

行政院國家科學委員會專題研究計畫成果報告

整合式 XML 文件管理系統及其在 Web 出版之應用 (II)

Integrated XML Document Management System and its Application to the Web Publishing (II)

計畫編號：NSC89-2413-H-032-017

執行期限：90 年 8 月 1 日至 91 年 7 月 31 日

主持人：林信成 執行機構：淡江大學資圖系

壹、中文摘要

本計畫旨在利用 XML 簡易、可鏈結、具結構性及可自訂標籤等特點，針對 XML 及現有的網路電子文件之間的關連及運作細節進行研究，藉由系統化的分析過程，設計一個兼具學術價值及研究性質的「整合性 XML 文件管理系統」，並探討其應用在 Web 線上出版之可行性，從而建立一個適用於電子出版的實驗環境。

本研究採實驗研究法，主要在針對電子文件中缺乏結構性、缺乏資料著錄、缺乏檢索效能、無法進行加值處理等不足，以 XML 建構符合網路電子出版需求的實驗系統。其主要貢獻在於遵循現行標準，進一步探討 XML 的設計原則及精神，以擴展 XML 未來的應用空間。本計畫所建立的實驗系統可為 Metadata 的研究建立一個實作的範例，此一實驗系統具備了資料庫的優點，便於日後的存取、蒐尋、傳遞，成為具有結構的文件，這樣的文件可以依欄位、關鍵字進行搜尋，可以截取文件中特定屬性的內容，增加資訊檢索的回現率(Recall Rate)及精確率 (Precision Rate)，並可提高資訊管理的效率。

關鍵詞：XML、標示語言、Metadata、電子出版

貳、緣由與目的

1996 年 7 月「XML 工作小組」(XML Working Group) 在 W3C (World Wide Web Consortium, 全球資訊網協會) 的贊助下成立(註 1)，當年 11 月提交 XML 初稿，並於 1998 年 1 月 10 日正式通過 XML 1.0 規範，成為 W3C 的一個建議標準(Recommendation)。由於 XML 具有可擴展性、結構性、自我描述性，並採用資料和樣式分離原則，使其在資料的管理、交換上擁有極為卓越之性能。XML 和 HTML(註 2)一樣都是從 SGML 演變而來的，只不過 HTML 是 SGML 的一個應用語言 (Application)，而 XML 卻是 SGML 的一個精簡子集 (Subset)。XML 將 SGML 去蕪存菁，捨棄約百分之二十複雜罕用的部份，承襲了其他百分之八十的特點，是以具備了 SGML 所沒有的簡易性與靈活性，又有著 HTML 所欠缺的擴展性與結構性。因此，稱 XML 為主導「第二代 Web」(Second-Generation Web) 的重要技術實不為過(註 3)。

不過，XML 並不是被發展出來取代 HTML 的，而是用以彌補其不足之處。依據 XML 的特性，可歸

納出以 XML 為核心技術的新一代 Web 出版將具備如下之特色：

- 電子文件具備自我描述性
- 電子文件更能有效整合
- 電子文件更具結構性
- 電子文件具備內容和外觀分離原則
- 標注語言具備多樣性及可擴展性

綜上所述，XML 不但能有效解決目前網路上電子文件的亂象，更有助於開創電子文件自動交換與傳遞的新契機。因此，有必要建立一個適合一般使用者使用的「整合性 XML 文件管理系統」(Integrated XML Document Management System)。藉由系統分析過程，將此一整合系統劃分為三個子系統。分別是：(1) 編輯子系統；(2) 核心子系統；(3) 出版子系統。其主要任務是基於 XML 規範，對電子文件的結構、內容、表現三要素，進行有效率的管理以利於展示、查詢、編排、維護等加值處理，藉由各個子系統中的編輯模組、剖析及轉換模組、檢視模組、排版模組、發行模組 ... 等，可以建立一個適用於網路環境的電子文件整合出版系統。

系統方塊圖如圖 1 所示，其中除了 Web 子系統是供前台使用者使用外，其餘皆為後台管理與系統運作之用。

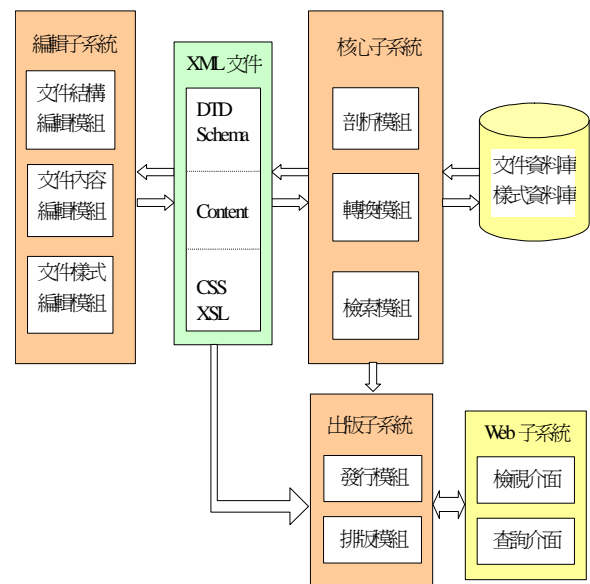


圖 1 整合式 XML 文件管理系統示意圖

參、結果與討論

經過兩年的研究與討論，前述的各項模組及子系統陸續完成，並實際植入以下的三個實驗系統之中進行運作：

- (1) 基於 XML 之資料交換系統(註 4)；
- (2) 基於 XML 之新聞管理與出版系統(註 5)；
- (3) 以 XML 與 CMARC 簡篇為基礎之編目系統；

茲分述如下。

(一) 基於 XML 之資料交換系統

一個理想的資料交換系統是由 N 個分散於網路上的異質系統所組成，彼此透過相同的協定進行資料傳遞與共享，如圖 2 所示。

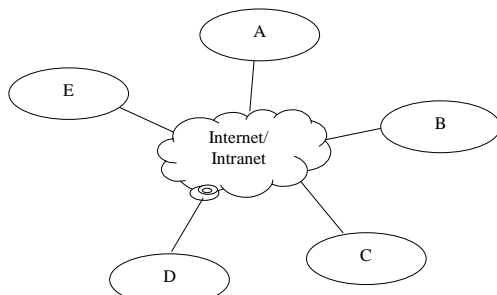


圖 2 理想的資料交換系統

為了簡化起見，本資料交換系統，其架構如圖 3 所示，圖中以「系統 A」及「系統 B」代表兩個分散於網路上的異質系統，彼此透過 XML 進行資料的傳遞與交換，若將此系統加以擴充，則可成為 N 個彼此透過 XML 進行資源共享的一般化整合系統。所規劃並擬完成之雛形系統，共分為七大功能模組，分別是：

- (1) 資料檢索模組 (Data Searching Module)
- (2) XML 轉換模組 (XML Transforming Module)
- (3) 資料傳送模組 (Data Transferring Module)
- (4) XML 剖析模組 (XML Parsing Module)
- (5) XML 資料呈現模組 (XML Data Presentation Module)
- (6) 資料儲存模組 (Data Storage Module)
- (7) Metadata 轉換模組 (Metadata Transforming Module)

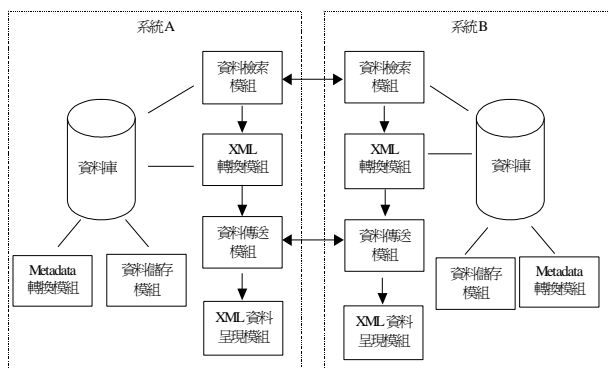


圖 3 XML-Based 資料交換系統

上述七模組彼此分工合作，其主要功能說明如下：

(1) 資料檢索模組 (Data Searching Module)

提供使用者檢索所需之資料。此模組為一聯合檢索機制，若是在系統 A 內檢索的資料不足或是查無使用者所須之資料時，使用者可啟動本模組之聯合檢索功能，可檢索遠端資料庫內之資料（即系統 B）。而檢索之結果將透過 XML 轉換模組以 XML 格式輸出，透過 XML 瀏覽器呈現，因此這份資料交換模型利用了 XML 在資料交換上的優勢，提供了極佳的 XML-Based 聯合檢索機制。

(2) XML 轉換模組 (XML Transforming Module)

XML 轉換模組與資料檢索模組關係密切，負責將資料檢索模組從資料庫中取出的資料轉換為 XML 資料，而其轉換後的 XML 資料結構完全依照彼此間所訂定共通之 DTD 規範來做資料的轉換，使得所有 XML 資料皆為合法的。

(3) 資料傳送模組 (Data Transferring Module)

資料傳送模組負責兩系統之間 XML 資料的傳送。XML 轉換模組所轉換後之 XML 資料將透過此模組傳送至對方之 XML 剖析模組，以做進一步之資料檢驗。

(4) XML 剖析模組 (XML Parsing Module)

在本系統中，XML 剖析模組負責檢驗所交換的 XML 資料是否具有合法性 (Validation)，以及是否符合系統間所訂定之共通 DTD。本模組首先將驗證此 XML 資料是否正確使用系統間所訂定共通資料結構規範之 DTD，接著將依照此共通之 DTD 來嚴格檢驗此文件是否為一合法的 XML 文件。若此文件為合法的 XML 文件，剖析器將把驗證過的 XML 文件交由 XML 資料呈現模組將資訊呈現給使用者。

(5) XML 資料呈現模組 (XML Data Presentation Module)

XML 資料呈現模組負責將資料內容呈現給使用者，並且具備解讀 XML 文件的功能。此模組透過 DSO(註 6)以及 DOM(註 7)來解讀 XML 文件，並結合該資訊所需之相關功能及超連結，並將之加以包裝、排版。本資料呈現模組具備相當大的相容性，只要是符合系統之間共通之 DTD 規範的 XML 文件皆可透過它來呈現，充分運用 XML 在資料交換上的優勢，達成資料可互通、交換使用的願景。

(6) 資料儲存模組 (Data Storage Module)

資料儲存模組亦具備了與 XML 資料呈現模組相同的 XML 文件解讀的能力。資料儲存模組能將所解讀後的資訊轉存回系統內之資料庫。資料儲存模組是相當重要的一環，在進行資料交換的過程，XML 是其中的橋樑，而資料的儲存模組將是其中的推手，亦是資料交換的終點站。

(7) Metadata 轉換模組 (Metadata Transforming Module)

進行資料交換的場合相當的廣泛，尤其是在開放性系統中，有可能遇上跨平台的資料的交換，而且不同的平台間，有可能資料的格式亦有所不同，此時便有賴訂共同的 Metadata 格式來加以規範。(註 8)有鑑於此，本研究也建構了一套 Metadata 轉換模組，透過使用者建立一份系統專屬的 Metadata 對照表，可彈性的針對內部的 Metadata 做外部異質的 Metadata 轉換，以 XML 的格式輸出，提供了跨平台 XML-Based

資料交換機制的雛形。

在典型的三層式 Web 架構上(使用者端— Web 伺服器—資料庫伺服器),資料往往是直接靠 CGI(註 9)、ASP(註 10) … 等中介軟體(Middle-ware)從資料庫取出直接呈現在網頁上。若是兩個傳統的 Web 平台欲進行資料的交換,除了直接把資料庫共享給對方之外,似乎較無其他方便的做法。而若是又遇上雙方的資料庫為異質平台的資料庫時(例如 Microsoft SQL Server 與 Oracle)所遇上資料交換的瓶頸將更大,且這樣的方法在資料庫的管理與安全上亦增加了困擾。而本研究之 XML-Based 資料交換系統在資料的格式上將資料庫中的資料經過了 XML 轉換,並且使用者介面也整合了 XML 技術,可直接對 XML 資料做解析,並將資訊呈現給使用者。只要系統與系統之間訂定共通之 DTD 來規範與驗證 XML 檔案的合法性,並於各系統上建構 XML 轉換引擎以及 XML 資料呈現模組(閱讀器)等 XML 相關介面,即可實現了新世紀 XML-Based Web 資料交換、資訊共享的願景。

(二) 基於 XML 之新聞管理與出版系統

本系統以電子新聞的管理與出版為例,藉由自訂的 Metadata 格式,以 XML 語法進行實驗性新聞資料庫之全文標誌,並自行設計一套新聞管理與出版系統,實際在 Web 環境中整合 XML 技術,探討與印證 XML 在電子文件管理與出版方面的優勢。此系統之特色為各個管理與出版模組皆以 XML 為基礎,系統內之所有資料亦採用 XML 格式,相較於傳統的資料處理模式來說,有著更彈性與更易加值處理的特點!再者,藉由 XML 優越的結構化與自我描述性,使得電子文件的「智慧化」程度得以提升,進而增進資訊檢索之精確度。

在新聞資料庫的 Metadata 方面,國內目前只有政大新聞系謝瀛春教授發表過有關科學新聞的內容標誌(註 11),國外則是以 NITF (News Industry Text Format)(註 12)與 XMLNews(註 13)為兩大主流,概述如下:

(1) 科學新聞內容標誌

此內容標誌依據新聞學及新聞寫作的相關原理,運用 XML 標記新聞內容,並以純淨新聞體裁之科學新聞為樣本,進行標誌工作。在新聞事件(event)的陳述上,以人(who)、事(what)、時(when)、地(where)、如何(how)、為何(why)六大方向來描述,將新聞內容以此為表達的重點,讓使用者清楚的明白發生的新聞事件始末。

(2) XMLNews

XMLNews 分成兩部份:XMLNews-story 定義新聞的內容,XMLNews-meta 則敘述新聞稿件。XMLNews-story 是借用另一個 XML 規格 NITF 而來的,事實上是 NITF 的 Subset。(註 14)NITF 是在 1998 由許多新聞機構及美國報業協會,共同研製的一個 XML 標準規格。NITF 的訂定是要取代以列印為主的 ANPA 1312 老舊規格(註 15),但由於 NITF 太龐大,許多 Tag 大多數人都不會用到,並且不夠靈活,要加一個 Tag 都不容易,所以才會有 XMLNews 的出現。

XMLNews-meta,是依據 W3C 標準的規則 RDF (Resource Description Framework) 制訂的,RDF 是

以物件導向的理念,用 XML 作規範,制訂出來一套描述資料的規則。XMLNews-meta 是描述新聞稿件的規格,其 DTD 中的 Element 有新聞機構代碼與名稱、版權、使用權、作者、新聞類別、產業名稱、語言名稱與代碼、人物名稱、發稿時間等多項。(註 16)

為了驗證經由 XML 註錄 Metadata 之電子文件,能有效提升檢索系統之精確度,本研究以一實驗性質之新聞資料庫為例,除對每篇新聞內容以 XML 註錄相關之 Metadata 外,並在其全文中以 XML 加註描述性標籤,此外,藉由自行設計的新聞管理與出版系統,測試加註了這些描述性資料之後的檢索結果。本系統之核心在於設計一套以 XML 為基礎之新聞管理與出版系統,因此,為了使系統開發及實驗能順利進行,本系統制訂了一個簡易的新聞資料庫 DTD,作為新聞資料之組織、檢索、出版之基礎。

本研究所規劃並擬完成之系統,其介面共分為前端使用者介面與後端管理者介面,如圖 4 所示,若依功能劃分則可區分為四大功能模組,分別是:

- (1) 資料出版模組 (Data Publishing Module)
- (2) 資料檢索模組 (Data Searching Module)
- (3) 資料編輯模組 (Data Editing Module)
- (4) 資料管理模組 (Data Management Module)

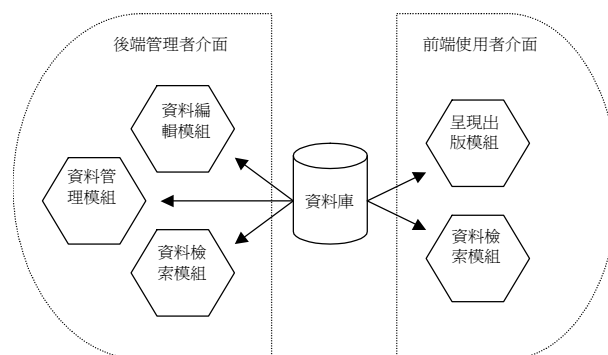


圖 4 基於 XML 之新聞出版系統功能模組

上述四大模組彼此分工合作,其主要功能說明如下:

(1) 資料出版模組

負責將資料內容呈現給使用者,並且具備解讀 XML 文件的功能。此模組透過 DSO 以及 DOM 來解讀 XML 文件,並結合該資訊所需之相關功能及超連結,加以包裝、排版,只要是符合系統之 DTD 規範的 XML 文件,皆可透過此模組呈現內容。

(2) 資料檢索模組

提供使用者檢索所需新聞資料。一般檢索模組僅提供欄位的檢索功能,並未提供針對內容某些特定目標加以檢索,如人名、地名等,本模組除提供新聞類別、關鍵字檢索功能之外,藉由 XML 將文件結構化的特性,可針對新聞內容的人、事、時、地、物加以檢索,提高檢索結果的查準率。

(3) 資料編輯模組

本模組提供管理者編輯新聞文件內容。對於 XML 並不了解或不熟悉者,皆可透過此模組,輕易的將所需之新聞內容編輯成符合系統之 DTD 規範的 XML 文件,並針對新聞內容給予人、事、時、地、物不同的

標記，提供檢索模組使用，另外經由 XML 文件內容與呈現資料分離的特點，同一份文件可選擇不同樣式做為出版的選擇。

(4) 資料管理模組

提供管理者異動／修改資料。透過網路可遠端開啟管理模組，對所需的新聞資料做新增、修改與刪除等動作，另外，對於異動過的資料，藉由資料管理模組即可做查詢的動作，檢視其 XML 內容是否可正確呈現，無需回到一般使用者介面。

透過新聞管理與出版系統的實作，將各資料模組以 XML 為基礎，系統內所有資料亦採用 XML 格式，以 XML 將新聞作結構化的處理，並自訂 Metadata 來描述其內容，搭配資料檢索模組，可確實針對新聞內容作精確檢索，其精確度優於傳統的全文檢索結果。此外，由於 XML 資料與樣式分離的特性，使得新聞呈現的樣式非常有彈性，可因不同的使用者需求做更改，而不用更動原始的新聞資料內容。以奇摩新聞為例，其合作的對象有聯合新聞網、中央社、中時電子報、台灣日報及路透社等國內外新聞媒體(註 17)，發佈其即時與每日新聞，在這樣的一個合作關係之下，每家媒體有著不同的新聞格式與其排版方式，一但需要發佈在同一網站之上，必定另外制定合作的標準，才能解決彼此間不相容的問題；若運用本新聞管理與出版系統，只需遵守 DTD 的規範，同一份新聞內容，就可依照不同的需求，迅速方便的在網站上出版，而無需另外做內容上的更改。

(三) 以 XML 與 CMARC 簡篇為基礎之編目系統

以中國機讀編目格式(CMARC)簡篇第四版為標準，提出一份可行之 DTD 作為書目資料庫結構，並透過系統實作方式，以 XML 為基礎架構一個編目系統，其目的在探討 XML 應用於圖書館自動化系統之潛力；此外，藉由系統實作，亦展現 XML 在結構化資料組織方面之優越性能。

CMARC 長久以來一直是圖書館自動化系統所採用的編目標準之一，但由於國內自動化系統並不一致，所以往往在做書目資料交換時會產生許多轉換上的問題。CMARC 的漸趨複雜，也使得資料的建檔格外變得有選擇性，書目資料在線上的展現模式，不管是簡略(brief)或是以簡短款目(short entry catalog)的呈現方式，若以使用者的使用效度來考量，編目員建檔並不需要建立多完整或多完美的書目記錄才是唯一的編目目標，事實上讀者只需要一些有效的檢索要項，每筆資料也只要一些具可獲性的檢索點，這就足以幫助讀者找到資料並發揮線上目錄的功能(註 18)。因此簡略編目對小型及地方圖書館來說，已足以滿足讀者的需求，並且使得小型自動化系統更容易開發，節省圖書館經費。

中國機讀編目格式可視為 Metadata 的一種，它是描述書本資料的資料，而 XML 是實現 Metadata 的語法(syntax)之一，要使 Metadata 具備互通性，可以在不同系統之間交換，則包裝 Metadata 的語法是非常重要的部份。猶如 ISO 2709 在不同的圖書館自動化系統之間穿梭自如，HTML 是使得 WWW 文獻能在不同系統間交換的主要功臣、而 SGML、XML 是電子圖書館/博物館/檔案館系統用來標示其 Metadata 及全

文資料的標準語法(註 19)。

本研究所採用的中國機讀編目格式為第四版(以下簡稱 CMARC4)，DTD 的內容是以 CMARC4 簡篇(註 20)為標準，再參考國家圖書館所制定之「NBINet 核心書目記錄必備項目說明」中的必備項目，並考慮系統的需求相互對照而制定的。

依照 CMARC4 的規範，一筆書目記錄由多個欄位所構成，而其欄位是由欄號、指標、分欄三個部分組合而成，以 XML 表示可更清楚其結構關係：

```

<CMARC4>
  <欄位名稱>
    <指標名稱/>
    <分欄名稱/></分欄名稱/>
  </欄位名稱>
  <欄位名稱>
    .....
  </欄位名稱>
  <欄位名稱>
    .....
  </欄位名稱>
</CMARC4>

```

圖 5 CMARC 記錄內容結構

本研究依照 CMARC4 的特性及其資料結構，在設計 DTD 時有三個要點如下：

(1) 指標為每一欄之第一個資料單元，用以指示該欄之內容。所以並不包含任何的資料，所以需宣告為空標籤。

(2) 為了使 DTD 可讀性增強，讓不熟悉 CMARC 的人也能了解，除分欄值使用英文字母外，其餘標籤皆用 CMARC4 中所定義的中文名稱。

(3) 以 CMARC4 簡篇格式為主，再參考國圖 NBINet 核心書目記錄必備項目，交叉對照後，整理出 DTD 中必備項目，如圖 6 所示。

欄號	指 標		分欄	欄、指標、分欄名稱
	1	2		
001				記錄識別欄
100			a	一般性資料 一般性資料
101	0 1 2		a	作品語文 原文 譯文 含譯文 正文語文
200	0 1		a f	題名及著者敘述項 題名不適宜做為檢索款目 題名做為檢索款目 正題名 第一種著者敘述
205			a	版本項 版本敘述
210			c d	出版項 出版者名稱 出版日期
215	0 1		a	稽核項 以中文著錄 以中文以外之語文著錄 數量
805			b	館藏記錄 次層單位代碼

圖 6 DTD 必備項目

此系統共分為六大模組，分別為：

- (1)編目管理模組：管理者可新增、刪除、修改、查詢資料。
- (2)資料查詢模組：提供使用者查詢資料。
- (3)XML 轉換模組：將資料庫的資料轉換為 XML 格式。
- (4)DTD 驗證模組：引用 DTD 驗證轉換所得之 XML 文件是否正確。
- (5)XML 格式呈現模組：將 DTD 驗證通過之資料以 XML 格式呈現。
- (6)HTML 格式呈現模組：將查詢所得資料以 HTML 格式顯示。

圖 7 編目資料編輯

本研究所提出之 DTD 僅為實驗系統所用，並非公開討論後所制訂的共同標準，研究的目的除作為本系統之資料庫結構外，對於將來國內相關單位若欲制訂 CMARC DTD 之規範與標準，亦可作為一個參考用的草案。配合網路與資料交換的發展之下，CMARC DTD 的制定是必然的，只是一個標準的訂定並不是由個人或少數人的力量即可完成，需經歷各相關團體單位的討論、協商認可，且現時國內各領域的 DTD 發展也是考慮的因素之一，如何讓 CMARC DTD 與其他的 DTD 做相關性的對照，或是彼此間有個參考、轉換的標準，有待各單位的互動及配合，未來的發展還有很長的一段路要走！

肆、計畫成果自評

本研究以 Web 為基礎，首先在網路上建置了兩家虛擬書店，彼此透過 XML 進行資料的傳遞與交換，驗證了 XML 在 Web 資料交換上的便利，若將此系統加以擴充，則可成為 N 個彼此透過 XML 進行資源共享的一般化整合系統；再者，本研究透過新聞管理與出版系統的實作，將新聞內容以 XML 作結構化的處理，並自訂 Metadata 來描述其語意，搭配資料檢索模組，可確實針對新聞內容作精確檢索，其精確度優於傳統的全文檢索結果；此外，本研究以中國機讀編目格式簡篇第四版為標準，制定其適用之 DTD，透過系統實作，架構一個基於 XML 的簡易編目系統，對於小型及地方圖書館來說，不但足以滿足讀者的需求，且小型自動化系統容易開發，可省下相當龐大的經費。本系統將機讀資料與 XML 相結合，除可透過 Web 查詢機讀資料外，並以系統之 DTD 對 XML 內容加以驗證，達成資料之結構性與組織化。

本研究在兩年的研究過程中，由於需考慮的層面非常的廣泛，因此遭遇了許多困難。例如，要能精確的與後端資料庫相互對應，並與當初使用者所建置的結構相同，是系統設計的最大考驗；其次，應用主題的選擇也是困難點之一，在第一年結束時，原本計畫第二年的系統實作能以圖書館利用教育做為主題，將其文件以 XML 的結構儲存、傳送及管理，但在後來的研究過程中，基於資料交換與出版的重要性，而將實作主題轉向，分別完成應用於網際網路之資料交換、新聞管理與出版、CMARC 編目等三個系統，來加以驗證當初系統規劃的編輯子系統、核心子系統及出版子系統。此三個系統所完成的結果如下：

(一) 基於 XML 之資料交換系統

本系統共分為七大功能模組，分別是：

- (1)資料檢索模組
- (2)XML 轉換模組
- (3)資料傳送模組
- (4)XML 剖析模組
- (5)XML 資料呈現模組
- (6)資料儲存模組
- (7)Metadata 轉換模組

依據圖 1 的內容，上述七大模組中的資料檢索模組、XML 轉換模組、XML 剖析模組即為核心子系統。此外，本系統還增加了資料傳送模組，負責 XML 資料交換，是原先系統所未規劃的。

(二) 基於 XML 之新聞管理與出版系統

依功能劃分則可區分為四大功能模組：

- (1)資料出版模組
- (2)資料檢索模組
- (3)資料編輯模組
- (4)資料管理模組

其資料出版模組與資料管理模組之功能，與原先規劃之出版子系統相符。

(三) 以 XML 與 CMARC 簡篇為基礎之編目系統

此系統共分為六大模組：

- (1)編目管理模組
- (2)資料查詢模組
- (3)XML 轉換模組
- (4)DTD 驗證模組
- (5)XML 格式呈現模組
- (6)HTML 格式呈現模組

上述之編目管理模組，其功能即為編輯子系統之需求。

總之，本研究內容不但與原計畫中的預期目標相符合，且在實際個別建構系統之後，更能深入瞭解 XML 的特點。並且藉由系統實作的驗證，更加深應用 XML 的信心，期望在不久的將來，XML 發展更成熟之際，能全面的應用於網路之上！

伍、註釋

註 1 「XML 工作小組」最初稱為「SGML 編審委員會」(SGML Editorial Review Board)。

註 2 HTML 規範由「全球資訊網協會」(World Wide Web Consortium, 簡稱 W3C)所制定，最新的建

- 議規範 "HTML 4.01 Specification, W3C Recommendation 24 December 1999", 可於 <<http://www.w3.org/TR/html4/>>取得。
- 註 3 Jon Bosak and Tim Bray, "XML and the Second-Generation Web", Scientific American, May 1999, also available at <<http://www.sciam.com/1999/0599issue/0599bosak.html>>.
- 註 4 林信成、陳勇任，「基於 XML 之網際網路資料交換離形系統設計」，教育資料與圖書館學 39 卷第 2 期（民國 90 年 12 月），頁 145-160。
- 註 5 林信成、陳勇任、楊翔淳。「基於 XML 之新聞管理與出版系統設計」。淡江大學資訊與圖書館學系編。出版與圖書館學術研討會論文集。臺北縣：淡江大學資訊與圖書館學系，民國 91 年。
- 註 6 DSO (Data Source Object, 資料來源物件) 可將 XML 文件視為一份文件資料庫進行資料存取的动作。
- 註 7 DOM (Document Object Module, 文件物件模型) 為一 W3C 所制定的介面標準。DOM 並非是針對 XML 量身訂做的，而是一套普遍適用於 HTML、XML 等文件的應用程式介面 (Application Programming Interface, API)。
- 註 8 Metadata, 在英文文獻中, 最常見的解釋為「data about data」, 直譯成中文則為「有關資料的資料」或「描述資料的資料」。目前已發展成熟或正發展的 metadata 格式眾多, 適用於不同領域及用途, 詳細內容可參考: 「陳雪華, 網路資源與 Metadata 之發展, 圖書館學刊第 12 期」。
- 註 9 CGI (Common Gateway Interface, 共通開道介面), 為一伺服器端與用戶端之間資料處理的介面。而所開發出來用於 Web 上處理伺服器端與用戶端之間資料的程式也俗稱 CGI 程式。
- 註 10 ASP(Active Server Pages, 動態伺服器網頁), 功能類似 CGI, 為 Microsoft 公司所開發, 專用於 Microsoft Web 平台。
- 註 11 謝瀛春、黃學碩、維習安、雷約翰、謝清俊, 「新聞內容的標誌-XML 之應用」, 海峽兩岸資料庫/數據庫與資訊/信息服務交流與合作論文集 (民國 90 年 1 月), 頁 205-212。
- 註 12 "NITF News Industry Text Format", available at <<http://www.nitf.org/site/index.html>>.
- 註 13 "XMLNews.org", available at <<http://www.xmlnews.org/>>.
- 註 14 "XMLNews.org XMLNews Specifications", available at <<http://www.xmlnews.org/XMLNews/>>.
- 註 15 "nitf introduction", available at <<http://www.nitf.org/site/intro.html>>.
- 註 16 "XMLNews-Meta Documentation", available at <<http://www.xmlnews.org/docs/xmlnews-meta.html>>.
- 註 17 "Yahoo! 奇摩新聞", available at <<http://tw.news.yahoo.com/>>.
- 註 18 魏令芳, 「簡略編目的發展與趨勢」, 大學圖書館 4:1 (民國 89 年 3 月), 頁 114。
- 註 19 陳昭珍, "XML, Metadata 與檔案資料數位化," <http://archives.sinica.edu.tw/main/article06.html> (1 Mar. 2002).
- 註 20 中國機讀編目格式修訂小組, 中國機讀編目格式=Chinese MARC format (附錄) 第四版 (國家圖書館, 民國 86 年), 頁 51-55。