

都柏林核心集第五次研討會的最新發展

都柏林核心集是一個簡易的元資料，用來支援著者著錄和描述網路上電子文件的特色，以協助資訊的檢索與過濾。1997年10月的第五次研討會後，都柏林核心集已經開始進行標準化的過程，使用都柏林核心集和計劃開發中的系統，正在全世界各地蓬勃發展中。 吳政叡

元資料 (Metadata) 為描述資料的資料，主要是描述資料屬性的資訊，用來支持如指示儲存位置、資源尋找、文件紀錄、評價、過濾等的功能。以圖書館的角度來看，就其本義和功能而言，元資料可說是電子式目錄，因為編製目錄的目的，即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。因此元資料是用來揭示各類型電子文件或檔案的內容和其他特性，其典型的作業環境是電腦網路作業環境。[註1] 換言之，元資料是因應現代資料處理上的二大挑戰而興起的：一是電子檔案成為資料的主流，另外一個是網路上大量文件的管理和檢索需求。

都柏林核心集為元資料的一種，是1995年3月由國際圖書館電腦中心 (OCLC) 和 National Center for Supercomputing Applications (NCSA) 所聯合贊助的研討會，在邀請五十二位來自圖書館、電腦、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。[註2] 都柏林核心集的15個基本項目有：主題和關鍵詞 (Subject)、題名 (Title)、著者 (Creator)、簡述 (Description)、出版者 (Publisher)、其他參與者 (Contributors)、出版日期 (Date)、資源類型 (Type)、資料格式 (Format)、資源識別代號 (Identifier)、關連 (Relation)、來源 (Source)、語言 (Language)、涵蓋時空 (Coverage)、版權規範 (Rights)。[註3]

在都柏林核心集的最新發展方面，第五次研討會已經於1997年10月6-8日在芬蘭的赫爾辛基舉行，根據澳洲國家圖書館的一位與會者--Bemal Rajapatirana的報告，與會者達成了如下的幾項共識：[註4]

(一) 加快標準化的腳步—由於都柏林核心集的15個基本項目架構，自第四次研討會以來已普遍獲得認同，同時都柏林核心集也得到世界各國很多研究者的肯定，並且嘗試建造系統，此時若無一定的標準來遵循，將使系統的建造者無所適從和系統的更改頻繁。因此基於都柏林核心集已趨成熟的共識，決定推派代表撰寫 RFC 的草案，呈交給 IETF 進行標準化的過程。

(二) 區分簡單和複雜兩種都柏林核心集格式—簡言之，所謂簡單 (simple) 和複雜 (complex) 格式的區分，一般而言主要是以有無使用任何修飾詞作為標準來劃分的。

(三) 語法上採用 HTML 和 RDF 格式為主—HTML 的格式目前是使用 4.0 版本，其寫法如下：[註5]

```
(1) <META NAME="DC.Subject"  
    SCHEME="LCSH"  
    LANG="EN"
```

CONTENT="Computer Cataloging of Network Resources">

(2) <META NAME="DC.Date.Created"

SCHEME="ISO8601"

CONTENT="1998-03-05">

由上面的例子可知，都柏林核心集的語言和架構修飾詞，是分別直接利用HTML 4.0中的LANG 和 SCHEME來表示，至於次項目修飾詞則是放在項目（或欄位）名稱之後，如DC.Date.Created。

(四) 成立工作小組—針對一些尚未有定論的議題，組成工作小組進行研討，主要有

(1) 內容或格式尚未有定論的基本項目，如Date、Relation、Rights Management等項目。

(2) 修飾詞。

(3) 特殊性議題，如都柏林核心集和Z39.50間的互換。

(五) 訂定次項目（或類別）修飾詞的制定原則

(1) 與基本項目一致，都是可省略的選擇項。

(2) 次項目須能進一步協助詮釋項目的內容。

在這次研討會中，除了已開發系統的介紹外，也有一些正在籌建中的都柏林核心集相關系統的宣佈，以下是它們的簡介：[註6]

(一) 丹麥政府決定自西元1997年起將所有政府的出版物上網，系統的主要規格之一，是採用都柏林核心集來描述文件和協助查詢。

(二) 荷蘭國家圖書館將發展一種新的全球資訊網服務，系統的主要做法是要在所有已蒐集的網頁中，加入都柏林核心集的資料，新的網頁將要求提供者先自行加入都柏林核心集的資料後再送呈，將來荷蘭國家圖書館的搜尋引擎會利用這些元資料來協助檢索。

(三) 英國的UKOLN正在推行一個名為BIBLINK的計劃，在出版社和國家書目中心間建立一條網路通訊管道，來直接交換書籍紀錄和資訊，這套系統是使用都柏林核心集作為其基本的格式。

(四) 在商業的應用上，一個稱為STARTS的協定正在發展中，它可以辨識網頁中的元資料，來協助使用者過濾和排比查詢的結果，STARTS已決定包含都柏林核心集。

綜觀以上的發展，顯示都柏林核心集已漸成熟和廣受肯定，以系統的實作而言，歐洲和澳洲可說是居於領先的地位，歐洲較注重都柏林核心集在圖書館相關服務上的應用，澳洲的DSTC則較偏重都柏林核心集在WWW相關服務上的應用。

以下介紹由澳洲的「分散式系統技術中心」（DSTC）所推動的系列研究計畫。澳洲的分散式系統技術中心（DSTC）是在1992年由十個機構參與設立，主要宗旨是研究建設一個全球分散式資訊系統所需的技術，並取得澳洲政府和其他參與機構七年（1992-1999）總共6千萬美元的研究經費贊助，因此正致力於推動一些與資訊管理和檢索相關的計畫，而分散式系統技術中心（DSTC）所

推動的這一系列計畫，是目前為止在都柏林核心集的應用和實作系統上，最著名和成功的系列研究計畫之一。其中三個跟都柏林核心集最密切相關的計畫為TURNIP [註7]、HotOIL [註8]、MetaWeb [註9]。

DSTC所推動的這一系列關於全球資訊存取（Global Information Access）的研究計畫，都是採用都柏林核心集做為描述資源的格式，因此是以都柏林核心集來協助使用者在WWW和網際網路上搜尋資料，並以都柏林核心集來和其他的檢索技術（如URN）、網路協定（如HTTP和Z39.50）等相結合，可以說是將都柏林核心集的檢索效用發揮的淋漓盡致。

DSTC在使用者檢索介面上的設計（如HotOIL研究計畫）非常值得借鏡，不但可以自動替使用者將查詢送至各個搜尋引擎和資料庫，還會將傳回的結果先行整理，大大的減輕了使用者過濾的負擔，這是作者所見過最好的介面設計之一。另外一個DSTC檢索系統的優良設計，是直接介面的設計上，應用了都柏林核心集的欄位來縮小查詢的範圍（如MetaWeb研究計畫），這將有助於檢索效能的提昇。

有鑒於元資料對資料著錄和檢索的重要性，作者建立了一個相關的實驗系統—元資料實驗系統（Metadata Experimental System，簡稱MES，網址：<http://140.136.85.194/mes> 或 <http://mes.lins.fju.edu.tw/mes>），作者建立MES目的，除了是讓讀者透過這個系統，對元資料及其未來的可能運作方式，有更具體的認知外；也希望利用此一實驗系統，來測試和驗證元資料的功能和效用，例如都柏林核心集這種簡易的資料描述格式，是否如制定者們所預期的，足以滿足大部分網路文件著錄和檢索的需求。MES是一開放性的實驗系統，歡迎任何人上站著錄自己的網頁或文件，以供他人查詢和檢索。

截至1997年10月止，祇包含兩種元資料類型—都柏林核心集（Dublin Core）和IETF正在規劃中的URI架構（包含URN，URL，URC）。[註10] 其中尤以都柏林核心集的功能最為完整，可達成直到第四次研討會的要求，採用與HTML 2.0完全相容的格式，作者將於近期內修改系統以配合HTML 4.0格式的使用。

元資料實驗系統有四個主要的特色：同時提供著錄和檢索兩種功能、開放性設計、使用URN作為資源（或文件）的唯一識別碼、提供模糊檢索功能。[註11] 由於元資料實驗系統的功能眾多，因此作者將各項作業依功能和類別分為若干的子系統，目前有以下六個子系統：註冊子系統、都柏林核心集子系統、查詢子系統、URL子系統、輔大圖書館資料查詢子系統、評價認證（SOAP）子系統。為了便利使用者充分利用多視窗來進行平行作業，系統的設計是將每個子系統單獨開一個視窗來處理，因此使用者在離開子系統時，請直接將該子系統的視窗關閉即可。因為都柏林核心集子系統較為複雜，因此又下分為兩個次系統—著錄次系統和查詢次系統。[註12]

註釋

註 1：吳政叡，「從元資料看未來資料著錄的發展趨勢」，*資訊傳播與圖書館學* 3 卷 2 期（民 86 年 12 月），頁44-45。

註 2：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>, p. 2.

註 3：同註1，頁46-47。

註 4：B. Rajapatirana, "The 5th Dublin Core Metadata Workshop: a report and observations," 2 Dec. 1997, <<http://www.nla.gov.au/nla/staffpaper/helsinki.html>>.

註 5：吳政勸，「元資料實驗系統和都柏林核心集的發展趨勢」，國立中央圖書館臺灣分館館刊 4 卷 2 期（民 86 年 12 月），頁17-18。

註 6：同註4，頁3-4。

註 7：R. Iannella and H. Sue, "Basic URN Service (BURNS)," <<http://www.dstc.edu.au/RDU/TURNIP/burns.html>>, p. 3.

註 8：N. Ward, "HOTOIL," <<http://www.dstc.edu.au/BDU/APAP/HotOIL/HotOIL.html>>, (26 Jan. 1998).

註 9：D. Campbell, "The MetaWeb Project," 22 January 1998, <<http://www.dstc.edu.au/RDU/MetaWeb/>>.

註10：吳政勸，「三個元資料格式的比較分析」，中國圖書館學會會報57期（民85年12月），頁41。

註11：同註1，頁48-49。

註12：同註5，頁20-22。