

都柏林核心集到中國機讀編目格式的轉換對照表

吳政叡 (Cheng-Juei Wu)

輔仁大學圖書資訊系專任副教授

Associate Prof.

Department of Library & Information Science

Fu-Jen University

E-mail: lins1022@fujens.fju.edu.tw

中文摘要

由於國外都柏林核心集已經漸趨成熟，也開始被廣泛使用在一些與資訊檢索相關的應用系統中，所以已經有都柏林核心集到UNIMARC和USMARC的轉換對照表。為了推廣都柏林核心集在國內的應用，本文介紹都柏林核心集和國內圖書館使用最廣泛的機讀編目格式—中國機讀編目格式（Chinese MARC）間的轉換，此外本文也詳細介紹了都柏林核心集的15個欄位和修飾詞的內容和應用方式。

Mapping Dublin Core to Chinese MARC

Abstract

This work discusses how to map the Dublin Core to the Chinese MARC that is used by most of the libraries in Taiwan. The Dublin Core is a relatively simple metadata for networked resources and is becoming increasingly popular among information retrieval related communities in recently years. The mapping table and a detailed analysis between these two formats are given in this work. In addition, the usage of each element and the three kinds of qualifiers of the Dublin Core is given as well.

關鍵字：元資料，中國機讀編目格式，都柏林核心集，Metadata，Chinese MARC，Dublin Core。

一、前言

都柏林核心集為元資料的一種 [註1]，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] 因此在研討會的報告中，將元資料定義為資源描述（Resource Description），而研討會的中心問題是

如何用一個簡單的元資料記錄來描述種類繁多的電子物件？

根據研討會的報告，都柏林核心集處理的對象，將祇限於「類文件物件」（Document-Like Objects，簡稱DLO），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。同時因為研討會的目標是發展一個簡單有彈性，且各種專業人員也可輕易了解和使用的資料描述格式，所以都柏林核心集祇規範那些在大多數情況下，必須提及的資料特性。

國際圖書館電腦中心（OCLC）主導創設都柏林核心集的目的，是希望能一方面解決MARC在應付網路文件上的困境，另一方面能有一套簡單的資源描述格式，讓眾多非圖書館的專業人士來使用，以最少成本來解決網路文件快速增加的問題，因此都柏林核心集是定位在簡單的資源描述格式，但提供一個基本資料庫，作為各種專業進一步加工處理的基礎。由於現在圖書館界的自動化系統，主要仍以機讀編目格式（MARC）為主，在實務的考量下，須要製作一套轉換方法，將都柏林核心集轉換成機讀編目格式，以便圖書館界來加以利用，甚而做進一步的加工處理。因此本文介紹都柏林核心集和國內圖書館使用最廣泛的機讀編目格式—中國機讀編目格式（Chinese MARC）間的轉換。

根據王振鵠教授在中國機讀編目格式第三版中的序言描述，中國機讀編目格式的發展歷程，首先是在民國 69 年 5 月成立「中文機讀編目格式工作小組」進行研定，主要是依據國際機讀編目格式（UNIMARC-1980）、美國國會圖書館書目機讀編目格式（MARC Formats for Bibliographic Data-1980）、ISO 2709 格式來制定，於民國 70 年 1 月正式刊行中國圖書機讀編目格式第一版，同年 7 月第二版修訂出版。隨後又修訂加入連續性出版品、地圖、音樂、視聽資料等，在民國 71 年發行新版，並且更名為中國機讀編目格式（Chinese MARC），民國 73 年又根據國際圖書館協會聯盟 1983 年編訂的 UNIMARC Handbook 修訂出版第二版，第三版則於民國 77 年出版，第四版甫於民國86 年 6 月修訂出版。

二、都柏林核心集欄位介紹

1997年10月公布的資料著錄項目列表如下：[註3]

(一) 主題和關鍵詞 (Subject)：作品所屬的學術領域，控制語彙用 scheme 註明出處如 LCSH，亦可包含分類號如杜威十進分類號 (Dewey Decimal Number)。

例子：Subject = 都柏林核心集。

(二) 題名 (Title)：作品名稱。

例子：Title = 都柏林核心集與元資料實驗系統。

(三) 著者 (Creator)：作品的創作者或組織。

例子：Creator = 吳政叡。

(四) 簡述 (Description)：文件的摘要或影像資源的內容敘述。

例子：Description = 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。

(五) 出版者 (Publisher)：負責發行作品的組織。

例子：Publisher = 松崗電腦圖書資料股份有限公司。

(六) 其他參與者 (Contributors)：除了著者外，對作品創作有貢獻的其他相關人士或組織。

[註：如書中插圖的製作者。]

(七) 出版日期 (Date)：作品公開發表的日期，建議使用如下格式—YYYY-MM-DD 和 參考下列網址：<http://www.w3.org/TR/NOTE-datetime>。在此網頁中共規範有六種格式，都是根據國際標準日期暨時間格式—ISO(國際標準組織)8601制定而成，是ISO 8601的子集合 (subset)，現在列舉和解說如下以供參考：[註 4]

(1) Year (年) -- YYYY。

例子：1997 (西元1997年)。

(2) Year and Month (年、月) -- YYYY-MM。

例子：1997-09 (西元1997年9月)。

(3) Complete date (完整日期) -- YYYY-MM-DD。

例子：1997-09-07 (西元1997年9月7日)。

(4) Complete date plus hours and minutes (完整日期加時、分) --
YYYY-MM-DDThh:mmTZD

[註：T用來隔開日期和時間，TZD表示本地時間和國際格林威治時間的差距(時間差)。]

例子：1997-09-07T19:05+08:00 (西元1997年9月7日台灣下午7點5分，而台灣所屬的中原標準時區與國際格林威治時間差8小時)。

(5) Complete date plus hours, minutes, and seconds (完整日期加時、分、秒) -- YYYY-MM-DDThh:mm:ssTZD

例子：1997-09-07T19:05:25+08:00 (西元1997年9月7日台灣下午7點5分25秒)。

(6) Complete date plus hours, minutes, and seconds (完整日期加時、分、秒) -- YYYY-MM-DDThh:mm:ss.sTZD

例子：1997-09-07T19:05:25.25+08:00 (西元1997年9月7日台灣下午7點5分25又1/4秒)。

(八) 資源類型 (Type)：作品的類型或所屬的抽象範疇，例如網頁、小說、詩、技術報告、字典等，建議參考下列網址：<http://sunsite.berkeley.edu/Metadata/types.html>。在上述網頁中將作品的類型粗分成六種，現在列舉和解說如下：[註 5]

(1) Text (文字) -- 作品的內容主要是文字(可夾帶影像、地圖、表格等)，例如書籍、文集、技術報告、小冊子等。

例子：<META NAME="DC.type" CONTENT="Text">。

(2) Image (影像) -- 相片、圖形、動畫、影片等。

(3) Sound (聲音) -- 各式各樣的聲音，例如演講、音樂等。

(4) Software (軟體) -- 可執行的程式(二進制檔)和程式的原始檔，但不包括各種互動式應用程式。

(5) Data(資料)-- 各種文字或數據資料的集合體，例如地理資料、書目記錄、統計數據、遙測資料等。

(6) Interactive (互動式應用) -- 設計給一個或多個使用者的互動式應用，例如遊戲軟體、線上聊天服務、虛擬實境等。

以上的六種類型又以第一種類型 (Text) 最為繁複，可再細分如下：

(1) Abstract (摘要) -- 其他文件的簡要敘述。

例子：<META NAME="DC.type" CONTENT="Text.Abstract">。

(2) Advertisement (廣告) -- 如徵人啟事。

(3) Article (論文)。

(4) Correspondence (書信) -- 可再細分為討論、電子郵件、信件、明信片四類。

例子：<META NAME="DC.type" CONTENT="Text.Correspondence.Email">。

(5) Dictionary (字典)。

(6) Form (表格)。

(7) Homepage (WWW首頁)。

(8) Index (索引)。

(9) Manuscript (手稿)。

(10) Minutes (會議紀錄)。

(11) Monograph (專論) -- 如書籍。

(12) Pamphlet (小冊子)。

(13) Poem (詩)。

(14) Proceedings (會議論文集)。

(15) Promotion (促銷文件)。

(16) Serial (連續性出版品) -- 可再細分為期刊、雜誌、報紙、時事通訊四類。

(17) TechReport (技術報告)。

(18) Thesis (學位論文) -- 可再細分為碩士、博士二類。

例子：Type = Text.Dictionary。

例子：Type = 文字.技術報告。

(九) 資料格式 (Format)：告知檢索者在使用此作品時，所須的電腦軟體和硬體設備，例如 text/html (MIME格式)、ASCII、Postscript (一種印表機通用格式)、可執行程式、JPEG (一種通用圖像格式)。亦可擴展至非電子文件，例如book (書本)、叢書、期刊。

例子：Format = text/html。

例子：Format = 叢書。

(十) 資源識別代號 (Identifier)：字串或號碼可用來唯一標示此作品，例如URN、URL、ISSN、ISBN等。

例子：Identifier (scheme = URL) = <http://mes.lins.fju.edu.tw/>。

(十一) 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 = <http://www.blm.gov/>。

(十二) 來源 (Source)：作品從何處衍生而來 (同內容範疇)，例如莎士比亞的某個電子書出自那個紙本。

(十三) 語言 (Language)：作品所使用的語言，建議遵循 RFC 1766 的規定，請參考下列網址：<http://ds.internic.net/rfc/rfc1766.txt>，RFC 1766 是使用 ISO 639的二個字母的語言代碼。[註 6]

例子：Language = en。(English) [註 7]

(十四) 涵蓋時空 (Coverage)：作品所涵蓋的時期和地理區域。

(十五) 版權規範 (Rights)：作品版權聲明和使用規範。可能值如下：[註 8]

(1) 空白 (Null)：無特別聲明，使用者須自行參考其他來源。

(2) 無限制 (No Restriction on Reuse)：可複製再傳播。

(3) 參考處 (URI or Other Pointer)：使用的相關說明，在所指定的出處。

例子：Rights = 無限制。

其中某些資料項，是針對電腦作業環境而設計的，如資料格式 (Format)，其他如資料類型 (Type)、關連 (Relation)、來源 (Source) 等，也和網路或電子作業環境有密切的關係。同時此資料描述格式可說是非常簡單和容易使用，幾乎所有的資料項都有自我解釋的功能，大部份人在短時間內就知道如何來使用。

三、修飾詞 (Qualifier)

為了豐富都柏林核心集的內涵和擴大其應用範圍，1997年3月在澳洲坎培拉 (Canberra) 的國家圖書館，舉辦了都柏林核心集的第四次研討會 (簡稱DC-4)，與會者是65位包括電子圖書館研究者、網際網路專家、圖書館員等人士。根據『The 4th Dublin Core Metadata Workshop Report』一文的描述 [註9]，會議的主要成果是「坎培拉修飾詞」(Canberra Qualifier)，正如文法中的修飾詞功用一樣，都柏林核心集中的修飾詞，是對項目的內容和語意，做進一步的界定或說明，使其意義更明確，目前包括三種修飾詞：

(一) 語言 (language) 修飾詞：指出資料項所使用的語言。

例如：Lang =en，指出此資料項是使用英文著錄的。

[註：此修飾詞不是用來指示原始文件 (或資源) 所使用的語言。]

(二) 架構 (Scheme) 修飾詞：指出項目內容的詮釋方法。

例如：SCHEME = LCSH，指出這個主題是來自LCSH。

(三) 次項目 (Subelement) 修飾詞：進一步指示資料項內容涵蓋的範圍或出處。[註 10]

例如：Creator.PersonalName = “C.J. Wu”，更明確的界定此資料的屬性是 (著者) 姓名。

至於「坎培拉修飾詞」的實作方式目前有二種建議：[註11]

(一) 完全遵循HTML 2.0規格，其寫法如下：

```
<META NAME="DC.Subject"
```

```
CONTENT=" ( SCHEME=LCSH ) ( LANG=EN ) Computer Cataloging of  
Network Resources">
```

此法的優點是與目前的系統和瀏覽器完全相容，缺點是修飾詞和項目的內容，混合在同一字串內，須有特殊程式將其從字串中抽離出來。

(二) 遵循HTML 4.0規格，其寫法如下：

```
(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。同時由HTML 4.0中也新加入兩項屬性 – LANG 和 SCHEME 來看 [註 12]，可知修飾詞的使用，已是都柏林核心集發展中的一個必然趨勢，因此第二種寫法已經成為都柏林核心集的正式寫法。

在都柏林核心集的三種修飾詞—語言修飾詞（LANG）、架構修飾詞（SCHEME）、次項目修飾詞（SUBELEMENT）中，語言修飾詞是目前發展最明確的一種，在OCLC的都柏林核心集的（半官方）首頁中，建議遵循RFC 1766的規定，使用 ISO 639的二個字母的語言代碼，如LANG="EN"。[註 13] 另外一個常用的語言代碼是 NISO Z39.53 的三個字母代碼。[註 14]

相對於語言修飾詞，架構修飾詞和類別修飾詞在都柏林核心集的15個項目中，則隨著各項目功能的差異，以及使用者地域的不同而有很大的變化，因此下面以逐項討論的方式進行。由於這二種修飾詞仍在發展中，以下的討論主要是綜合『Dublin Core Qualifiers』[註15]、『Guide to Creating Core Descriptive Metadata』[註16]、『Provisional Report of the Dublin Core Subelement Working Group』[註17]、『Date Ranges and DC.Date Subelement Definitions』[註18]、『Relations Working Group』[註19] 五篇文章而來。

(一) 主題和關鍵詞（Subject）：有下列的架構修飾詞。

- (1) LCSH（美國國會圖書館主題標題表）-- Library of Congress Subject Heading。
- (2) UDC（國際十進分類號）-- Universal Decimal Classification。
- (3) DDC（杜威十進分類號）-- Dewey Decimal Classification。
- (4) NLM（美國國立醫學圖書館分類號）-- National Library of Medicine。
- (5) MeSH（醫學標題表）-- Medical Subject Headings。
- (6) Colon（冒號分類法）-- Colon Classification。
- (7) JEL（經濟期刊文獻分類法）-- Journal of Economic Literature Classification。
- (8) RCHME（英文索引典）-- English Heritage Thesaurus (ISBN 1-873592-20-5)。
- (9) AAT（藝術與建築索引典）-- Art & Architecture Thesaurus。

- (10) ULAN (藝術家索引) -- Union List of Artist's Names 。
- (11) TGN (地理名詞索引典) -- Thesaurus of Geographic Names 。
- (12) SHIC2 (社會歷史和產業分類法) -- Social History & Industrial Classification 。
- (13) TGM1 (國會圖書館地理名詞索引典) -- LC Thesaurus for Graphic Materials I: Subject Terms 。
- (14) MSC (數學分類法) -- Mathematical Science Classification 。
- (15) YKL (芬蘭公共圖書館分類法) -- Finnish Public Library Classification 。
- (16) SAB (瑞典圖書館分類法) -- Swedish Library Classification 。

(二) 題名 (Title): 有下列的次項目修飾詞。

- (1) 完整題名 (Long)。
- (2) 短題名 (Short) -- 題名的簡稱。
- (3) 其他題名 (Alternative) -- 題名的另外通稱。
- (4) 副題名 (Subtitle)。
- (5) 書背題名 (Spine) -- 題名取自書背。
- (6) 翻譯題名 (Translated) -- 原著翻譯書的題名。

[作者註: 在都柏林核心集次項目修飾詞工作小組的草案中, 則是較簡略的分為主要題名 (Main) 和其他題名 (Alternative) 兩種。]

(三) 著者 (Creator): 主要有下列的類別修飾詞, 寫法則依照一般慣例, 直接加在項目名稱的後面, 如

```
<META NAME="DC.Creator.Email"
CONTENT="lins1022@fujens.fju.edu.tw">
```

- (1) 姓名 (Name)。
- (2) 電子郵件位址 (Email)。
- (3) 郵件地址 (Postal)。
- (4) 電話號碼 (Phone) -- 建議使用 (+國家碼 區域碼 本地電話碼) 的格式, 例如輔仁大學的總機是 +886 02 29031111 。
- (5) 傳真號碼 (Fax) -- 格式請參照上面的項目 (4) -- 電話號碼。
- (6) 任職機構 (Affiliation) -- 著者任職機構的名稱。

- (7) 住家電子郵件位址 (HomeEmail)。
- (8) 住家地址 (HomePostal)。
- (9) 住家電話 (HomePhone) -- 格式請參照上面的項目 (4) -- 電話號碼。
- (10) 住家傳真 (HomeFax) -- 格式格式請參照上面的項目 (4) -- 電話號碼。
- (11) 網址首頁 (Homepage) -- WWW的首頁。

[作者註：在都柏林核心集次項目修飾詞工作小組的草案中，是先分為個人名稱 (PersonalName) 和團體名稱 (CorporateName) 兩種，然後再細分如上。]

(四) 簡述 (Description)：有下列的架構修飾詞。

- (1) URN – 外在說明文件的 URN 編號。
- (2) URL – 外在說明文件的 URL 位址。

(五) 出版者 (Publisher)：次項目修飾詞參照項目(三)著者。

(六) 其他參與者 (Contributors)：次項目修飾詞請參照上面的項目(三)著者。由於可能參與者的種類繁多，雖然都柏林核心集的第四次研討會中，並未設立角色 (Role) 修飾詞，但為能明白標示參與者所扮演的角色，建議加入以下的角色修飾詞。

- (1) 編者 (Editor)。
- (2) 插圖者 (Illustrator)。
- (3) 攝影者 (Photographer)。
- (4) 裝訂者 (Binder)。
- (5) 翻譯者 (Translator)。
- (6) 電腦資料創造者 (MachineReadableCreator)。[作者註：此處泛指將非數位文件轉成數位資料的人。]
- (7) 贊助者 (Sponsor)。
- (8) 編纂者 (Compiler)。
- (9) 著錄者 (Cataloger)。
- (10) 聯絡者 (Contact)。
- (11) 評論者 (Reviewer)。
- (12) 校對者 (Proofreader)。

(13) 行銷者 (Distributor)。

(七) 出版日期 (Date)：雖然在前面第二節的欄位介紹中，曾經提過建議使用ISO 8601的子集合，即如下格式- YYYY-MM-DD。但仍有以下的格式也經常被使用，祇是使用時務必要用架構修飾詞來指示使用的標準和格式。

(1) (IETF) RFC 822 – 例如 Sun, 21 Dec 1997 21:37:15 +0800 (星期, 日 月年 時: 分: 秒 國際格林威治時間差)。

(2) (ANSI) X3.30-1985 -- 例如 19971221 (YYYYMMDD)。

(3) (ISO) 31-1:1992 -- 例如 1997-12-21 (YYYY-MM-DD)。

出版日期有下列的次項目修飾詞：[註 20]

(1) 創造日期 (Created) -- 文件初次創造的日期。

(2) 發行日期 (Issued) -- 文件的正式發行日期。

(3) 接受日期 (Accepted) -- 如論文和條約的正式接受日期。

(4) 存取日期 (Available) -- 文件可存取的開始日期或期間。

(5) 典藏日期 (Acquired) -- 文件到館或典藏的開始日期。

(6) 取得日期 (DataGathered) -- 原始資料取得的開始日期，如氣象資料。

(7) 可使用日期 (Valid) -- 文件內容成立的期間。

例子：公車時刻表可有如下的日期資訊：

Date.Created =1997-12-01。

Date.Issued =1997-12-15。

Date.Available =1997-12-01/1998-12-31。

Date.Valid =1998-01-01/1998-12-31。

(八) 資源類型 (Type)：次項目修飾詞參照前面的欄位介紹項目八 -- 資源類型。

(九) 資料格式 (Format)：無架構修飾詞，內定使用 MIME格式。

(十) 資源識別代號 (Identifier)：主要有下列的架構修飾詞。

(1) 全球資源定位器 (URL)。

(2) 全球資源識別名稱 (URN)。

(3) 國際標準書號 (ISBN)。

(4) 國際標準叢刊號 (ISSN)。

(十一) 關連 (Relation)：架構修飾詞請參照上面的項目十 -- 資源識別代號，次項目修飾詞則如下所列。[註 21]

(1) 父階層 (IsParentOf) -- 現在文件是項目十一所指示文件的上一階層。

(2) 子階層 (IsChildOf) -- 現在文件是項目十一所指示文件的下一階層。

(3) 屬於 (IsMemberOf) -- 現在文件是項目十一所指示文件的一部份。

(4) 衍生 (IsDerivedFrom) -- 現在文件是項目十一所指示文件衍生而來的。

(5) 書目資料 (HasBibliographicInfoIn) -- 項目十一所指示文件含有現在文件的書目資料。

(6) 版本資訊 (IsRevisionHistoryFor) -- 項目十一所指示文件含有現在文件的版本資訊 (修改歷史)。

(7) 評論 (IsCriticalReviewOf) -- 項目十一所指示文件是現在文件的評論。

(8) 概要 (IsOverviewOf) -- 項目十一所指示文件是現在文件的概要。

(9) 評比等級 (IsContentRatingFor) -- 項目十一所指示文件含有現在文件的內容評比等級資訊。

(10) 補充資料 (IsDataFor) -- 項目十一所指示文件含有現在文件的補充資料，如數據資料和程式。

[作者註：在都柏林核心集次項目修飾詞工作小組的草案中，是先分為類型 (Type) 和指示器 (Indicator) 兩種，然後類型 (Type) 再細分如下：

(1) 屬於 (IsPartOf) -- 現在文件是項目十一所指示文件的一部份。

(2) 版本 (IsVersionOf) -- 現在文件是項目十一所指示文件的某個版本。

(3) 格式 (IsFormatOf) -- 現在文件是項目十一所指示文件的另一種格式或呈現形式 (即兩者的內容相同)。

(4) 引用 (References) -- 現在文件引用項目十一所指示文件。

(5) 衍生 (IsBasedOn) -- 現在文件是項目十一所指示文件衍生而來的。

(6) 要求(Requires)-- 現在文件須要利用到項目十一所指示文件。

以作者的看法，前面的關連 (Relation) 次項目修飾詞版本，似乎較工作小組草案的完整。在寫法上都柏林核心集關連修飾詞工作小組的草案中，規定使用 <Relation.Type> <Relation.Indicator > 的方式，例如 Relation = IsPartOf http://mes.lins.fju.edu.tw/。]

(十二) 來源 (Source)：架構修飾詞請參照上面的項目十 -- 資源識別代號。

(十三) 語言 (Language)：架構修飾詞有

(1) ISO 639。

(2) Z 39.53。

(十四) 涵蓋時空 (Coverage)：次項目修飾詞有二

(1) 空間 (PlaceName)。

(2) 時間 (PeriodName)。

架構修飾詞在配合空間座標時

(1) 經緯度 (LatLong) -- 例如 -37.24.43/+121.58.54。

當次項目修飾詞是時間，架構修飾詞參照項目七 -- 出版日期。

(十五) 版權規範 (Rights Management)：有下列的架構修飾詞。

(1) URN – 外在版權規範說明文件的 URN 編號。

(2) URL –外在版權規範說明文件的 URL 位址。

四、格式轉換對照表

以下是作者根據中國機讀編目格式第四版，製作的一份從都柏林核心集對映（轉換）到中國機讀編目格式（Chinese MARC）的摘要表格。

表1. 都柏林核心集對映到中國機讀編目格式的對照表。

都柏林核心集	修飾詞	中國機讀編目格式
Title (題名)		200 \$a 正題名
		200 \$d 並列題名，510\$a 並列題名
	Subtitle	200 \$e 副題名
	Long	553 \$a 完整題名 512 \$a 封面題名

	Spine	516 \$a 書背題名
	Alternative	517 \$a 其他題名
Creator (著者)	PersonalName	200 \$f 第一著者敘述
	PersonalName	700 \$a 主要著者
	PersonalName	702 \$a 其他著者
	CorporateName	710 \$a 團體主要著者
	CorporateName	712 \$a 團體其他著者
Subject (主題和關鍵詞)	LCSH 、 MeSH, etc.	610 \$a 非控制主題詞彙 606 主題標題
	UDC	675 國際十進分類號
	DDC	676 杜威十進分類號
	LCC	680 美國國會圖書館分類號
	CCL	681 中國圖書分類號 682 農業資料中心分類號
	NLM	686 美國國立醫學圖書館分類號 687 其他分類號
Description (簡述)		330 \$a 摘要註
Publisher (出版者)	Name	210 \$c 出版者，經銷者等名稱
	Postal	210 \$b 出版者，經銷者等地址
Contributors (其他參與者)	PersonalName	702 \$a 其他著者
	CorporateName	712 \$a 團體其他著者
		200 \$g 第二及依次之著者敘述
Date (出版日期)		210 \$d 出版，經銷等日期
Type (資源類型)		204\$a 資料類型標示
Format (資料格式)		300 \$a 一般註
Identifier (資源識別代號)		001 系統控制號
	ISBN	010 \$a 國際標準書號 (ISBN)

	ISSN	011 \$a 國際標準叢刊號 (ISSN)
		020 \$b 國家書目號
	URL	856 \$u URL
Source (來源)		300 \$a 一般註
Language (語言)		101 \$a 作品語文
Relation (關連)		300 \$a 一般註
Coverage (涵蓋時空)	PeriodName	122 \$a 作品涵蓋時間
	PlaceName	123 \$d 西經、\$d 西經、\$e 北緯、\$f 南緯
Rights (版權規範)		300 \$a 一般註

為了使讀者便於和中國機讀編目格式第四版相對照，在欄位名稱和符號格式的寫作上，將力求與中國機讀編目格式一致。以下是針對上述表格的詳細說明和例子：

- (一) 題名 (Title)：第一個題名放入 200\$a 當作正題名，其餘題名資料若無法詳加分辨，則當作其他題名放入 517\$a。若能透過修飾詞或人工分辨，則分別處理如下：並列題名同時放入200 \$d和510\$a，副題名 (scheme = subtitle) 放入 200 \$e，完整題名 (scheme = long) 放入 553 \$e，書背題名 (scheme = spine) 放入513 \$a，封面題名放入 512 \$a。

例子一：200 1 \$a元資料實驗系統。

- (二) 著者 (Author or Creator)：目前都柏林核心集暫定使用次項目修飾詞—PersonalName和CorporateName來區分個人著者和團體著者，中國機讀編目格式將個人著者放入 700/702，而團體著者是放入 710/712。另外中國機讀編目格式將姓和名放在不同分欄，雖然對人類而言處理這些差異並不困難，但用電腦來自動辨識時將較為困難和費工夫。另外一個困難是都柏林核心集也不刻意區分主要著者，因此當著者超過一人時，若無法透過修飾詞或人工分辨出主要著者時，則將第一個著者視為主要著者。總言之主要著者放入 200 \$f 和 700 \$a，其餘的其他著者置入 200 \$g 和 702\$a。請注意如同並列題名一般，相關著者資料須同時出現在 200 的著錄段和 700 的著者及輔助檢索段。

例子一：200 1 \$a元資料實驗系統\$f吳政叡。

例子二：700 1 \$a吳\$b政叡。

例子三：710 02 \$a輔仁大學圖書資訊系。

(三) 主題和關鍵詞 (Subject and Keywords)：都柏林核心集已規劃有架構修飾詞來區分詞彙的來源依據，因此未使用架構修飾詞者，視為非控制主題詞彙放入610\$a。若有使用架構修飾詞，則放入 606\$a，並以分欄 2 來表示所使用的標題系統。

例子一：610 01 \$a都柏林核心集。

例子二：606 1 \$2lc\$aWord processing 。

若所屬資料為分類碼，則國際十進分類號 (UDC) 放入 675\$a、杜威十進分類號 (DDC) 放入 676\$a、美國國會圖書館分類號 (LCC) 放入 680\$a、中國圖書分類號 (CCL) 放入 681\$a、農業資料中心分類號放入 682\$a、美國國立醫學圖書館分類號 (NLM) 放入 686\$a，其餘置入 687\$a，再利用分欄 d 來表示所使用的分類號系統。

例子三：676 \$a025.313。

(四) 簡述 (Description)：放入 330\$a。

例子一：330 \$a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。

(五) 出版者 (Publisher)：出版者名稱放入 210 \$c，地址則置於 210 \$b。

例子一：210 \$c松崗電腦圖書資料股份有限公司\$b台北市敦化南路一段 339 號 5 樓\$d民85。

(六) 其他參與者 (Contributors)：將所有的其他參與者當成其他著者，作法如同前面的著者欄位，所有的其他參與者置入 200 \$g，同時將個人參與者放入 702\$a，而團體參與者是放入 712\$a，例子請參見前面的著者欄位。

(七) 出版日期 (Date)：祇取西年紀年放入 210 \$d，例子參見前面的出版者欄位。

(八) 資源類型 (Resource Type)：放入 204\$a。

例子一：204 0 \$a地圖。

(九) 資料格式 (Format)：此處主要是來標示電腦的檔案格式，建議使用MIME 所規範的術語，例如 text/html，目前在中國機讀編目格式中的幾處電腦相關欄位，如欄位 135、736、786等均不太合適放入，因此暫時放入 300\$a。

例子一：300 \$atext/html (MIME)。

(十) 資源識別代號 (Resource Identifier)：中國機讀編目格式的欄位 001—系統控制號是必備欄且須惟一，可由系統自動產生，或取自其他來源，例如國際標準書號 (ISBN)、URL 等。

例子一：001 <http://mes.lins.fju.edu.tw/mes/default.htm>。

若有架構修飾詞來區分資源識別代號的系統，則國際標準書號放入 010、國際標準叢刊號 (ISSN) 放入 011、國家書目號 (National Bibliography Number，簡稱 NBN) 放入 020、URL 放入 856\$u。

例子二：010 0 \$a957-22-2155-8。

(十一) 來源 (Source)：目前在中國機讀編目格式中並無適當的對映欄位，因此暫時放入 300\$a。

(十二) 語言 (Language)：放入 101\$a，因為現行的都柏林核心集建議使用 ISO 639的二個字母語言代碼，但是一般的機讀編目格式如 UNIMARC 和中國機讀編目格式是使用 Z39.53，因此須要建立一個兩者的對照表來進行轉換工作。

例子一：101 0 \$achi。

(十三) 關連 (Relation)：中國機讀編目格式無特定欄位相對映，放入 300\$a (一般註)。

例子一：300 \$a (type = IsChildOf) <http://140.136.85.194/default.htm>。

(十四) 涵蓋時空 (Coverage)：若架構修飾詞是 Temporal，放入 122\$a，若是 Spatial則放入欄位 123。

例子一：122 2 \$ad1972\$ad1998。

例子二：123 1 \$aa\$b253440\$de0550000\$ee0950000\$fn0310000
\$gn0110000。

(十五) 版權規範 (Rights Management)：中國機讀編目格式無特定欄位相對映，放入 300\$a (一般註)。

例子一：300 \$a使用無限制。

雖然中國機讀編目格式已有一些電腦相關欄位如欄位 135、736、786等，

但由以上的對照和分析，顯示出中國機讀編目格式在一些對網路文件相當重要的新欄位，如（電腦）資料格式、關連、版權規範、來源等，目前均無法適當的處理，祇能暫時放入欄位 300（一般註）中，這是有待加強的部份。

五、結語

元資料（Metadata）最常見的英文定義是 "data about data"，可直譯為描述資料的資料，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。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（Dublin Core）是眾多元資料種類中的一種，國際圖書館電腦中心（OCLC）主導創設都柏林核心集的目的，是希望能一方面解決MARC在應付網路文件上的困境，另一方面能有一套簡單的資源描述格式，讓眾多非圖書館的專業人士來使用，以最少成本來解決網路文件快速增加的問題，因此都柏林核心集是定位在簡單的資源描述格式，但提供一個基本資料庫，作為各種專業進一步加工處理的基礎。除了本文所介紹的都柏林核心集到中國機讀編目格式的轉換對照表外，其餘與都柏林核心集相關的格式轉換尚有以下幾種，有興趣的讀者請自行參閱相關的網頁。

（一）從都柏林核心集到UNIMARC（http://www.ukoln.ac.uk/metadata/interoperability/dc_unimarc.html）。

（二）從都柏林核心集到USMARC（Dublin Core/MARC/GILS Crosswalk）。

（三）從都柏林核心集到EAD（<http://www.oclc.org:5046/~emiller/DC/crosswalk.html>）。

（四）從都柏林核心集到IAFA/ROADS Templates（http://www.ukoln.ac.uk/metadata/interoperability/dc_iafa.html）。

（五）從都柏林核心集到Z39.50 tag set G（<http://www.roads.lut.ac.uk/lists/meta2/0733.html>）。

（六）從SOIF到都柏林核心集（http://www.ukoln.ac.uk/metadata/interoperability/soif_dc.html）。

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22]

- (一) 加快標準化的腳步—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。與NIST方面的接觸也在進行中，例如都柏林核心集和Z39.50間的互通，就已在積極進行中。
- (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，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，寫法請參見文中的介紹。至於RDF格式則尚在發展中。
- (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。

這次研討會的另一個特點，是在第二天針對都柏林核心集的實作系統進行展示和討論，這是以前四次研討會所沒有的，也說明都柏林核心集已漸趨成熟和受到肯定。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23]

- (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（請參見下面第四章中關於DSTC的介紹）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。

註釋

註 1：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 期（民 85 年 12 月），頁39-41。

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：S. Weibel and E. Miller, "Dublin Core Metadata Element Set: Reference Description," 2 Oct. 1997, <http://purl.oclc.org/metadata/dublin_core_elements>.

註 4：M. Wolf and C. Wicksteed, "Date and Time Formats," 15 Sept. 1997, <<http://www.w3.org/TR/NOTE-datetime>>。

註 5：R. Tennant, "Dublin Core Resource Types," 23 Sept. 1997, <<http://sunsite.berkeley.edu/Metadata/minimalist.html>>.

註 6：H. T. Alvestrand, "Tags for the Identification of language," March 1995, <<http://ds.internic.net/rfc/rfc1766.txt>>, p. 2.

註 7："Guide to Creating Core Descriptive Metadata," 13 April 1996, <<http://www.ckm.ucsf.edu/people/jak/meta/mguide3.html>>, p. 7.

註 8：S. Weibel and E. Miller, "Image Description on the Internet: A Summary of the CNI/OCLC Image Metadata Workshop," D-Lib Magazine (Jan. 1997), <<http://www.dlib.org/dlib/january97/oclc/01weibel.html>>, p. 5.

註 9：S. Weibel, R. Iannella, and W. Cathro, "The 4th Dublin Core Metadata Workshop Report," D-Lib Magazine (June. 1997), <<http://www.dlib.org/dlib/june97/metadata/06weibel.html>>.

註10：P. Miller, "Provisional Report of the Dublin Core Subelement Working Group," 11 Feb. 1998), <http://purl.oclc.org/metadata/dublin_core/wsubelementdraft.html>, p. 1.

註11：吳政叡，「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18。

- 註12 : B. Marsh, "Syntactic Considerations for the Dublin Core," 2 Nov. 1997, <http://purl.oclc.org/metadata/dublin_core/syntax.html>, p. 7.
- 註13 : 同註7 , 頁7 。
- 註14 : "NISO Z39.53 Language Codes," 1994, <<http://www.sil.org/sgml/nisoLang3-1994.html>>, p. 7.
- 註15 : J. Knight and M. Hamilton, "Dublin Core Qualifiers," 1 Feb. 1997, <<http://www.roads.lut.ac.uk/Metadata/DC-SubElements.html>>.
- 註16 : 同註7 。
- 註17 : 同註10 。
- 註18 : J. Kunze and E. Jul, "Date Ranges and DC.Date Subelement Definitions," (14 Feb. 1998), <http://purl.oclc.org/metadata/dublin_core/wdatedraft.html>.
- 註19 : D.Bearman, "Relations Working Group," (14 Feb. 1998), <http://purl.oclc.org/metadata/dublin_core/wrelationdraft.html>.
- 註20 : 同註18 , 頁1-3 。
- 註21 : 同註15 , 頁10-11 。
- 註22 : 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23 : 同註22 , 頁3-4 。