

淺談開放式元資料獲取協定的應用與內涵

A Novel Way of Exposing and Harvesting Metadata Using the OAI-PMH

余顯強

Shien-Chiang Yu

輔仁大學圖書資訊學系兼任講師

Lecturer

Department of Library and Information Science

Fu-Jen Catholic University

【摘要 Abstract】

開放式檔案分享機制是邁向開放式數位圖書館應該具備的服務之一，而 OAI-PMH 是發佈和獲取元資料的開放式標準，藉由此項標準可使各個資料提供者與服務提供者之間系統的溝通更為容易，使得數位典藏的資料能夠保有 metadata 的原始結構，透過標準且簡單的程序達到分享、使用與加值，有助於使用者更方便地檢索與獲取網路資源，滿足文獻資訊檢索的需求。本文旨在探討開放式元資料獲取協定的規範、目的、組成元件、內涵及其相關的應用。

Providing open archive sharing is one of the essential functions of Open Digital Libraries. OAI-PMH is an Open Archive Initiative Protocol of Metadata exposing and Metadata harvesting. OAI-PMH enhances interoperability among many existing and independent Digital Libraries systems. This protocol combines and processes information from each individual archive. Also, it increases functionality in the aspects of discovery, presentation and analysis of data originating from compliant archives. This paper attempts to discuss the concept, standard, propose, and composition of OAI-PMH.

關鍵詞 Keyword

開放式檔案協會 開放式元資料獲取協定 元資料 數位圖書館 聯盟檢索

Open Archives Initiative (OAI) ; Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) ;
Metadata ; Digital libraries ; Federated searching



壹、前言

能夠透過地獲得所需的資源，而不需要逐一到各個數位圖書館內搜尋，是多數檢索者引領企盼許久的事，然而其中最主要的阻礙之一便是因為許多數位圖書館之間使用不同且特有的技術，因此，缺乏相互溝通的機制已經是現今數位圖書館之間，所面臨重大問題之一。(註 1)數位圖書館或學術機構的資料庫與系統彼此互不隸屬，相關資料或不同領域的資料不僅分散儲存而且難以整合，使得資料的分享與流通有所限制。尤其在學術單位，不同研究室或系所，經常會有合作的研究計劃或類似的研​​究需求。因此如何提供各研究室、系所、研究社群之間，分享或找出研究方向的相關文獻，以供研究進行或學術著作的引用參考是急需解決的問題。於是，1999年7月由 Paul Ginsparg, Rick Lue, Herbert Van de Sompel 等人發出了一份針對學術電子文件(e-print)分享計劃的 Universal Preprint Service 會議邀請(註 2)，並於 10 月在美國新墨西哥州 Santa Fe 所舉辦的會議中推動了 OAI (Open Archives Initiative)組織的成立。(註 3)

OAI 組織參考了由 C. Mic Bowman、Peter B. Danzig、Darren Hardy、Udi Manber、Michael Schwartz 等人所建置的 Harvest 系統架構(註 4)，並經過了先期版本測試一段時間之後，於 2001 年 1 月 21 日正式公佈了 OAI 元資料獲取協定(Open Archives Initiative Metadata Harvesting Protocol, OAIMH) (註 5)，同年 7 月 2 日修訂發表 1.1 版(註 6)，最新版本為 2002 年 6 月 4 日正式發表的 Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) 2.0 版。(註 7) OAI-PMH 制定的目的是提出分散式檔案之間相互溝通的技術，提供一簡單、低門檻的方式來建立一個能夠在各系統間溝通，並有效地分享及使用元資料的方法。而其主要的目標除了提供元資料的交換與使用，還包括下列

三點：(註 8)

- (一)簡化文件內容有效的傳播。
- (二)提升電子化文件的存取。
- (三)擴展存取數位化資料種類的範圍。

OAI 已經將 OAI-PMH 提交 W3C 計劃申請為國際標準，以期能成為全球元資料分享的開放性標準。藉由此協定具備應用程式獨立，且可互相運作的優點，以提供和提升 Web 上多種從事文件內容出版發行的社群應用與分享的服務功能。此一協定並不是要取代目前的 Z39.50 等開放式資料存取標準，而是提供一個容易實作(Easy-to-implement)和容易部署(Easy-to-deploy)的替代性解決方案。它的出現只是要讓 Internet 網路上這麼多的資訊服務提供者，以一組標準的陳述語法，可以讓不同的服務提供者及資料提供者互相溝通沒有阻礙。因此現今許多數位圖書館/博物館的計劃中，均將此一協定納入研究與實作範圍之內。

貳、技術規範與應用目的

OAI-PMH 的內涵主要包含元資料的發佈(Expose)與獲取(Harvest)兩個類型，因此在協定上分別定義兩個部分：

- (一)定義一個資料提供者能夠透過 HTTP 為基礎的協定，發佈其 metadata 的機制。
- (二)定義一個能夠從儲存器獲取含有 metadata 資料錄的機制。

因此，加入 OAI 組織的單位，依據任務的不同區分為資料提供者(Data provider)與服務提供者(Service provider)兩個角色。加入者須依據本身服務提供的種類，向 OAI 登記註冊成為資料提供者或服務提供者。OAI 在接受登記註冊後，會執行相對的驗證以確保登記之系統能夠完全符合 OAI 協定的規範。除此之外，登記註冊還有下列目的：(註 9)

- (一)能夠提供 OAI 確認過之儲存器的明細，使



服務提供者能夠很容易的知道有哪些能夠獲取資料來源的儲存器。

(二)確保資料提供者能夠提供符合 OAI 協定規範的機制。

(三)提供 OAI 能夠監督協定的使用，以及規劃未來的活動與策略。

OAI-PMH 是以 XML 相關的技術，發佈與獲取元資料的開放式標準。在技術的規範上盡量以實作簡單為原則，而由於元資料在應用領域的差異與語意內涵的深度，在資料的發佈與獲取上必須考慮通用的問題，以及元資料本身的限制。因此以下主要是說明 OAI-PMH 在應用上常令人誤解的一些迷思：(註 10)

(一)OAI-PMH 發佈和獲取的物件，並不包含如文件、影像、聲音等全文資料(Full-content)。OAI-PMH 只是一個用來交換元資料的協定，其餘文件格式、內容均需透過其他程式應用技術輔助，並不在此協定規範內。不過 OAI 承諾此協定將會提供全文資料的交換能力，至少在最新發表的 2.0 版本，仍未有此方面的協定規範。

(二)OAI-PMH 不是提供兩個檔案系統之間資料交換的協定，而是明確的將兩者之間切割成資料提供者(發佈元資料的實體)和服務提供者(獲取元資料並提供前端使用者增值服務的實體)。當前後端明確地分為資料提供者與服務提供者這兩個實體，並不是每一個加入 OAI-PMH 服務的單位必須只能擇一為之，例如 Cite Base(註 11)便是一個既為資料

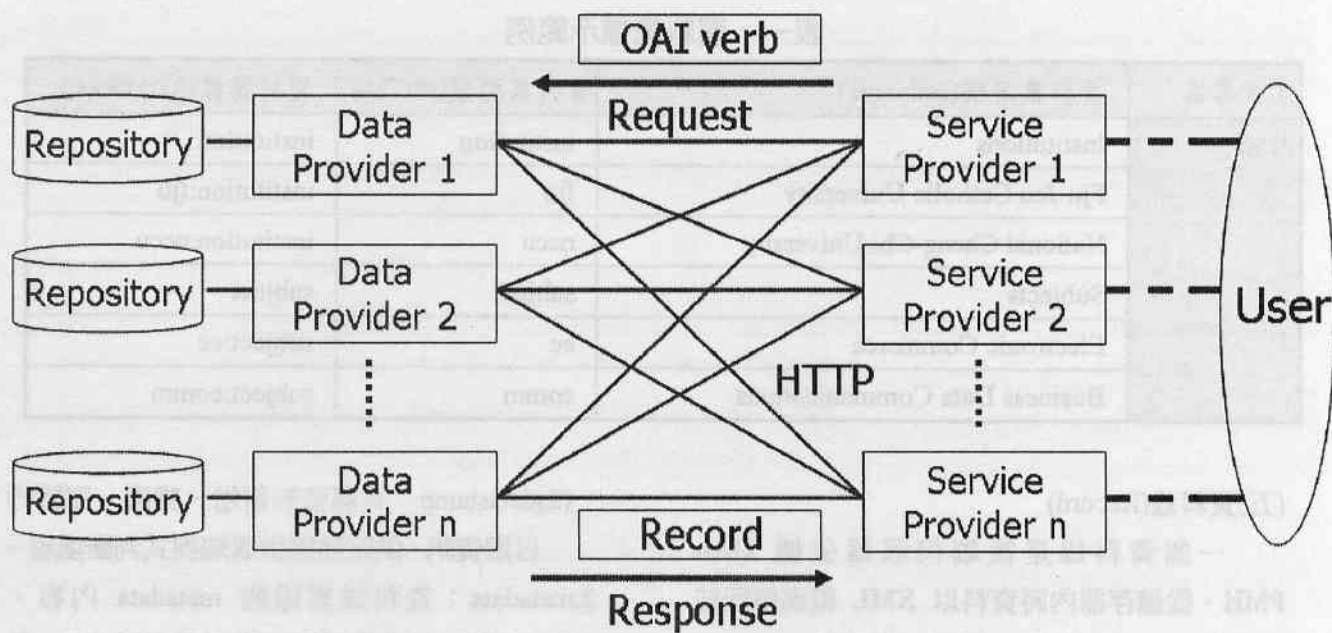
提供者，也是服務提供者。

(一)OAI-PMH 在傳輸的元資料格式並不只限於都柏林核心集元資料(Dublin Core, DC)，雖然 OAI 組織強調所有的資料提供者至少能夠提供 Dublin Core 元資料格式作為資料交換的基本格式。但為了考量一些特定領域的應用需求與其元資料結構，OAI 組織並不限制只能使用 Dublin Core 格式，而且還鼓勵發展特殊的格式以便能提供特殊領域的使用需求。不僅如此，OAI 也在多次會議中宣告支援所有的 metadata。(註 12)

參、組成元件

OAI-PMH 最主要的概念即是基於電子化文件「獲取」的目的，只要是提出 OAI-PMH 需求的前端應用程式便是一個獲取者(Harvester)，一個獲取者可以是一個服務提供者，也可以是一個前端的使用者。通常是由服務提供者向資料提供者提出需求，以便將取得的電子文件加值(Value-added)處理。因此 OAI-PMH 本身並非是由資料提供者驅動服務的互動行為，也不是由服務提供者(或獲取者)直接向資料提供者「擷取」所需的資料，而是如圖一所示，服務提供者透過 OAI-PMH 定義的指令向資料提供者提出資料獲取的請求，資料提供者再依據本身系統所能提供的資料範圍、種類、格式做適當的回應。因此提出獲取需求的主導權在「獲取者」這一方，通常是服務提供者，而資料提供者則有權依據本身的資料情況決定回應結果。





圖一：OAI-PMH 運作架構圖

OAI-PMH 整體的運作核心主要是在 HTTP 協定上傳輸使用 XML 文件的協定，而前後端整個運作環境，包含下列五個主要的組成元件：(註 13)

(一)資料提供者

提供其文件內容，並以 OAI-PMH 作為發佈 metadata 的協定。主要工作為維護一個或一個以上支援 OAI-PMH 來將其內容以 metadata 發佈的儲存器(Web 伺服器)。

(二)服務提供者

透過 OAI-PMH 協定向資料提供者取得資料，並將獲得的 metadata 在其上建立增值服務。

(三)資料儲存器(Repository)

透過 HTTP，接受 OAI-PMH 所提出存取資料需求的伺服器。

(四)資料集(Set)

非必備功能，主要是為了方便取得部分範圍所需資料的目的。儲存器內可將不同類別的資料區分成不同的群組，並以階層式架構表示，以節點(Node)作為各分類的區分，因此每一個節點即稱之為資料集。如表一所示，依據資料來源學校分類，分為輔仁大學與政治大學兩個資料集，各資料集之下可再細分資料集，因此服務提供者向後端資料提供者獲取資料時，便可指明特定資料集範圍的資料。



表一：資料集標示範例

元素名稱	資料集名稱(setName)	資料集標籤(setTag)	資料集標示(setSpec)
內容	Institutions	institution	institution
	Fju-Jen Catholic University	fju	institution:fju
	National Cheng-Chi University	nccu	institution:nccu
	Subjects	subject	subject
	Electronic Commerce	ec	subject:ec
	Business Data Communications	comm	subject:comm

(五)資料錄(Record)

一個資料錄是後端伺服器依據 OAI-PMH，從儲存器內將資料以 XML 編碼傳回前端的 metadata。如表二所示，資料錄 metadata 結構包含下列三個部分：

1.header：標記有關此筆資料錄的資訊，主要包含兩部分：

(1)unique identifier：儲存器內資料的唯一性標示。

(2)timestamp：有關資料新增、維護、刪除的日期資訊，供服務提供者端程式判斷處理。

2.metadata：資料錄實際的 metadata 內容。OAI-PMH 規定資料提供者必須支援 DC，是否提供其他元資料格式由資料提供者端的系統決定。

3.about：非必備的部分。提供有關資料的相關說明，如版權宣告、使用權限等，這部分的結構在 OAI-PMH 中不做規定。

表二：資料錄範例

```

<header>
  <identifier>oai:arXiv:cs/0112017</identifier>
  <timestamp>2002-02-28</timestamp>
  <setSpec>cs</setSpec>
  <setSpec>math</setSpec>
</header>

<metadata>
  <oai_dc:dc
    xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/

```

(續下表)



(接上表)

```

http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
  <dc:title>Using Structural Metadata to Localize Experience of Digital
    Content</dc:title>
  <dc:creator>Dushay, Naomi</dc:creator>
  <dc:subject>Digital Libraries</dc:subject>
  <dc:description>With the increasing technical sophistication of both
    information consumers and providers, there is increasing demand for
    more meaningful experiences of digital information. We present a
    framework that separates digital object experience, or rendering,
    from digital object storage and manipulation, so the
    rendering can be tailored to particular communities of users.
  </dc:description>
  <dc:description>Comment: 23 pages including 2 appendices,
    8 figures</dc:description>
  <dc:date>2001-12-14</dc:date>
  <dc:type>e-print</dc:type>
  <dc:identifier>http://arXiv.org/abs/cs/0112017</dc:identifier>
</oai_dc:dc>
</metadata>
<about>
  <provenance
    xmlns="http://www.openarchives.org/OAI/2.0/provenance/"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance/
    http://www.openarchives.org/OAI/2.0/provenance/oai_provenance.xsd">
    <originDescription>
      <baseURL>http://the.oa.org</baseURL>
      <identifier>oai:r2:klik001</identifier>
      <datestamp>2002-01-01</datestamp>
      <metadataPrefix>oai_dc</metadataPrefix>
      <harvestDate>2002-02-02T14:10:02Z</harvestDate>
    </originDescription>
  </provenance>
</about>

```



肆、協定內涵

OAI-PMH 主要是透過指定的指令(HTML 傳遞變數名稱)，提供前端向後端伺服器提取所需資訊的協定。由於 OAI-PMH 是基於 HTTP 之上的應用協定，因此其指令集即是透過 HTTP 所使用中的 GET 或 POST 方法，藉由前端與後端傳輸之變數名稱與其內容，觸發後端對應之伺服器應用程式，依據變數內容處理後，遵照 OAI-PMH 定義之 XML Schema 所規範的格式將結果傳回前端。

依據 OAI-PMH 及 HTTP 規範，前端向後端提出 GET 或 POST 請求時，至少需傳遞一個的參數，形式為：key=value，使用多個參數時必須以「&」隔開。GET 方法傳送的請求必須使用「？」連接在 URL 之後，例如：

`http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai:arXiv:quant-ph/9901001&metadataPrefix=oai_dc`

此一命令使用 OAI-PMH 的 GetRecord 指令，向資料提供者要求取得系統標示為 oai:arXiv:hep-th/9901001，且記錄的元資料格式為 oai_dc(OAI 定義之 Dublin-Core Schema)的資料。不過這是一個未經過編碼的命令，因為要求的系統標示包含「:」、「/」字元，而「:」字元在命令中常用於附加參數(如密碼、埠號等)；「/」字元在命令中則表示「目錄」的區隔符號(Delimiter)，因此實際送出上述 HTTP 命令時，必須轉換成：

`http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai%3AarXiv%3Aquant-ph%2F9901001&metadataPrefix=oai_dc`

後端資料提供者的系統對 OAI-PMH 請求的回應也必須使用 XML 格式回應，並加上適當的標示。每一個回應的 XML 資料的都應有如下標記：

- (一)字碼種類標示必須為 UTF-8(Unicode)編碼。
- (二)回應的資料必須置於請求中指令同名的元

素之內(例如使用 GetRecord 請求資料，則後端回應的資料會放置於<GetRecord>元素之內)。

- (三)對於所有的回應，根元素為 OAI-PMH 宣告 xmlns、xmlns:xsi、xsi:schemaLocation 名稱空間
- (四)對於所有的回應，根元素的第二個元素是 responseDate，依據資料提供者端的系統時間，標示回應發送的時間。
- (五)對於所有的回應，根元素的第三個子元素是 request，紀錄原先前端提出請求命令，內容必須依據 XML 實體參數編碼(例如「&」要替代成「&」，以便符合 XML 的語法規範)。

依據 OAI-PMH 2.0 版本，所提供之指令(變數名稱)共計下列六種：

- (一)GetRecord：從一個儲存體內的項目提取指定的資料錄/metadata。此一指令提供兩個必備的參數(Argument)。

1.identifier：標明提取資料的 Id，此 Id 代表於某一儲存器內提取一個項目之 metadata 的唯一標示。格式為「schema:Repository ID: local ID」。

2.metadataPrefix：指定提取之資料需呈現之格式類型。前端系統可藉由 ListMetadataFormats 指令向儲存器詢問能提供之格式類型種類。

- (二)Identify：用於取得儲存器相關之資訊，包括行政管理、系統標示、特殊社群等資訊。而後端伺服器回應之資訊，依據協定規範必須包含下列元素：

1.repositoryName：儲存器名稱。

2.baseURL：儲存器的 BASE-URL。

3.protocolVersion：儲存器支援的 OAI 協定版本。



4.adminEmail：儲存器系統管理者的 E-Mail。

(三)ListIdentifiers：向儲存器取得能夠獲取資料錄的 Id 明細。此一指令提供四個參數。

1.until：(非必備)指定資料日期標示 (datastamp)的迄止日期。

2.from：(非必備)指定資料日期標示的起始日期。

3.set：(非必備)指定資料集種類。

4.resumptionToken：用於流量控制使用之憑證標記，提供爾後前端使用此一憑證標記向後端 Service provider 提取相同之回應。

(四)ListMetadataFormats：取得儲存器能夠支援 metadata 的格式種類。此一指令提供一個必備參數 identifier，標明提取資料的 Id，此 Id 代表於某一儲存器內提取一個項目之 metadata 的唯一標示。格式為「schema:Repository ID: local ID」。

(五)ListRecords：向儲存器取得指定範圍的所有資料錄。此一指令提供五個參數。

1.until：(非必備)指定資料日期標示的迄止日期。

2.from：(非必備)指定資料日期標示的起始日期。

3.set：(非必備)指定資料集種類。

4.resumptionToken：用於流量控制使用之憑證標記，提供爾後前端使用此一憑證標記向後端 Service provider 提取相同之回應。

5.metadataPrefix：指定提取之資料需呈現之格式類型。前端系統可藉由 ListMetadata Formats 指令向儲存器詢問能提供之格式類型種類。

(六)ListSets：取得儲存器的資料集結構。此一指令提供一個必備參數 resumptionToken，用於流量控制使用之憑證標記，提供爾後前端使用此一憑證標記向後端 Service provider 提取相同之回應。

伍、相關應用與研究

OAI 最基本的主要目的是希望建立一套包含搜尋引擎與連結系統之間的使用者服務機制，提供散聚各處且不同性質之電子化文件的分享使用。因此在最初 OAI 會議召開之初，先成立了 Universal Preprint Service(UPS) Prototype 計劃。(註 14)此計畫主要是建立跨檔案之間使用者服務的可行性研究。研究中的資料來源結合了如表三所示的檔案機構。透過此六個檔案機構，提供將近 200,000 筆記錄。



表三：研究開放式檔案之機構

檔案中機構	名稱	服務網址
arXiv	Los Alamos E-Print Archives	< http://arXiv.org >
CogPrints	Cognitive Sciences Eprint Archive	< http://cogprints.soton.ac.uk/search >
NACA	Digital Library for the National Advisory Committee for Aeronautics	< http://naca.larc.nasa.gov >
NCSTRL	Networked Computer Science Technical Reference Library	< http://ww.ncstrl.org >
NDLTD	Networked Digital Library of Theses and Dissertations	< http://www.theses.org >
RePEc	Research Papers in Economics	< http://netec.ncc.ac.uk/WoPEc.html >

雖然此實驗性計畫使用的協定是改良於 Dienst 的 NCSTRL+, 不過最後將成果呈現於 Santa Fe 的 OAI 會議上, 提供了下列在 OAI-PMH 制定的項目與流程的主要參考與建議依據:

1. 資料提供者和服務提供者間的界定與區別
2. 建立多樣獲取(Rich harvesting)的方法
3. 確定資料提供者和服務提供者整合的框架
4. 電子文件唯一性標示名稱空間的重要
5. 規範處理的元資料

之後, 包含美國、歐洲等地便開始有利用 OAI-PMH 技術框架格式的實驗計畫, 茲介紹如下:

一、Arc(註 15)

Arc 是第一個應用 OAI-PMH 協定提供結盟單位之間的檢索服務。其最初是 Old Dominion 大學應用 Universal Preprint Service (UPS) 協定及多種不同的技術所建置的服務計劃。為了提昇聯盟檢索(Federated searching)的應用效率, Arc 將「OAI 層」架設於原系統上, 因此一個服務提供者能夠同時從其他資料提供者和服務提供者獲取資訊。而此研究也對 OAI 提出了下列改善建議:

(一)在分散式環境中獲取資料的過程, 不同檔案

系統必須有不同的格式與命名轉換方式, 因此協定必須規範資料正規化。

(二)無限制的收集資料容易導致資料提供者系統的負荷超載、資料被非法轉製或濫用, 因此資料提供者的安全保護必須能夠限定擷取的來源, 不過如此卻使此系統難以實作。因此協定上合法的資料提供者與服務提供者需要註冊的機制。

二、National Science Digital Library(NSDL)(註 16)

NSDL 計畫經費是由美國 National Science Foundation 所支援, 為一個多方面參與的計畫, 其目的是創造一個線上網路的學習環境與資源, 主題包括科學、數學、工程和基礎教育等, 未來將與 San-Diego Super Computer Center 合作, 經由 OAI-PMH 獲取元資料, 再加以重製與最佳化後儲存於元資料儲存器, 使這些資料能夠透過 SDLIP(註 17)協定檢索。

三、Cyclades(註 18)

此計畫是由 European Commission 與義大利、德國、英國等國家共同推行的計畫, 主要目標是希望經由 OAI 發展更好的網路存取與傳



遞服務，以協助學者和會員在面對大量、繁雜的電子文件時，可以檢索到需要的資料。

四、Digital Library Federation Testbed(註 19)

此為 2000 年 10 月由 W. Mellon 基金會在劍橋所召開的會議之後所延伸的一個後續的研究計劃，主要是為了探索元資料格式下一階段在獲取和標示上的可行方案，以便解決在科技、組織和資源上使用的問題。而會議的結論便是決定採用 OAI 技術來解決大量館藏的元資料發佈，以及實驗使用這些元資料的服務方法。

陸、OAI-PMH 與 Z39.50 的比較

本質上，OAI-PMH 與 Z39.50 均能達成聯盟檢索的需求，也就是使用者能夠透過單一的介面從多個不同數位圖書館取得所需的資訊。透過一次查詢的動作，便可從許多不同來源獲得結果，讓使用者不需逐一查詢各個數位圖書館。Z39.50 已經是圖書館系統使用很久一段時間的開放式檢索協定，主要是由「原始系統」(或稱 Z39.50 Client) 依靠線上即時連結一個以上的「目標系統」(或稱 Z39.50 Server)，使用複雜的通訊協定集執行所需的查詢行為，再逐一取得由目標系統獲得的查詢結果、排序、去除重覆，最後將處理結果呈現出來。

由於下列的三項原因，應用 Z39.50 協定的分散式檢索系統之間必須考量回應時間、結果大小、網路頻寬的條件，使 Z39.50 被批評非常難建立一個高品質的聯盟檢索服務：(註 20)

- 一、不同的自動化系統對 Z39.50 查詢請求的語意常會有不同的解譯，而導致不正確的執行結果。
- 二、連接多個目標系統時，由於取得的資訊必須線上即時在原始系統內組合整理，導致嚴重影響系統的效率。
- 三、使用者在獲得結果之前，必須等待資料在系統之間傳輸與處理的時間。

因此，就 OAI-PMH 與 Z39.50 兩者作一比較，OAI-PMH 以 XML 作為資料處理的依據，著重在容易建置，並期望解決所有 metadata 的分享問題。服務提供者可以使用 OAI-PMH 事先取得所有後端資料提供者所有的資料，增值或建立索引之後，間接提供前端使用者所需的資訊，達成聯盟檢索的需求。在這種模式之下，OAI-PMH 服務提供者是建立在資料中介者(Broker)的角色。而 Z39.50 以 MARC 和 Dublin Core 為資料處理的依據，著重在點對點(Peer to peer)的協定服務模式，透過線上即時查詢目標系統的資訊，然後處理所有遠端回應的資料集之後，再呈現給前端使用者。強調的是即時性的聯盟檢索服務，中間不存在任何中介者的角色，前端也不強調資料增值的處理與分析。

因此，OAI-PMH 與 Z39.50 兩者協定不僅應用目的不相同，使用的方法也不相同，雖然兩者都可達成聯盟檢索服務的需求，但彼此之間卻無法取代彼此。不過一個很好的方式是將 Z39.50 視為一個橋接器(Bridge)或閘道(Gateway)，除了處理 Z39.50 的命令之外，也能處理 OAI-PMH 的命令，而其中需要增加的處理能力是能支援 metadata 綱要，即可結合彼此的優點。

柒、結論

開放式檔案分享機制是邁向開放式數位圖書館應該具備的服務之一，而所有數位圖書館之間也應該能共同分享彼此的資源，提供使用者透通的資訊取得管道。而 OAI-PMH 是發佈和獲取元資料的開放式標準，藉由此項標準可使各個資料提供者與服務提供者之間系統的溝通更為容易，使得數位典藏的資料能夠保有 metadata 的原始結構或 Dublin Core 格式，並透過標準且簡單的程序達到分享、使用與增值，有助於使用者更方便地檢索與獲取網路資源，滿足文獻資訊檢索的需求。OAI-PMH 並不是要取代目前的 Z39.50 等開放式資料存取標



準,而只是提供一個容易實作和容易部署的替代性解決方案。它的出現只是要讓 Internet 網路上這麼多的資訊服務提供者,以一組標準的陳述語法,可以讓不同的資訊服務提供者及資料提供者互相溝通沒有阻礙。透過這樣一個標準,數位圖書館、學

術單位之間就可以在一共通的協定之下,輕易的分享與取得電子文件相關的典藏資訊或全文,進而提供使用者正確且快速地取得所需資訊的服務。

(收稿日期:2002年9月12日)

註釋:

- 註 1: Xiaomin Liu et al., "Arc-An OAI Service Provider for Digital Library Federation," D-Lib Magazine 7:4(April 2001), <<http://www.dlib.org/dlib/april01/liu/04liu.html>>.
- 註 2: Paul Ginsparg, Rick Luce, and Herbert Van de Sompel, Call for participation in the UPS initiative aimed at the further promotion of author self-archived solutions, July 1999, <<http://www.openarchives.org/ups-invitation-ori.htm>>.
- 註 3: Herbert Van de Sompel, Carl Lagoze, "The Santa Fe Convention of the Open Archives Initiative," D-Lib Magazine 6:2 (Feb. 2000), <<http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>>.
- 註 4: C. Mic Bowman et al., "The Harvest Information Discovery and Access System," Computer Networks and ISDN Systems 28:2(Dec. 1995), pp.119-125.
- 註 5: The Open Archives Initiative Protocol for Metadata Harvesting, v.1.0, <<http://www.openarchives.org/OAI/1.0/openarchivesprotocol.htm>>.
- 註 6: The Open Archives Initiative Protocol for Metadata Harvesting, v.1.1, <<http://www.openarchives.org/OAI/1.1/openarchivesprotocol.html>>.
- 註 7: The Open Archives Initiative Protocol for Metadata Harvesting, v.2.0 <<http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>>.
- 註 8: 余顯強、陳瑞順,「開放式檔案存取協定-OAIMH」, 書藝 38(民國 91 年 5 月), 頁 26。
- 註 9: 同前註。
- 註 10: S. Warner, "Exposing and Harvesting Metadata Using the OAI Metadata Harvesting Protocol: A Tutorial," High Energy Physics 4(June 2001), <<http://library.cern.ch/HEPLW/4/papers/3>>.
- 註 11: Cite Base, <http://cite-base.ecs.soton.ac.uk/>.
- 註 12: Clifford A. Lynch et al., "Metadata Harvesting and the Open Archives Initiative", ARL Bimonthly Report 217, (Aug. 2001), <<http://www.arl.org/newsltr/217/mhp.html>>.
- 註 13: 同註 8, 頁 27-28。
- 註 14: Herbert Van de Sompel et al., "The UPS Prototype: An Experimental End-User Service across E-Print Archives," D-Lib Magazine 6:2(Feb. 2002), <<http://www.dlib.org/dlib/february00/vandesompel-ups/02vandesompel-ups.html>>.
- 註 15: 同註 1。
- 註 16: National Science Digital Library, <<http://www.ehr.nsf.gov/duet/programs/nsdl/>>.



- 註 17 : A. Paepcke, et al., "Search Middleware and the Simple Digital Library Interoperability Protocol," D-Library Magazine 6:3 (Mar. 2000), <<http://www.dlib.org/dlib/march00/paepcke/03paepcke.html> >.
- 註 18 : Cyclades, an open collaborative virtual archive environment, <<http://www.ercim.org/cyclades/>>.
- 註 19 : Digital Library Federation, DLF evaluation of the Open Archives Initiative, <<http://www.clir.org/diglib/architectures/testbed.htm>>.
- 註 20 : 同註 12。

蔡其基
Chang-ying Tsai

臺灣師範大學圖書館學系教授

Section Chief, The Library, Chang Jung Christian University

tsaiman@mail.cjcu.edu.tw

陳其基

Leong-yeon Chin

國立中央大學資訊管理系教授

Associate Professor, Department of Information and Library Science, National Central University

【摘要/Abstract】

近年來，隨著資訊技術的飛躍發展，數位化資源的獲取與利用已成為圖書館界關注的焦點。本文探討了開放式元資料獲取協定（OAI）在數位圖書館建設中的應用。文章首先介紹了OAI的定義、標準及協定內容，並分析了其在提高館際互聯、促進資源共享方面的優勢。其次，文章討論了OAI在數位圖書館建設中的應用現狀，包括其在館際互聯、資源共享、館藏建設等方面的實踐。最後，文章對OAI在數位圖書館建設中的應用前景進行了展望。

Over the past few years, the Taiwan government has been pushing an agenda of massive infrastructure construction projects aggressively. In order to alleviate the financial burden on government fiscal tanks as well as to improve the efficiency and effectiveness of the private sector, many of these projects are carried out with the Build-Operate-Transfer (BOT) model since the decision of "Strains for the Management of Private Sector Participation in Public Infrastructure Projects" by the Executive Yuan. This article investigates the feasibility of applying the BOT model in public library projects. Issues on the management are also elaborated using the Kaohsiung City for Library for the Monthly Bookend Children and the Yang-Ming Children's Discovery Community Library as objects of study. The authors seek to add to the volume of the research and literature on the application of public library projects with the BOT methodology.

關鍵字/Keywords

開放式元資料獲取協定、數位圖書館、資訊管理、圖書館學
BOT ; Public library ; Risk management ; Outsourcing

