 四. 結論

本文主要是介紹三個與元資料 (metadata) 相關的發展，它們分別是 FGDC 制定的 Digital Geospatial Metadata、1995 年 OCLC/NCSA 研討會制定的 Dublin Core、和 IETE 的 URI 工作小組所負責發展的 URC。這三個 metadata 格式各有其著重點，如 FDGC 的 Digital Geospatial Metadata 著重在地理性資料方面，而 OCLC/NCSA 的 Dublin Core 則是針對基本書目資料而設計的。以下就不同面相來比較這三個 metadata 格式的異同。

1. 發展時間：Digital Geospatial Metadata 成形時間最早，公布於 1994 年 6 月 8 日；其次是 Dublin Core，公布於 1995 年 3 月；URC 則尚正在發展中。
2. 處理對象：Digital Geospatial Metadata 可說完全是針對地理性資料而設計的；Dublin Core 則處理所謂的類文件物件 (DLO)，基本上是類似印刷文件 (如書籍等) 的網路電子文件書目資料；URC 則最為廣泛，祇要是網路上的檔案均在其涵蓋範圍內。
3. 系統完備性：Digital Geospatial Metadata 發展的最為完備，有著錄內容的標準格式可遵循，有資料交換所軟體負責資料的查詢和交換，可在 WWW 下運行，也已實際上在網際網路上運作。Dublin Core 則祇有著錄格式，其他配合的軟體，和與 USMARC 的對照和轉換工作則正在發展中。URC 所涉及的二個主體，URL 可說是已發展完備，事實上已是 WWW 的重要支柱之一；另一個主體 URN 雖然尚未完全定型 (成為標準)，但已具雛型；至於 URC 本身則由於牽涉廣泛，尚在研議中，目前還在釐清所須需求的階段。
4. 格式詳盡程度：Digital Geospatial Metadata 雖然是祇限於地理性資料，但可說是完整的地理資料

描述格式。Dublin Core 祇規範了 13 個核心資料項，並非一完整的資料描述格式，必須要能與其他完整的格式連結，目前最主要的連結對象是指向 USMARC。URC 雖尚未成形，但由現存規劃和討論文獻看來，其所含資料項必遠超過 Dublin Core，至於專精程度可能不及於 Digital Geospatial Metadata，所以最可能的詳盡程度是介於二者之間。

5. 接受程度：Digital Geospatial Metadata 有美國聯邦政府支持，已儼然成為地理性資料方面的官方標準。Dublin Core 的格局過小，且目前似乎祇有部份圖書館界的人士支持，將來命運未知。URC 雖然尚在發展中，但由於 URL 已隨著 WWW 的盛行而被廣為接受，又有 IETE 的支持，將來被全球網路界接受的可能性極高。

6. 作業環境：三者都是建基於網路作業環境，即資訊檢索是以全球網路為範圍，其他處理功能的設計，也以網路為其假想作業環境。

由這三個 metadata 的發展可知，metadata 是晚近才興起的一個概念，其興起的背景來自網路的普及，以及網路成為資訊檢索的主要作業環境。所以本文所介紹的三個 metadata 都強調網路作業環境，也都是協助在網路上做檢索的資料描述格式。

## 註釋

註 1：NASA, Directory Interchange Format Manual, Version 1.0, July 13 1988. NSSDC/WDC-A-R&S 88-19.

註 2：E.P. Shelley and B.D. Johnson, "Metadata: Concepts and Models," in Proceedings of the Third National Conference on the Management of Geoscience Information and Data (Adelaide, Australia: Australian Mineral Foundation, 1995), pp 1-5.

註 3：Clint Matthews, Chin-Chih Chang, Jim Hull, and Sejin Choi, "Metadata," <<http://www.seic.okstate.edu/gis/metadata.html>> (16 Feb. 1996).

註 4：IFLA, "UNIMARC: An Introduction," <[www.nlc-bnc.ca/ifla/VI/3/P1996-I/unimarc.htm](http://www.nlc-bnc.ca/ifla/VI/3/P1996-I/unimarc.htm)> (26 Sept. 1996).

註 5：徐小鳳，自動化書目的資訊服務（台北市：學生書局，民 73），頁 43。

註 6：李德竹，我國圖書館自動化系統線上目錄及其顯格式之研究，圖書館學刊 7 期（民 80 年 11 月），頁 4。

註 7：NIST, "National Information Infrastructure (NII): General Information," 1996, <<http://nii.nist.gov/nii/niiinfo.html>>.

註 8：吳政叡，淺談電子圖書館的發展，圖書館學刊 25 期（民 85 年 6 月），頁 23-31。

註 9：National Science Foundation, Department of Defense Advanced Research Projects Agency, and National Aeronautics and Space Administration, "Digital Library Initiative," in NSF 93-141, 1993.

註 10：Infoseek Corp., "Infoseek Home Page," <<http://www.infoseek.com/>> (18 Feb. 1996).

- 註 11 : FGDC, "Content standards for digital geospatial metadata -- FGDC," 1994, <<http://fgdc.er.usgs.gov/fgdc.html>>.
- 註 12 : William Clinton, Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure (Washington D.C.: The White House, 1994).
- 註 13 : FGDC, "Making your Spatial Data Discoverable: Building an NSDI Clearinghouse Node," <<http://www.fgdc.gov/clearinghouse/itraining/toc.html>> (26 Sept. 1996).
- 註 14 : Brandon Plewe, "GeoWEB Site Home Page," 1995, <<http://wings.buffalo.edu/geoweb>>.
- 註 15 : Peter Schweitzer, "Metadata Validation Service," <<http://www-mel.nrlmry.navy.mil/meta-val.html>> (26 Sept. 1996).
- 註 16 : Carolyn Abeyta, "Geospatial Support Staff Metadata Tutorial," 1995,<[http://www.blm.gov/gis/meta/barney/tut\\_met1.html](http://www.blm.gov/gis/meta/barney/tut_met1.html)>.
- 註 17 : BLM, "Metadata and WWW Mapping Home Page," 1996 <<http://www.blm.gov/gis/nsdi.html>>.
- 註 18 : Stuart Weibel, Jean Godby, Eric Miller, and Ron Daniel, "OCLC/NCSA Metadata Workshop Report," 1995, <[http://www.oclc.org:5047/oclc/research/publications/weibel/metadata/dublin\\_core\\_report.html](http://www.oclc.org:5047/oclc/research/publications/weibel/metadata/dublin_core_report.html)>.
- 註 19 : Rebecca Guenther, "Mapping the Dublin Core Metadata Elements to USMARC", 1995, <<gopher://marvel.loc.gov/00/.listarch/usmarc/dp86.doc>>.
- 註 20 : Tim Berners-Lee and Robert Cailliau, "WorldWideWeb: Proposal for a HyperText Project," 1990, <<http://info.cern.ch/hypertext/WWW/Proposal.html>>.
- 註 21 : T. Berners-Lee, L. Masinter, and M. McCahill, "Uniform Resource Locators (URL)," 1994, <<ftp://ds.internic.net/rfc/rfc1738.txt>>.
- 註 22 : Ron Daniel Jr., "A Global Distributed Directory Service for the World Wide Web," <[http://www.acl.lanl.gov/URI/URC\\_proposal/index.html](http://www.acl.lanl.gov/URI/URC_proposal/index.html)>.
- 註 23 : K. Sollins and L. Masinter, "Requirements for Uniform Resource Names," 1994, <<ftp://cnri.reston.va.us/internet-drafts/draft-ietf-uri-urn-req-00.txt>>.
- 註 24 : 陳亞寧，從編目作業探討網路資源的整理（台北市：文華，民 83），頁 111。
- 註 25 : J. Rhine, "Interpedia Homepage," 1994, <<http://www.hmc.edu/interpedia/index.html>>.
- 註 26 : Ron Daniel Jr., "An SGML-based URC Service," 1995, <<http://www.acl.lanl.gov/URI/URCspeg/spec.asc>>.
- 註 27 : Ron Daniel Jr., "URC Scenarios and Requirements," 1995, <[http://www.acl.lanl.gov/URI/Scenarios/scenarios\\_3.asc](http://www.acl.lanl.gov/URI/Scenarios/scenarios_3.asc)>.

