

中文摘要

隨著文明的進展，資訊累積的速度越來越快，資訊量亦越來越大。對於紙本資訊，圖書資訊學已經發展出有效的資訊組織方法；然而對於數位資訊，仍然沒有一套可資遵循的作法。近來，由於數位資訊所佔的比重越來越大，各國學者專家積極投入數位資訊組織的研究，並借重圖書資訊學對於紙本資訊處理的理論，從事詮釋資料(metadata)的研究。基於中文藏品的特殊性，必須發展適合中文資訊整理與組織的模式，包括：1. 從資訊加值的角度，探討與處理資訊之屬性，並建立中文資訊詮釋格式以及著錄規則；2. 從資源分享的角度，建立詮釋資料交換之共通架構，並建構詮釋資料管理工具。

我國國科會自八十七年開始進行「數位博物館」專案計畫。有一些圖書資訊學界的教授們，共同執行該專案計畫之下的一個分項計畫，名為：「資源組織與檢索之規範」(Resources Organization and Searching Specification, 簡稱 ROSS)。ROSS 之研究目標涵蓋在中文環境數位圖書館/博物館中，有關於資訊組織與檢索各個層面的重要議題；包括資料的儲存與管理系統的設計、使用者資訊需求及查詢資料的方式、各系統間整合檢索等。

ROSS 詮釋資料研究群在國科會「數位博物館專案計畫」中已獲得一些研究成果，包括 1. 以 Dublin Core 為核心欄位的中文詮釋資料格式(Metadata Interchange for Chinese Information, 簡稱 MICI-DC)；2. MICI-DC XML DTD；3. 詮釋資料管理系統 Metalogy (XML/metadata 輸入及資料庫管理)雛形系統；4. MICI Z39.50 Profile；5. Z39.50 Client /Server 雛形系統等。本計畫將近一步發展詮釋資料管理系統 Metalogy, Z39.50 Client /Server 系統，並發展 MICI 的使用者著錄手冊，做為在語意上詮釋資料著錄的參考。

根據美國陳劉欽智教授的說法，目前美國與中國大陸在中文詮釋資料格式的發展方面遠較台灣緩慢，因此 ROSS 詮釋資料研究群受邀參與本國際計畫，並主導中文詮釋資料格式的發展。本研究群期望藉著參與本計畫，探討多語言詮釋資料資訊查詢檢索的可行性，以便與國際間作交換。另外，希望所規劃設計之詮釋資料，不僅可以符合國科會「數位博物館」專案各個主題計畫中藏品的特性，還可以與國際上其他國家數位圖書館/博物館之詮釋資料相容。更期望所發展的詮釋資料管理系統 Metalogy，可提供美國、中國大陸相關館藏單位實際輸入資料之用。

關鍵字：數位圖書館，詮釋資料，資訊檢索，詮釋資料管理系統

Abstract

With the progress of civilization, information accumulates rapidly at an explosive rate. Furthermore, it is essential to design a user-oriented, information resource learning system to enhance users' information skills. Concerning printed information, Library and Information Science has developed effective information organization methods; however, for digital information, standard method is yet to emerge. Recently, due to rapid growth of digital information, researches on digital information organization are carried out in many developed countries, and many researchers are trying to formulate metadata based on printed information organization theory developed by Library and Information Science. Due to the uniqueness of Chinese language, our research team held that we should develop a model that will be appropriate for Chinese information organization, which include: First, from the aspect of value-added information, one has to investigate and handle information attributes, establish metadata formats for Chinese information resources, and provide a guide to best practice. Second, from the aspect of resource sharing, one has to develop an interoperable metadata exchange framework and metadata management tools.

National Science Council (NSC) launched the Digital Museum Project in 1998. Several professors from Library and Information Science worked together under a sub project called Resources Organization and Searching Specification (ROSS). ROSS's aims are to study relevant issues on information organization and retrieval in Chinese digital libraries/museums (DL/Ms), which include data storage and management system design, user demand and retrieval behaviors, and integrated retrievals across different systems.

Members of ROSS metadata research team consist of professors, programmers and research assistants. Results achieved in 1998 included 1. development of a Dublin Core-based metadata format, Metadata Interchange for Chinese Information (MICI-DC); 2. MICI-DC XML DTD; 3. metadata management system prototype, Metalogy (XML/metadata input and database management); 4. MICI Z39.50 Profile; 5. Z39.50 Client/Server prototype. Besides, ROSS is working on MICI-DC tagging guide to provide semantic interpretations for the cataloguers. In addition, keeping updated about the development of major metadata (mainly Dublin Core), XML and RDF are also essential. We plan to invite staffs from related DL/Ms in Taiwan to participate in this

project by using MICI-DC and Metalogy system to record their collections. Furthermore, ROSS will hold discussion meetings periodically to gather users' opinions, so to further improve metadata format and Metalogy system.

According to Dr. Ching-chih Chen, current development of Chinese metadata format in US and China is slower than Taiwan. Thus, ROSS metadata research team was invited to participate in this international project and to take lead in the development of Chinese metadata format. Through the participation of this project, we hope to investigate the feasibility of multilingual information retrieval, so to enhance global information exchange. In addition to be compatible with various collection attributes of the DL/M Project of NSC, the metadata format design should be compatible with metadata from DL/Ms of other countries as well. It is hoped that Metalogy may be used and implemented by DL/Ms in the States and in China.

Keywords: Digital Libraries, Metadata, Information Retrieval, Metadata Management System

目 次

壹、計畫背景.....	6
貳、計畫目標.....	8
參、研究方法.....	9
肆、文獻分析.....	11
伍、研究成果.....	24
陸、結論.....	39
柒、參考資料.....	43

壹、計畫背景

最近數年，數位圖書館／博物館的研究隨著網際網路的發展相當受到世人的重視，各先進國家無不熱烈支持有關的各項計畫。我國有悠久的歷史文化，累積豐富的文化資產，是世界的瑰寶。另外，關於臺灣本土的珍貴文獻與歷史文物，許多機構與學術研究單位有豐富的典藏。為了避免損壞這些寶貴的資源，因此無法開放讓大眾參觀。現在，透過無遠弗屆的網際網路連結，能夠將這些寶貴的資源呈現於全球資訊網上，既能夠擴大瀏覽的族群，也能夠讓終將損壞的珍貴資源以另一種方式維持其恆久的生命。

臺灣地區目前從事珍貴藏品數位化的機構主要有臺灣大學、中央研究院、國家圖書館、故宮博物院、歷史博物館、科學博物館等。另外還有如臺灣科學教育館、臺灣博物館、科學工藝博物館、臺北市立美術館等單位，其部分館藏文物也都有數位化之成果。¹將這些寶貴的資源數位化並呈現於全球資訊網上是首要的工作，但是更重要的是：必須針對這些資源的特性加以組織與整理，才能有效地檢索、使用，因此，在數位化計畫中，metadata 扮演很重要的角色。

基本的數位化系統，從使用者的角度而言，應包含的功能主要有查詢、瀏覽及其他相關網路資源連結。而通常資料量大的數位圖書館系統都會以資料庫來管理其書目記錄及數位化物件、或連結的網址，為避免連結網址的失效，及數位化物件變更存放的位置，所以也會考慮使用 Handle System 之類的機制來做識別處理。

以目前的檢索技術而言，提供查詢的方法不外為：全文檢索或欄位查詢。全文檢索可以不必對資料做描述，但檢索效能較差，對於無文字的影像、聲音或視訊資料而言，全文檢索技術也無法派上用場，所以以人工描述 metadata，建立欄位化的書目資料，對於數位圖書館而言，是絕對必要的工作。因此，在一數位圖書館或博物館等系統中，metadata 扮演重要的角色。Metadata 的訂定需由瞭解使用者需求及資訊屬性入手，同時也需考慮到資訊系統之間的互通性，所以相關標準的採用非常重要，如國際間已發展出來的格式：Dublin Core (Dublin Metadata Core Element Set) (<http://purl.oclc/dc/>)、EAD (Encoding Archival Description) (<http://lcweb.loc.gov/ead/>)、FGDC (Federal Geographic Data committee) (<http://www.fgdc.gov/>)、GILS (Government Information Locator Service) (http://www.access.gpo.gov/su_docs/gils/index.html)、TEI (Text Encoding Initiative Headers) (<http://www.uic.edu/orgs/tei/>)；以及 metadata 的交換語言，如 SGML/XML/HTML 等。目前發展中的數位圖書館系統不管用那一種 metadata，

其交換語言大多選用 XML 或 SGML，尤其是 XML 語言可取 SGML 之長、補 HTML 之短，也是目前網際網路界極力推廣的語言。

貳、計畫目標

本計畫的目標涵蓋在中文環境數位博物館/圖書館中，有關於資訊組織與檢索各個層面的重要議題，包括：資料的儲存與管理系統的設計、使用者資訊需求及查詢資料的方式、各系統間整合等。本計畫目標除了支援國科會「數位博物館專案計畫」主題計畫中有關於資訊組織與檢索相關的部分，主要發展重點在於制定中文數位博物館/圖書館環境中資源組織與檢索之相關規範，以及發展一套能夠實做的系統。這些規範要考量能夠與國際相關標準相容的問題，使中文資訊能被國外之整合檢索軟體檢索，並促進國內數位博物館/圖書館系統國際化的重要管道。通透性的整合檢索是電子資訊檢索系統中不可或缺的一項功能，也是中文數位博物館/圖書館系統國際化相當重要的課題。因此，本計畫積極探討並開發符合國際標準之整合檢索系統，可提供台灣、美國、中國大陸相關館藏單位實際輸入資料之用。

總而言之，本計畫的目標可以概述於下列兩點：

- 一、從資訊加值的角度，探討與處理資訊之屬性，建立中文資訊詮釋資料格式，並制訂著錄規則，包括：資訊加值理論的建構、資訊屬性之內涵結構的分析、資訊屬性之語法、語意、語用，及其關係的探討，中文資訊詮釋格式欄位項目的分析、以及其語意著錄規則的發展等。
- 二、從資源分享的角度，建立詮釋資料交換之共通架構，並建構詮釋資料管理工具，包括：詮釋資料之理論建構、資訊交換之問題及需求分析、建構詮釋資料管理工具等。

參、研究方法

一、從資訊加值的角度，探討與處理資訊之屬性，建立中文資訊詮釋資料格式，並制訂著錄規則。此部份執行方法與實施步驟包括：文獻分析、觀察分析、訪問座談、參與相關組織、擬定詮釋資料、測試並修正詮釋資料、以及發展詮釋資料著錄的規則等。茲分別敘述如下：

(一) 文獻分析：

以文獻分析方式瞭解國內外數位圖書館/博物館的發展與特色；深入探討幾種與本計劃相關之詮釋資料，特別是目前仍在發展修飾語的 Dublin Core 等，也分析各詮釋資料間連結及重複利用的可行性。

(二) 觀察分析：

藉由原件、影印本及掃描之電腦檔，觀察分析珍藏文獻之特性，例如其原件類型、數量、規格、檔案結構、原件維護情形等。藏品觀察分析包括台灣、大陸與美國三地相關單位的藏品。

(三) 專家訪問與座談：

透過訪問、座談等方式以瞭解藏品之內容特徵及屬性，使詮釋資料格式設計較完善，以更精確地描述藏品相關資訊。專家訪問與座談包括在台灣、大陸與美國三地相關單位舉行。

(四) 參與相關國際組織：

繼續參與博物館資訊交換的機構，如 CIMI (The Consortium for the Computer Interchange of Museum Information)；以及網路資源詮釋資料之國際性計畫 DCMI (Dublin Core Metadata Initiative)，以期所設計的詮釋資料能夠符合國際標準。

(五) 擬定詮釋資料：

根據以上文獻分析、觀察分析、訪問座談所得之結果，著手擬定詮釋資料。

(六) 測試詮釋資料並修正：

將所擬定之詮釋資料初稿，提供在台灣、大陸與美國三地相關單位實際輸入資料做進一步測試，並針對不適之處做修正。

(七) 發展詮釋資料著錄規則：針對每一欄位做定義的制訂、語意的解釋，並提供著錄的例子，以供台灣、大陸與美國三地藏品著錄單位使用。

二、從資源分享的角度，建立詮釋資料交換之共通架構，並建構詮釋資料管理工具，

包括：詮釋資料之理論建構、資訊交換之問題及需求分析、建構詮釋資料管理工具等。

- (一) 文獻分析法：蒐集並分析國內外相關之 XML/Metadata 系統設計之相關文獻、技術與目前系統開發情況。並探討相關的研究成果，如 W3C 的 RDF (Resource Description Framework)，及所相關的技術利用。
- (二) 系統分析：系統分析部份主要包括：
 - 1. 確定此系統應包含的功能。
 - 2. 確定工作流程。
 - 3. 確定每一個部份所需要開發的程式。
- (三) 系統設計：
 - 1. 採用 Delphi 語言來開發主要的程式。
 - 2. 在 web 上的程式則採用 Java 語言來開發。
 - 3. 系統平台為 Windows NT、SQL server、以及 Oracle server。
- (四) Web 檢索引擎設計：

設計檢索引擎，提供使用者可以透過全球資訊網的介面來查詢資料庫的內容。
- (五) 測試詮釋資料管理系統並修正：

將所設計之詮釋資料管理系統，提供在台灣、大陸與美國三地相關單位實際輸入資料做進一步測試，並針對不適之處做修正。

肆、文獻分析

一、Metadata 之意義與功能

(一) Metadata 之意義

所謂 metadata，在資訊組織界最普遍的解釋是 "data about data"，意指有關資料的資料，即資料之描述性資訊，如圖書館的 MARC (Machine-readable cataloging) (<http://www.nlc-bnc.ca/marc/emarc.htm>) 記錄，即為一種 metadata。但 Renato Innella 認為這種解釋，忽略了 metadata 最重要的概念，即其結構性，他認為應將 metadata 定義為 "structure data about data" 此結構二字，使得採用 metadata 做組織資訊的方式和全文索引(full-text indexing)有所區隔。¹ 對資料做描述以建立索引的作法並非始於今日，但 metadata 一詞卻是在網路資訊出現後，才普遍被使用的，且多指網路資源等電子資訊的描述。國內翻譯名詞有詮釋資料²，元資料³，超資料⁴，紛歧顯見。

(二) Metadata 之功能

Locan Dempsey 認為 metadata 的功能主要有：⁵

1. 找到資訊之位址 (location)；
2. 蒐尋資訊 (discovery)；
3. 紀錄資訊 (documentation)；
4. 評估資訊 (evaluation)；
5. 選擇資訊 (selection)；
6. 其他。

Renato Iannella 及 Andrew Waugh 也提出類似的觀點，他們認為 metadata 主要用來：⁶

1. 概述資料之內涵
2. 讓使用者查到該資料
3. 讓使用者決定該資料是否是他所要的
4. 避免使用者存取該資料（如禁止兒童使用某類資料）
5. 讓使用者檢索、複製資料
6. 指示應如何解譯該資料（例如說明資料之格式、編碼、加密的情況）
7. 用來決定可檢索哪一個資料（若資料可多種格式存在）
8. 說明資料使用的合法情況
9. 說明資料之歷史，如說明其原始資料為何，及其他的改變

10. 說明資料之聯絡人，如擁有者
11. 指示該資料與其他資源的關係
12. 控制資料的管理

但就資料庫系統而言，metadata 則直指資料庫管理系統之架構(schema)，且其對 metadata 的討論偏重在如何藉由 metadata 的使用，以達到多資料庫系統(multi-databases)的互通(interoperability)等問題。香港科技大學的 Pamela Drew 及 Jerry Ying 認為傳統的 metadata 被定義為：有關資訊源的靜態觀點(a static view over some source information)，此觀點通常由 metadata 應如何讓使用者蒐尋及存取資訊源之角度來思考。但是現在，因整個廣域網路中資訊源需互連，故而傳統 metadata 的範圍與角色也隨之改變。所以，他們認為，現在的 metadata 應滿足三項需求：⁷

1. 促使系統互通，而不只是在提供摘要性資訊；
2. 當越來越多的資訊被電子化時，metadata 模組應能讓電腦連接資訊源並自動擷取 metadata；
3. Metadata 管理系統應能定期核對原始資訊源，以確保 metadata 資訊的正確性。

Jeff Rothenberg 認為和一資料庫有關之 metadata 有三類：⁸

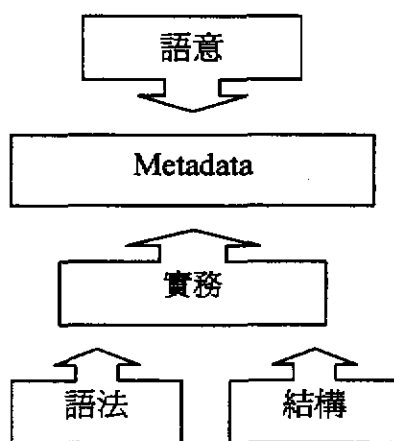
1. 協助或限制檢索的 metadata；
2. 作為分享與互通之用的 metadata；
3. 表達資料特性及作為索引資料的 metadata。

同時 Jeff Rothenberg 更進一步的將 metadata 分為三種層次，每一層中所包含之資訊如下：

1. 資料庫層次之 metadata (database level metadata)
2. 資料項目層次之 metadata (data-element level metadata)
3. 資料值層次之 metadata (data-value level metadata)

(三) Metadata 相關之課題

如果將 Jeff Rothenberg 的三種層次再進一步詮釋，我們也可以說 metadata 的設計必需考慮三方面的問題，一是語意(semantic)、二是結構(structure)、三是語法(syntax)。這三者的關係如下圖所示：



圖一：Metadata 的三個問題

此外，分散式資訊的整合檢索雖不是 metadata 的直接內涵，但在網路環境中是一個重要的課題，而且相關標準的設計和 metadata 息息相關。茲分別說明如下：

1. 語意問題 (semantics)：

Metadata 的語意內涵乃根據資料的特性及使用者的需求而定，它探討的問題包括應該描述資料的那些屬性？屬性的名稱如何命名？描述該屬性時所根據的資訊來源為何？以及該屬性內容是否應根據索引點或權威檔來描述？AACR2 (Anglo-American Cataloguing Rules Second Edition)(<http://www.libraries.psu.edu/iasweb/personal/jca/aacr/>)、Dublin Core、AAT (Art and Architecture Thesaurus)(http://shiva.pub.getty.edu/aat_browser/)、LCSH (Library of Congress Subject Heading)(<http://lcweb.loc.gov/cds/lcsh.html>)、LCNA (Library of Congress Name Authority File)等屬於此類標準。

2. 結構問題 (structure)：

決定了資料應描述屬性後，接下來的問題是建立屬性的結構。屬性的結構在呈現屬性之間的關係，使得屬性的描述清晰易懂、不疊床架屋，且便於程式的撰寫。MARC、DTD (Document Type Definition)、RDF (Resource Description Framework)(<http://www.w3.org/RDF/>)等所關心的即是這個層次的問題。

3. 語法問題 (syntax)：

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言。

4. 整合檢索的問題

即使每個系統都根據相同的 metadata 描述資料，但各系統的索引模式與檢索指令仍會不同，因此使用者往往必需學習不同系統的檢索介面及指令以查尋不同的系統。當資料庫很少的時後，這樣的學習還可忍受，若資料庫很多時，太多檢索指令及介面的學習則毫無效率可言。更何況在電子圖書館/博物館/檔案館領域中，不同學科領域或資訊類型往往使用不同的 metadata。因此分散式資訊的整合檢索在網路環境中是一個非常重要的課題。Z39.50⁹、OAI (Open Archives Initiative)、START (The Stanford Protocol for Internet Retrieval and Search) 等即屬此類的標準。

二、 台灣地區 Metadata 之發展現況

有關 Metadata 的研究在臺灣起步甚早，而實際設計 Metadata 則自 1987 年開始。在國科會的數位博物館專案開始之前，我們即於 1987 年 3 月在臺大電子圖書館與博物館 (National Taiwan University Digital Library/Museum, 簡稱 NTUDL/M) 計劃中成立了「Metadata 研究群」，著手於中文詮釋資料交換格式 (Metadata Interchange for Chinese Information, 簡稱 MICI) 之相關研究。Metadata 研究群於民國 87 年執行國科會「資源組織與檢索之規範」(Resources Organization and Searching Specification, 簡稱 ROSS) 計畫，89 年則繼續執行「資源組織與檢索之規範與系統實作」，作為國科會「數位博物館專案」之下技術支援的分項計劃。ROSS 計畫的研究目標涵蓋在中文環境數位圖書館/博物館中，有關於資訊組織與檢索各個層面的重要議題。¹⁰

而中央研究院也在民國 87 年 11 月正式成立「Metadata 工作小組」，負責中研院數位典藏計畫 (Digital Archive Project) 中有關 Metadata 的規劃、設計與實施等各項事宜。¹¹ 研究目的是要先整合中研院計算中心與資訊所各個相關實驗室的既有成果，參考國際研究趨勢，並且與研究同仁的密切溝通配合，建立一套符合全世界華文使用者習性的多媒體數位典藏的整合管理與服務環境，提供收數位典藏系統的基礎功能。¹² 中研院 Metadata 工作小組除了以中研院數位典藏計畫為主外，同時也支援其他相關計畫的 metadata，如「臺灣地震數位知識庫」，以及中研院內部的各項「數位圖書博物館計畫」。¹³

此外，在 1990 年故宮博物院的數位博物館計畫中，筆者與 Ross 研究群根據藏品的類型及系統的需求，設計了九份 Metadata，這些 Metadata 目前大部份皆已實際在使用。而在臺灣省公共圖書館自動化諮詢委員會的地方文獻數位化計畫中，筆者也設計了四份

符合各縣市收藏特色使用的 Metadata。^{14,15}

國家圖書館也於 1990 年成立「Metadata 研究小組」，主要的目的乃在：1. 探討國內外 Metadata 發展現況；2. 積極加入國內外 Metadata 研究組織及相關計畫；3. 發展各主題領域或資料類型之 Metadata；4. 研究 MARC 與 Metadata 轉換之語法、語意等相關標準；4. 發展 Metadata 管理系統。此外，國家圖書館也收集了國內各單位已發展出來的 Metadata 格式，出版《中文詮釋資料(Metadata) 格式彙編》¹⁶，根據該彙編，目前臺灣地區已設計出來的 Metadata 格式約有二十六種，如下所列：

1. 古籍善本詮釋資料及著錄範例（國家圖書館 Metadata 研究小組設計）
2. MICI-CD for NPM（數位博物館計畫資訊組織與檢索規範研究小組設計）
3. 器物 MICI-DC、DTD 及著錄範例（數位博物館專案故宮文物之美計畫 Metadata 小組設計）
4. 書畫 MICI-DC、DTD 及著錄範例（數位博物館專案故宮文物之美計畫 Metadata 小組設計）
5. 文獻 MICI-DC、DTD 及著錄範例（數位博物館計畫故宮文物之美計畫 Metadata 小組設計）
6. 人名權威檔及其 DTD（數位博物館計畫故宮文物之美計畫 Metadata 小組設計）
7. 主題權威檔及其 DTD（數位博物館計畫故宮文物之美計畫 Metadata 小組設計）
8. 地名權威檔及其 DTD（數位博物館計畫故宮文物之美計畫 Metadata 小組設計）
9. 時代權威檔及其 DTD（數位博物館計畫故宮文物之美計畫 Metadata 小組設計）
10. 參考書目、DTD 及著錄範例（數位博物館計畫故宮文物之美計畫 Metadata 小組設計）
11. 展覽、DTD 及著錄範例（數位博物館計畫故宮文物之美計畫 Metadata 小組設計）
12. 淡水河溯源（數位博物館計畫資訊組織與檢索規範研究小組設計）
13. 臺大圖書館古文書及其 DTD（臺大圖書館設計）
14. 蝴蝶生態面面（國立自然科學博物館、暨南大學、數位博物館計畫資訊組織與檢索規範研究小組設計）
15. 蘭嶼生物/文化多樣性數位博物館之詮釋資料（國立自然科學博物館、暨南大學蘭嶼生物/文化多樣性數位博物館計畫設計）
16. 平埔文化（中央研究院數位博物館專案平埔文化計畫設計）
17. 人文與自然資源地圖詮釋資料（中央研究院數位博物館專案人文與自然資源地圖計畫設計）
18. 傅斯年圖書館拓片詮釋資料及著錄範例（草稿）（中央研究院）
19. 傅斯年圖書館善本書明人文集詮釋資料及著錄範例（草稿）（中央研究院）

20. 傅斯年圖書館善本書詮釋資料及著錄範例(草稿)(中央研究院)
21. 國土資訊系統相關數值資訊詮釋資料規格及範例(內政部資訊中心設計)
22. 楊英風數位美術館詮釋資料(交通大學圖書館設計)
23. 戲曲唱片詮釋資料及 DTD(公共圖書館自動化及網路諮詢委員會地方文獻數位化小組及宜蘭文化局設計)
24. 皮偶劇本詮釋資料及 DTD(公共圖書館自動化及網路諮詢委員會地方文獻數位化小組及高雄縣文化局設計)
25. 地方發展古照片詮釋資料及 DTD(公共圖書館自動化及網路諮詢委員會地方文獻數位化小組及台中縣文化局設計)
26. 客家文物詮釋資料及 DTD(公共圖書館自動化及網路諮詢委員會地方文獻數位化小組及新竹縣文化局設計)

三、詮釋資料格式之設計步驟

Metadata 的設計並非新的觀念或新的事物，事實上，Metadata 的設計過程，類似設計資訊系統時都需做的系統分析以及檔案設計，而其內涵包括資料格式的語意及語法設計以及系統實作三部份。以下即就筆者設計故宮博物院 Metadata 之經驗，將 Metadata 的設計分為七個步驟分別說明之。

(一) 分析藏品或資料之屬性

設計 Metadata 的第一個步驟是了解藏品或要處理之資訊的性質，如要建立的是書畫資料庫，則需先具備書畫方面的知識，以及書畫這類「資料」的知識，並從中抽離出書畫的共同屬性，而不是以寫一幅書畫鑑賞文章的角度去思考。

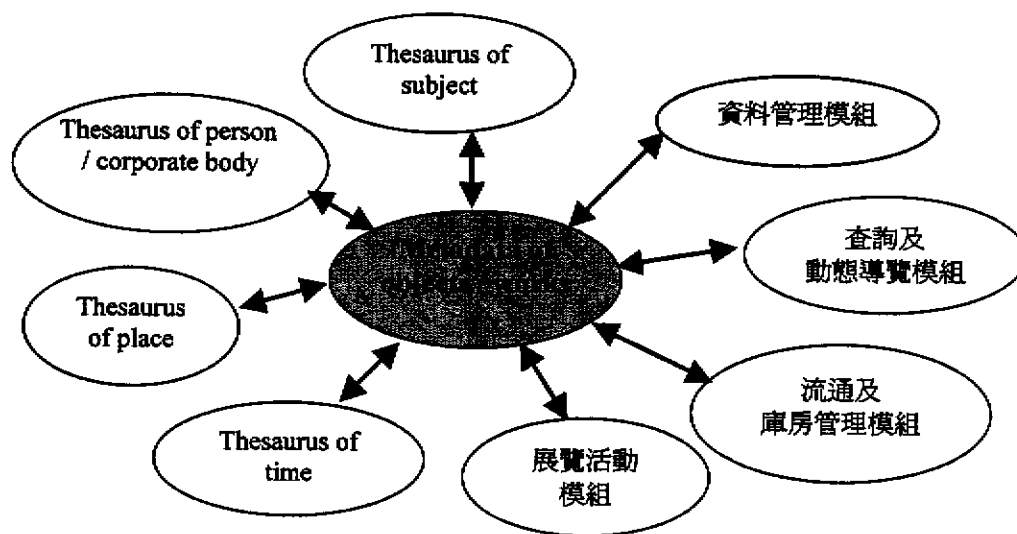
(二) 訪談使用單位之需求

此外，訪談 Metadata 的使用者或書畫專家是設計 Metadata 不可缺少的步驟。每一領域的使用者皆有該領域之使用習慣，若強求不同領域的人用相同的語言描述不同的資料，則此 Metadata 將得不到該領域的認同。在設計故宮的各種 Metadata 之前，了解他們的想法、需要就花了將近三個月的時間，Metadata 設計略具雛型之後，他們又提出了許多的修正意見，至少又用了三個月的時間，才將內容確定下來。

(三) Metadata 類型分析

除了藏品本身的描述之外，尚需將人物、主題、地點、時間方面的描述分開來，有如圖書館的名稱權威檔及主題權威檔一樣，應獨立建檔。這類型的 Metadata 不但可以

支援藏品資料的查詢，本身也是一個會被獨立查詢的資料庫，如故宮的人名權威檔，即類似一套人名小百科，本身就非常有價值。所以針對故宮的需求而言，初期我們所設計出來的 Metadata 類型有九種，包括：書畫、器物、文獻、主題、人名、地名、時代、展覽、參考書目等，而就一數位博物館而言，它應該包含的系統模組有：資料管理、查詢及動態導覽、流通及庫房管理、展覽活動等，其關係如下圖所示。



圖二：數位博物館 Metadata 及系統模組關係圖

(四) Metadata 語意設計

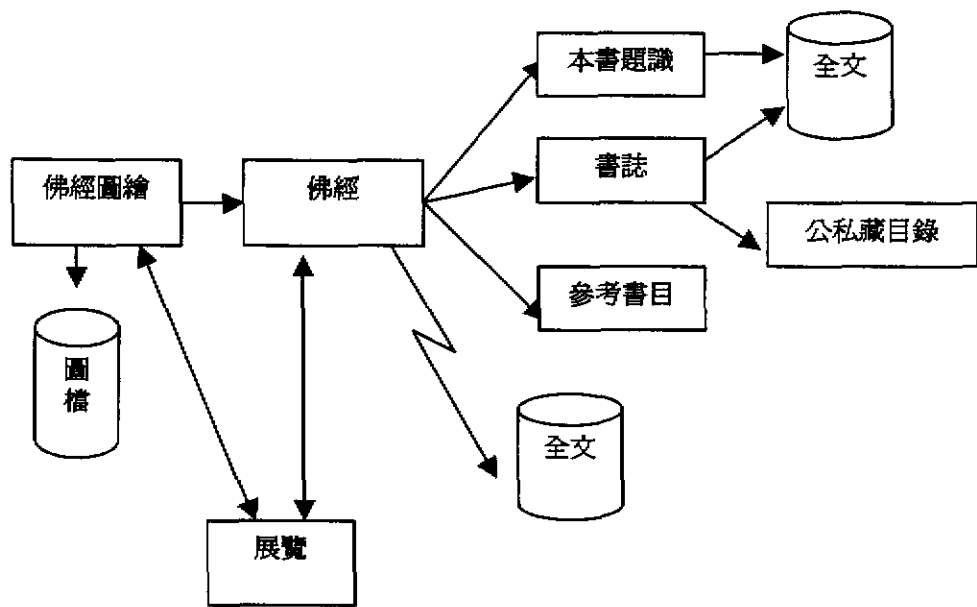
抽離了描述物件之共同屬性之後，進一步就應以較專業的用語或大家都使用的語詞來表示該屬性。如作品的名稱應稱為「題名」或「品名」或就用「藏品名稱」，作者名稱應用「作者」或「創作者」或「畫家」等，各領域所用的名詞，在系統的外部呈現可能不同，但內部的表達方式可以一致。屬性的名稱確定之後，還應確定描述的先後，這對不同領域的人，意義也不一樣。

(五) Metadata Model 設計

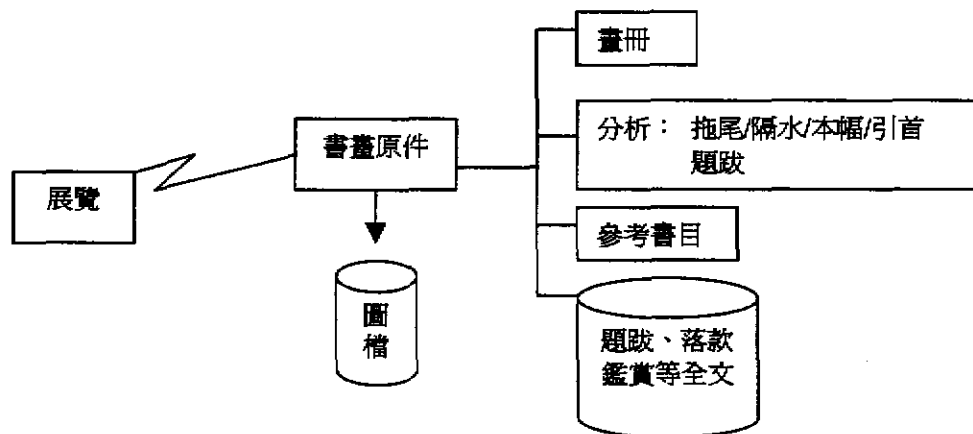
設計書畫之 Metadata 和寫一篇書畫鑑賞文章最大的不同，在於一篇書畫鑑賞文的作者會將和該幅書畫相關的知識都放在這篇文章中，而 Metadata 的設計必須知道應該如何將這些知識切分開來，對系統及對使用而言最具彈性，就好像關連式資料庫的 table 一樣，應建立幾個 table，是系統人員最具功力的地方，也是系統好壞的重要關鍵。如

果系統人員對資料的使用、屬性及系統架構沒有整體的了解，則可能會被使用單位牽著鼻子走，而照著使用者的說法將所有的資料放在同一筆記錄中；因為對使用者而言，他的資料使用習慣是整體的，但對系統而言，卻需做適度的分割。

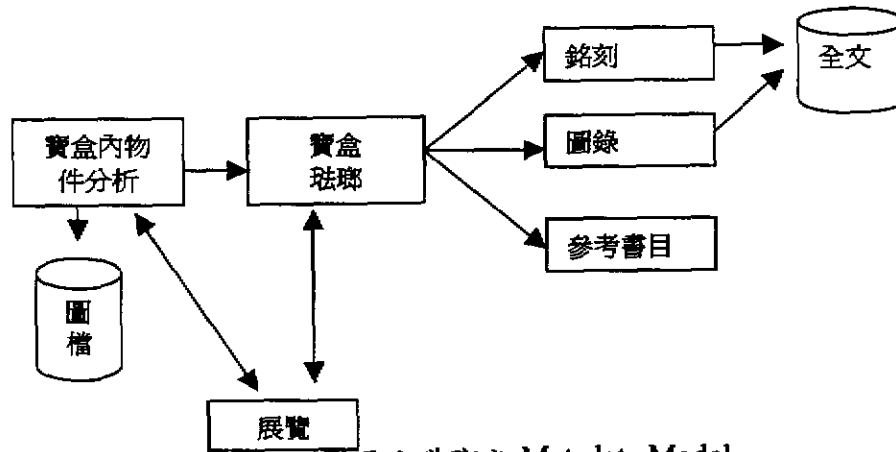
針對故宮的書畫、琺琅器、以及佛經圖繪三種今年要描述的藏品而言，要描述的主體當然是這三種資料本身，然而這三種資料本身還可以再分析，如一幅中國的畫有可分為：拖尾、隔水、本幅、引首、題跋，在某一次的展覽中，有可能只展覽一幅畫的引首或隔水，所以必需將這些資料分析出來；而一幅佛經圖繪一定是在某一套佛經中，所以佛經圖繪要描述，佛經也要描述，到底這些相關的描述應該放在同一筆記錄中，或分別在不同的記錄中描述，全視使用單位對各部份描述的詳細程度而定。換言之，「作品」與「作品」一定會有縱向的關係，也會有橫向的關係，這種關係如何表現，最困擾資料的描述者。為了說明資料的關係，筆者設計了這三種資料的 Metadata Model，並以圖例表示之。



圖三：佛經圖繪與佛經之 metadata model



圖四：書畫之 Metadata Model



圖五：琺瑯之 Metadata Model

(六) 交換語意及語法設計

如果資料庫不會與別人交換，則不必考慮交換語法的設計，只需將要描述的欄位設計出來即可，如果要與他人的系統相容，則需考慮交換的語意及語法。由於書畫等類型的 Metadata 在國內乃首次設計，所以在語意上們參考國外已設計出來的 Metadata，並做某種程度的調整。因為 Dublin Core 是目前較常被使用的 Metadata，所以在設計藏品本身的 Metadata 時，結構上，第一層即採用 Dublin Core 的 15 個欄位，第二層複分如果 Dublin Core 已有 qualifier 者，也盡量一致，不適用或不足之處則自己加上，這種做法，保留了第一層次的相容性。如果要轉成另一種 Metadata，如 MARC，則需有轉換表，不過可以在 XML 時代，這個部份可以交給 TXSL 來做。

在語法上，則一致採用 XML 標誌語言，做為交換的標準，所以每一種 Metadata 都有一種 DTD。

(七) 建立著錄範例，撰寫著錄手冊

Metadata 的欄位以及 Metadata Model 設計出來後，接著就應請使用單位找各種資料類型實際描述，看看是否還有需要修改之處，若已不須修改，即可以舉例的方式撰寫著錄手冊，讓每一個人皆有共識，照著手冊的指示著錄，以免前後原則不一。

(八) 系統實作設計

雖然語意及 XML 語法皆已設計完畢，但在系統實作時，尚需考慮到每一個欄位較

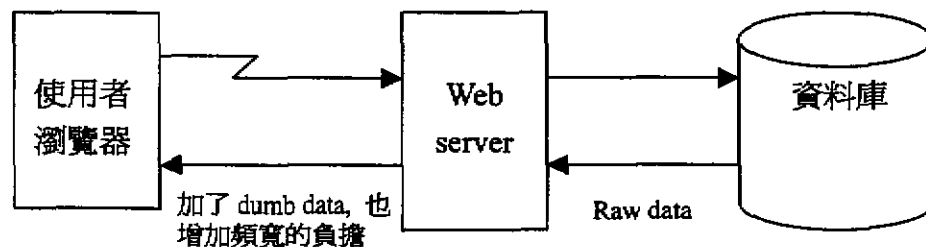
(八) 系統實作設計

雖然語意及 XML 語法皆已設計完畢，但在系統實作時，尚需考慮到每一個欄位較細部的屬性設定，如該欄位是否有預設值，是否以代碼表或權威檔控制輸入的資料，欄位長度為何，是否做檢索，檢索點的名稱為何，查詢顯示欄位及格式為何等。如「國立故宮博物院 MICI-DC & Metalogy 著錄範例與著錄手冊」所示，即為故宮書畫處的 Metadata 格式在我們所開發的 Metalogy 系統中的相關設定值。

四、以 XML 技術設計 Metadata 系統設計

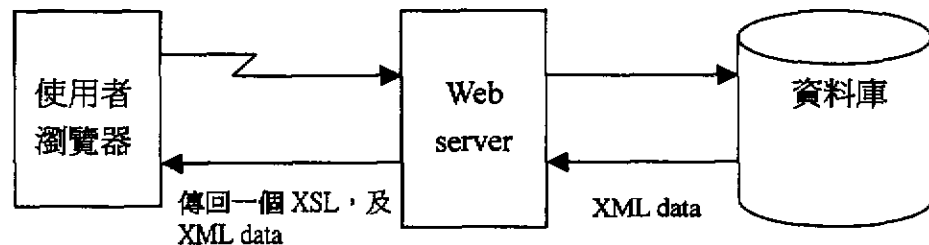
無論就數位圖書館相關計畫或電子商務等網際網路深入的應用而言，XML 已成網際網路界積極推廣的語法。XML 是 W3C 在 1998 年底提出的標準，它是從 SGML 衍生出來的簡化格式，也是一種元語言 (meta-language)，可以用來定義任何一種新的標示語言。XML 的制定是為了補足 HTML 的不完美，使得在 Web 上能夠傳輸、處理各類複雜的文件，它去除了 SGML 複雜不常用及不利於在 Web 傳送的選項功能，讓使用者可以很容易地定義屬於自己的文件型態，程式設計師也能在更短的時間開發 XML 相關應用程式。¹⁷

一般而言，目前在 Web 介面上做資料查尋都要透過 Web server 及 CGI 等程式將使用者的 query 轉成資料庫查詢語言，找到資料後，再透過這支程式將結果轉成 HTML 格式傳回使用者端。由於 HTML 語言主要是做顯示的標誌，所以使用者端無法做資料的再利用，此外，由於每一傳回使用者端的資料都加了顯示處理的訊息，使得原本的資料量變大而減慢了檔案的傳輸。如下圖所示：



圖六：目前的 HTMLweb 與資料庫查詢關係

然而如果使用 XML，則資料庫可傳回 XML 的資料到 Web server，Web server 只要傳一次 XSL(顯示樣式)，再將 XML 資料送回給使用者端，讓所有的顯示在使用者端處理即可，如此不但節省網路頻寬，使用者端也可以將有結構的 XML 資料做再利用，如下圖所示。



圖七：XML 之 web 平台與資料庫關係

由上述的說明我們可以得知以 XML 處理資料的潛力。綜合來說，XML 最適用於下列事件的資料處理：

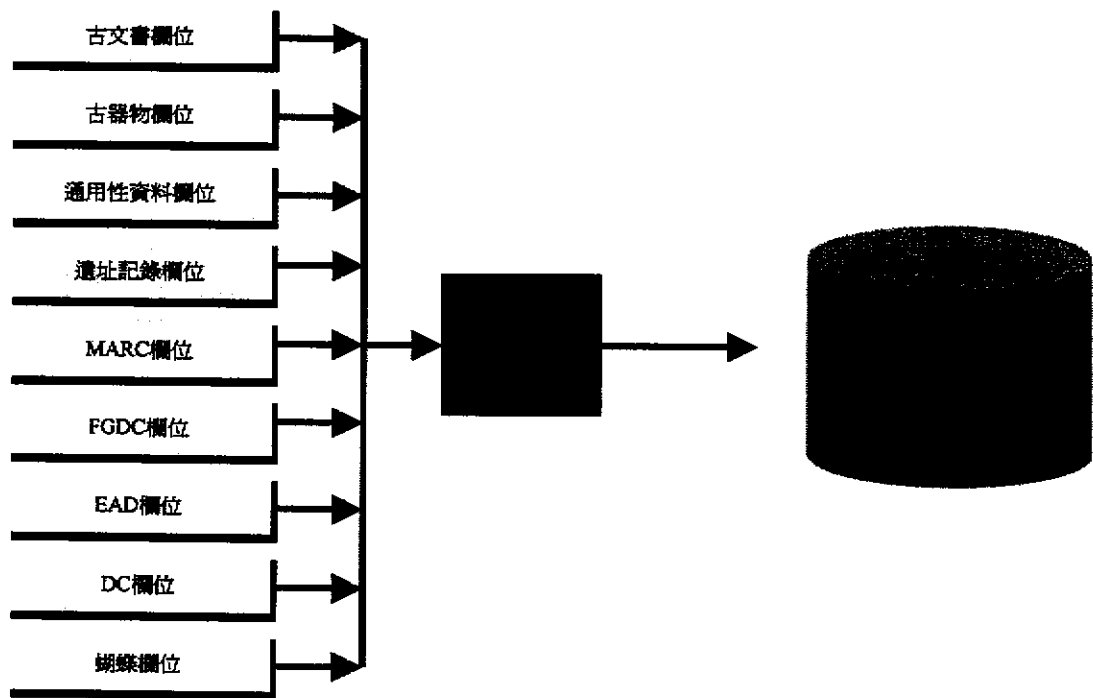
1. 資料庫之間的資料交換(Database interchange)
2. 使用者端的資訊處理(Client-side processing)
3. 使用者可以不同的樣式觀看資料(User views of the data)
4. 資訊過濾(Information filtering)

XML 的成功因素除了在於其架構取 SGML 之長，補 HTML 之短外，獲得大廠商的支持也是重要的原因。這些大廠商包括：Sun、Microsoft、IBM、Netscape、Oracle 等。因此目前實際採用 XML 的領域相當多，主要有：

1. 電子商務 (Electronic commerce)
2. 數位圖書館(Digital Libraries)
3. 以 XML 為基礎發展各類型資料之標誌語言，如：CML(Chemical markup Language), MathML, SVG(Scalable Vector Graphic), SMIL(Synchronized Multimedia Integration Language), RDF(Resource Description Framework), CDF(Channel Definition Format), OFX(Open Financial Exchange)
4. 線上銀行(Online banking)
5. 推播技術(Push technology)
6. 進階的檢索引擎(Advanced search engines)
7. 全球資訊網為基礎之控制系統(Web-based control systems)
8. 代理人系統(Agents)

XML 具有 SGML 的特性，但又不似 SGML 那麼複雜；XML 便於在網路上傳輸，又可以加強 HTML 所沒有的彈性及精確性。所以它已成為網路界及資料庫界積極支持的語法。因此，當我們要設計一個 Metadata 管理系統時，也決定採用這種語法，做為資料庫間交換資料的主要依據。

由於目前已發展出來的 Metadata 格式相當多，此外，為了能表達各類型各主題資料之屬性，尚未發展 Metadata 格式的領域可能還會發展出該領域的 metadata 格式，所以一個有彈性的 Metadata 管理系統，不能只是針對一種 metadata 格式來開發，而應該能讓使用單位自己決定要用那一種格式。所以，除了以傳輸及資料再處理的觀點來看 XML 的優點之外，從資料庫建置的角度而言，若能以 DTD 做為資料庫建置的基礎，也可以節省不同屬性的資料建置時間。因此，如下圖所示，任何 Metadata 只要寫出其 DTD 出來，就能建立一個新的資料檔，如此，對使用單位而言，建資料庫就不再是困難的事了。



圖八：metadata 以 DTD 作為資料庫建置基礎

伍、研究成果

一、中文詮釋資料交換格式

中文詮釋資料交換格式 (MICI) 乃以 Dublin Core 十五個基本欄位為主要的架構。但是為了描述文化遺產豐富的屬性、更精確表達描述藏品的語意，MICI 按照藏品類型特色，在相關欄位下，搭配欄位修飾語 (element qualifier) 的使用，不僅可以擴大其應用的範圍，同時也具有國際性。我們將以 DC 核心集為基礎並加上自行定義的欄位修飾語的詮釋資料格式稱為 MICI-DC。

(一) 都柏林核心集 Dublin Core

Dublin Core 起源於 1995 年由 OCLC (Online Computer Library Center) 與國家高速電腦應用中心 (National Center for Supercomputer Applications, NCSA) 於美國俄亥俄州的 Dublin 共同召開了一個詮釋資料研討會(OCLC/NCSA Metadata Workshop)，提出為網路資源提供詮釋資料的議題，以期支援跨領域的資源搜尋。都柏林核心集(Dublin Core，簡稱 DC)詮釋資料是一種專為網路上電子資源所設計的詮釋資料格式，原為網路資源著作者設計，目的為輔助跨領域網路資源的搜尋；但卻引起博物館界、圖書館界、政府機關與商業團體的廣大響應。因此 DC 工作小組邀聚各界學者專家，探討如何在 DC 核心集上，發展跨學界的國際共識。參與並推廣 DC 的國家遍及世界，共有 20 多個國家。DC 的優勢是簡單易用，若能推廣成為標準，具備跨領域的語意共通性後，就能進行跨領域的資訊交換。此外，DC 具有延伸性，能夠記錄架構、並描述較為複雜的語意。目前 DC (語意—資源的描述) 與 RDF (架構) 及 XML (語法) 結合，三者各自獨立、但彼此互補。

DC 是一種詮釋資料格式，基於國際的共識，定義了在網際網路的作業環境中，描述資源的基本資訊 (<http://purl.org/dc/>)。由於 DC 的目的是要讓資源的創造者或出版者自行建立資源的詮釋資料，格式簡單易用為其一大特點。都柏林核心集的格式雖然簡單，但為了符合使用者的不同需要，仍然維持了相當的延展性 (extensibility) 及可變性 (modifiability)。十五個欄位名稱如下：題名 (Title)、創作者 (Creator)、主題 (Subject)、簡述 (Description)、出版者(Publisher)、貢獻者 (Contributor)、日期 (Date)、類型 (Type)、資料格式 (Format)、辨識資料 (Identifier)、來源 (Source)、語文 (Language)、關連 (Relation)、時空涵蓋

範圍 (Coverage)、權限範圍 (Rights) 等。

簡易 DC (Simple Dublin Core or Unqualified Dublin Core) 是沒有用到修飾語的 DC 詮釋資料。換言之，欄位是以十五個 DC 詮釋資料欄位來表達屬性，並未進而定義著錄架構、修飾語值或是其他的處理資訊。DC 修飾語 (Qualified Dublin Core) 的詮釋資料則定義了著錄架構、控制詞彙、或是表示出詮釋資料的欄位值是 compound 或是 structured 的。DC 修飾語應用了其他資訊，以提高詮釋資料的精確度。例如，日期可以進而細分為最後修改日期、出版日期，並以特定的著錄方式著錄，以免產生模稜兩可的解釋。主題欄位可以從某個特定的控制詞彙選單 (如 US Library of Congress Subject Headings, LCSH) 來著錄。

(二) MICI-DC

目前 MICI-DC 可處理的資料類型包括：古文書、地圖、圖像/照片資料、器物以及故宮的書畫、器物與文獻 (佛經、佛經圖繪)。除採用 DC 制訂之修飾語外，也根據各典藏單位之需求自行定義相關修飾語，使用者可自行選擇 DC15 個欄位與修飾語，並視需求調整欄位順序，在遵循國際標準外，同時提供最大彈性以符合使用者的需求。詳細 MICI-DC 表如下：

表一：MICI-DC 欄位與修飾語列表

Element 欄位	Qualifier 修飾語		
Type 資源類型	Aggregation Level 藏品層次	單件/合集 Item/ Collection	
	Original Surrogate 原件與否	原件/代理物件 Original/ Surrogate	
	Cultural Natural 文化/自然	文化/自然 Cultural / Natural	
	DC Type 都柏林核心集建議 選項	互動式資源 interactive resource	
		資料組 dataset	
		事件 event	
		影像 image	
		聲音 sound	
		服務 service	
		軟體 software	
合集 collection (同「藏品層次」選項)			
文字 text			
	LocalLevel 自訂藏品類型		
Format 資料格式	Medium 媒體類型		
	Extent (Size, duration) 檔案大小或數量單位	Quantity 數量	
		Dimension 尺寸	Name 尺寸名稱
			MeasurementsUnit 尺寸數據和單位
	Position 尺寸位置		
Title 題名	Main 主要題名		
	Subtitle 副題名		

Element 單位	Qualifier 修飾語		
	Alternative 其他題名 (如翻譯題名、土語名)		
Description 簡述	Acquisition 入藏 (取得)	Method 方式	
		Source 來源	
Price 價格			
Physical Description 稽核項		Illustration 插圖	
		Color 色澤	
		Material 材質	
		Attachments 附件	
		FormWholeObject 全器形別或形式	
		Part Of Object 局部形式	
	Scale 比例尺		
	Abstract 內容摘要/ Synopsis 內容提要		
	Place 地點 (如製造地點等)		
Collection Or Site Information 採集或遺址資訊		Locality 採集或遺址位置	
		Name 採集或遺址名稱	
		Date Gathered 採集或發掘日期	
		Field Number 採集田野編號	
		Method of Collection 採集或發掘方式	
		Type Of Site 遺址性質	
		Coordinates 遺址座標	
		Coordinates of Object 遺址內物件座標	
		Phonomena 伴隨出土現象	
		Accompanying Object 伴隨出土物	
		Cultural Layer 文化層位	
		Geological Period 地質學時期	
		Age 年代測定	
Environmental Details 環境生態細節			
Seal Type 印記種類	From Artist 作者款印	Artist Inscription Seal Locality 作者款印位置	
		Artist Inscription 作者款識	
		Artist Seal 作者印記	
	About Colophon 題跋	Colophon Locality 題跋位置	
		Colophon Writer 題跋人	
		Colophon Seal 題跋印記	
		Colophon Full Text 題跋全文	
	About Label 題籤	Label Locality 題籤位置	
		Label Writer 題籤人	
		Label Seal 題籤印記	
		Label Full Text 題籤全文	
	About Loose Leaf 夾片	Loose Leaf Writer 夾片題跋人	
		Loose Leaf Seal 夾片題跋印記	
		Loose Leaf Full Text 夾片題跋全文	
	From Collector 收傳印記	Collector Seal Locality 收傳印記位置	
		Collector Seal Inscription 收傳印記印文	
	Inscription 銘刻 (包括內容全文)		Series Number 銘刻序號
			Position 銘刻位置
Category 銘刻種類			
Style 書寫方式			

Element 類位	Qualifier 修飾語			
		Content 內容	InscriptionContent Full Text 銘刻內容全文	
			InscriptionContent Image 銘刻內容影像	
	Decoration 紋飾	Series Number 紋飾序號		
		Position 紋飾位置		
		Category 紋飾種類		
	Transcription 釋文			
	Mount 裝潢	Volume Cover 冊面		
		Protective Covering Case 函套		
		Book Case 書盒		
	Release 版刻	Edition Name 版本名稱		
		Binding 裝訂		
		Style Form 版式	Border/ Column 欄式	
			Center Boundary/ Row 中央界行	
			Block Heart 版心	
			Frame Mark 版匡標記	
		Lines Per Page 行款		
	Font 字體			
	Exhibition 展覽	Exhibition Name 展覽名稱		
		Exhibition Size 展覽尺寸		
		Object Description 展品描述	Exhibition Description 展覽說明	
			Recommendation 鑑賞文章	
			Web Description 網頁說明	
	Condition 保存現況			
Grade 評等				
Notes 附註項				
Subject 主題和關鍵 詞	Subject Descriptor 控制詞彙	Primary Subject 主要主題		
		Secondary Subject 次要主題		
		Other Subject 其他主題		
		Situation 使用場合		
		Function 功能		
		Technique 技術	Series Number 技術序號	
			Position 技術位置	
			Category 技術種類	
		Style And Movement 風格派別		
		Personal Name 姓名		
		Corporate Body 團體/族群		
Keywords 關鍵詞彙 (非控制詞彙)				
Creator 創作者	Personal Name 姓名			
	Dynasty 朝代			
	Birth Place 籍貫			
	Corporate Body 團體/族群			
	Role 角色			
Contributor 其他貢獻者	Personal Name 姓名			
	Dynasty 朝代			
	Birth Place 籍貫			
	Corporate Body 團體/族群			

Element 欄位	Qualifier 修飾語	
	Role 角色	
Publisher 出版者		
Date 日期時間	Cataloging Date 編目日期	
	Created 製作時間	
	Issued 出版時間	
	Acquired 取得時間(入藏時間)	
	Modified 最後修改日期	
Identifier 識別資料	CallNumber 排架號	
	AccessionNumber 登錄號	
	URI 一致性資源識別號	
Source 來源		
Relation 關連	References 相關參照	Reference Work 參考作品
		Collection Catalogue 收藏著錄
		Research Material 參考資料
	Has Part 包含	Title 題名
		Creator 主要作者
		Contributor 次要作者
		Pagination 起始面頁碼
	Is Part Of 包含於	
Citation 引用出處		
Language 語文	Cataloging Language 編目語文	
	Item Language 作品語文	
Coverage 時空涵蓋範圍	Place 空間類別	Place Of Use 使用地點
		Scope Of Coverage 涵蓋範圍等級(地圖)
		Place Of Event 事件地名
	Time 時間類別	Period Of Use 使用時間
		Date Of Event 事件時間
	Lat Long 經緯度	
Rights 權限範圍	Owner Name 收藏者(單位)-名稱	
	Owner Country 收藏者(單位)-國家	

(三) MICI-DC 著錄手冊

為讓使用者更易於著錄 MICI-DC，我們製作了 MICI-DC 著錄手冊，根據 15 個欄位及其修飾語做語意上的說明，並提供相關範例，讓一般使用者可以自行選擇彈性運用，不需假手他人。請參見附錄一。

二、 XML/Metadata 系統發展實例—Metalogy 系統介紹

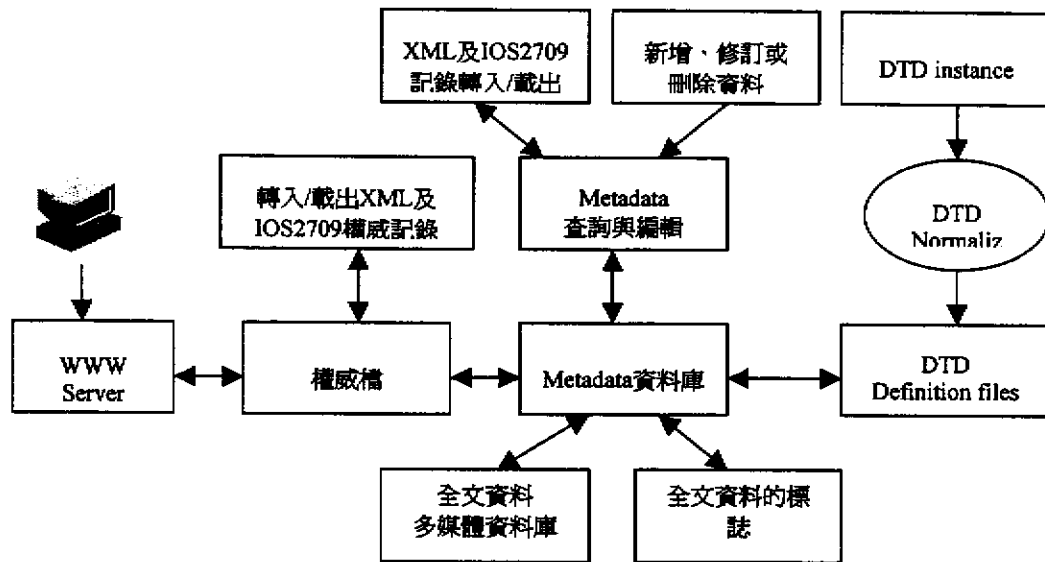
XML 具有 SGML 的特性，但又不似 SGML 那麼複雜；XML 便於在網路上傳輸，又可以提供 HTML 所沒有的彈性及精確性，所以它已成為網路界及資料庫界積極支持的語言。因此，當 ROSS 要設計一個 metadata 管理系統時，也決定採用這種語言，做為資料庫間交換資料的主要依據。但是除了語法外，如前所述，尚有語意的問題，由於目前已發展出來的 metadata 格式相當的多，同時，為了能表達各類型各主題資料之屬性，尚未發展 metadata 格式的領域可能還會發展出該領域的 metadata 格式，所以一個有彈性的 metadata 管理系統，不能只是針對一種 metadata 格式來開發，而應該能讓使用單位自己決定要用那一種格式。所以，開發一個 XML/Metadata 的通用性系統正好可以滿足這種需求，這也是 Metalogy 這個系統最主要的特色。詳細設計理念及架構說明如下：

(一) Metalogy (version 1.0) 系統特色與架構

Metalogy 是 ROSS 在國科會數位博物館專案下¹⁸，費時一年多所發展出來的 XML/Metadata 管理系統，這個系統可以用來建立數位博物館、數位圖書館、數位檔案館或任何主題資料之資料庫。它具有利用 DTD 建立資料庫、編輯 metadata、編輯權威檔(或索引典)、查詢(含 Window 介面及 Web 介面)、及轉入及轉出 XML 記錄等功能，此系統之特色主要有下列數端：

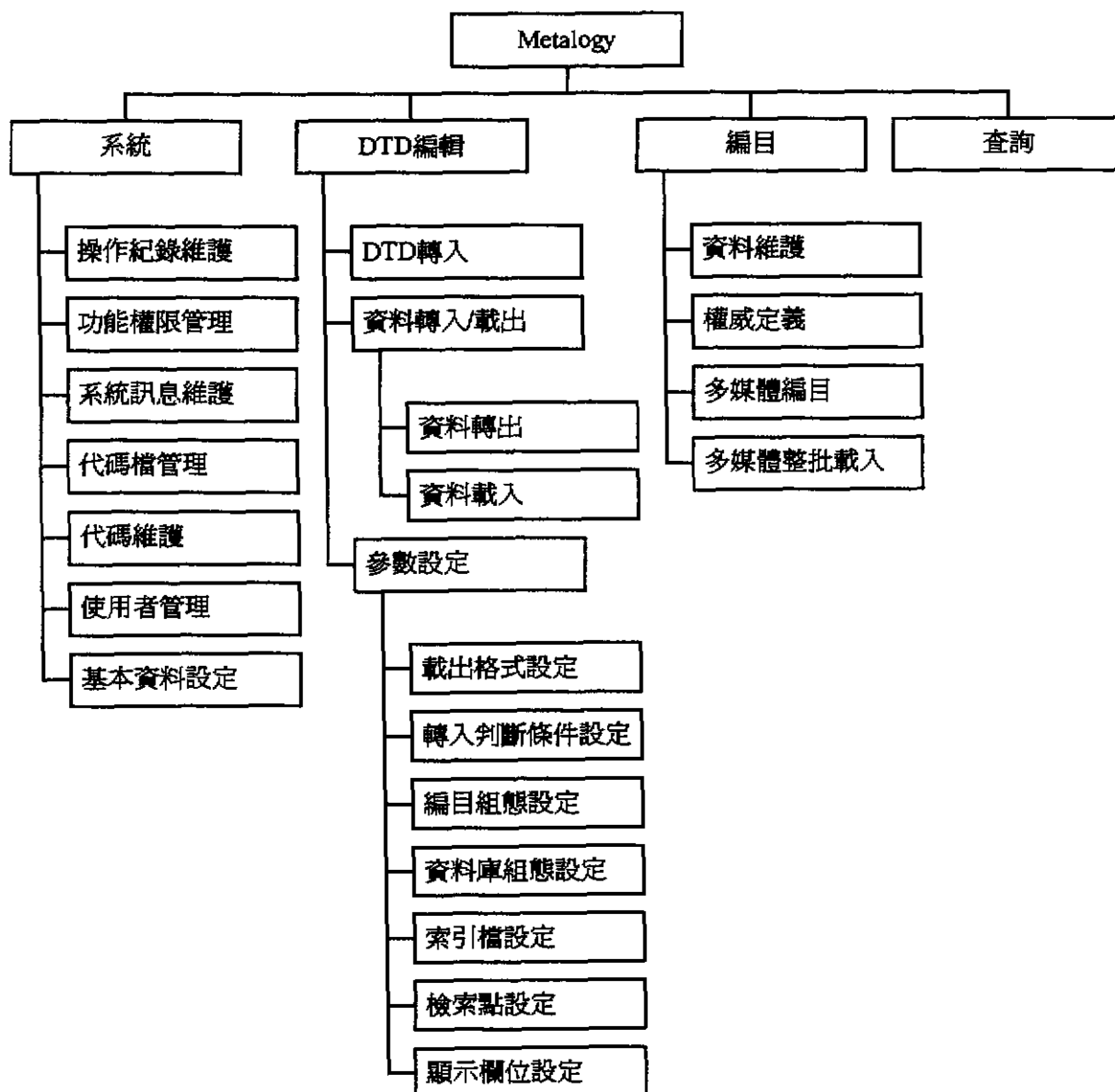
1. 系統 schema 主要依據輸入之 DTD 決定。
2. 系統允許使用多種不同之 DTD。
3. 多種不同格式之資料能夠同時被檢索。
4. 允許使用者依據 schema 調整內定之欄位格式與使用條件。
5. 具有親和的介面讓使用者自訂超連結、索引、檢索及顯示欄位。
6. 資料轉入與轉出能依循所屬之 DTD 格式。
7. 系統能判定轉入之資料是否符合指定之 DTD 格式，以及執行重複查核。
8. 系統能處理欄位式結構、多媒體結構、與全文。
9. 系統包含存取控制(access control)、異動記錄(transaction log)等系統管理功能。
10. 提供有 Web 查詢功能，可讓一般使用者透過 WWW 查詢資料庫。

Metalogy 系統之架構主要如下圖所示：



圖九：Metalogy 系統架構圖

目前此系統已大致開發完成，並有故宮博物院、宜蘭文化局、新竹文化局、台中市文化中心、高雄縣文化局等單位採用中。已設計出來的功能包括下圖各項：



圖十：Metalogy 功能架構圖

1. 系統：主要在供使用單位做系統管理之用，如權限設定、訊息維護等。
2. DTD編輯：包括的功能有將DTD轉入系統，相關參數設定、XML file的轉入及轉出。
3. 編目：即為各種Metadata記錄建檔及維護工作。
4. 查詢：此為 two-tier 的client端查詢介面，查詢功能包括精確檢索、模糊檢索、同音檢索，及簡單查詢及複合查詢兩種模式。

此外，Metalogy也另以Active X 開發Web 版的查詢介面。

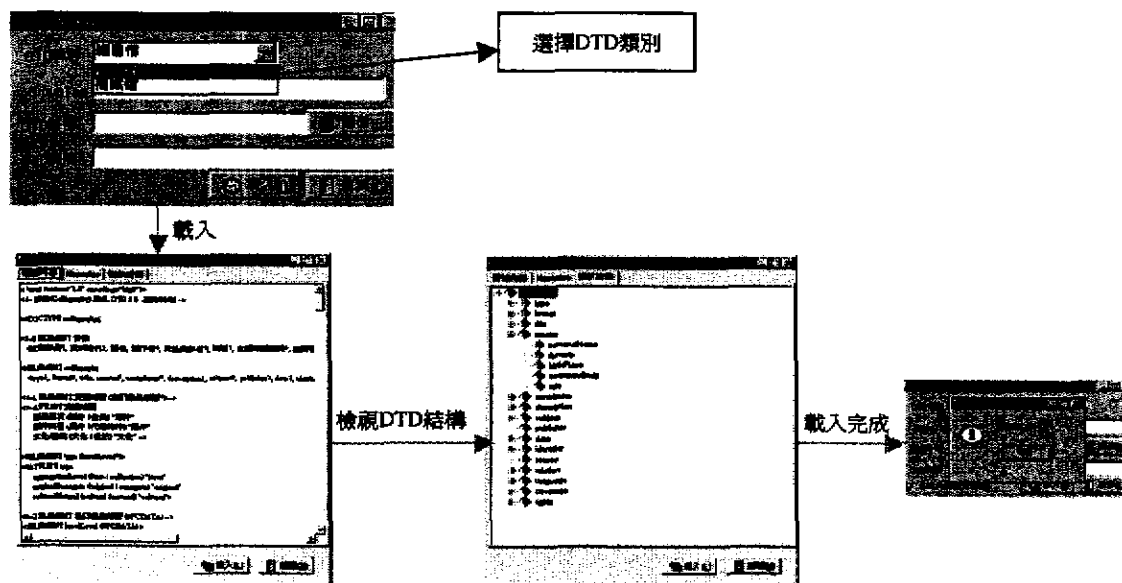
(二) 系統開發工具及目前已發展出來之功能

Metalogy 所使用之開發工具為 Delphi 5.0，Web Searching 所用之程式語言為 ASP，後端可用的資料庫管理系統為 Oracle 及 SQL server。

目前 Metalogy 已開發出來的功能包括：

1. 載入 DTD，建立資料庫

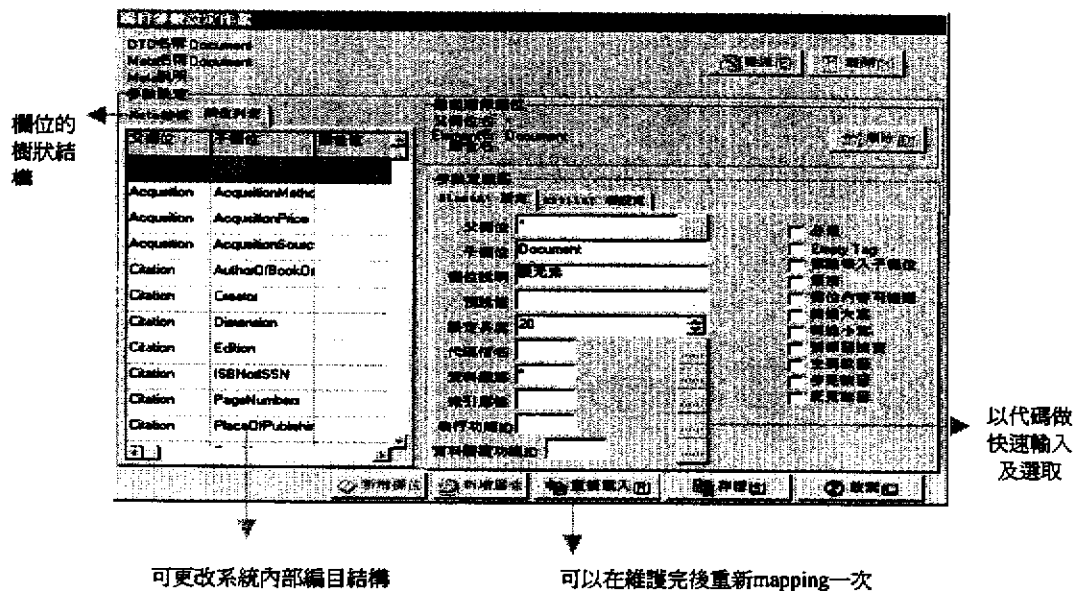
只需要載入任何一種 XML DTD，即會建立對應之資料庫，並產生編目畫面。



圖十一：DTD 載入畫面

2. 定義資料庫之 schema

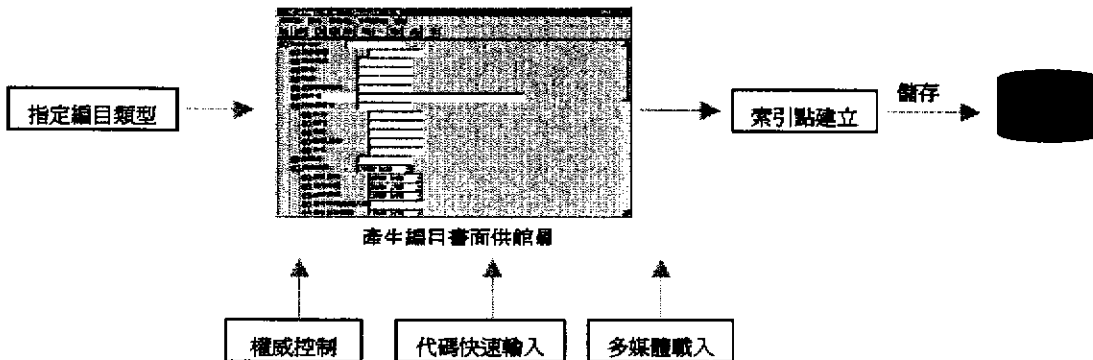
由於 DTD 在宣告上並不具備資料格式、特殊處理轉換、輸入長度、所屬權威款目、索引種類等欄位定義，雖然在轉入 DTD 時，系統會依照該 DTD 自動產生，但仍需要人工檢視、定義。



圖十二：定義資料庫 schema

3. metadata 編目功能

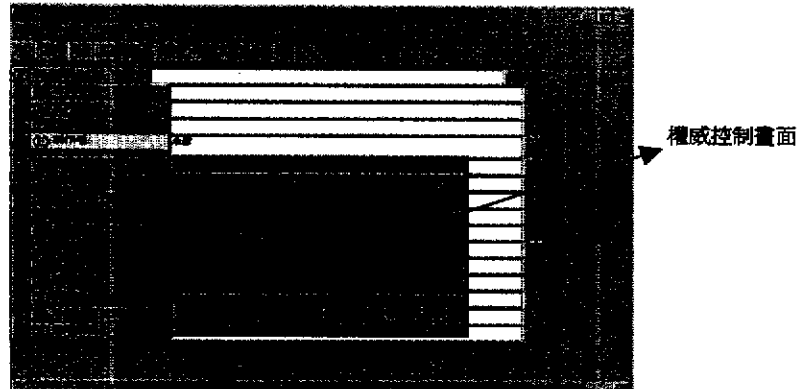
選擇所要編目的資料庫後，可以新增、修改或刪除某筆記錄。當編輯某一筆記錄時，能依據資料庫之 schema，針對某一欄位提供複製、刪除、插入子欄位、代碼、權威控制、與連結多媒體檔案等功能。並可直接呼叫查詢畫面，讓使用者能快速查檢所欲維護之記錄。



圖十三：編目畫面

4. 索引典建立及權威控制功能

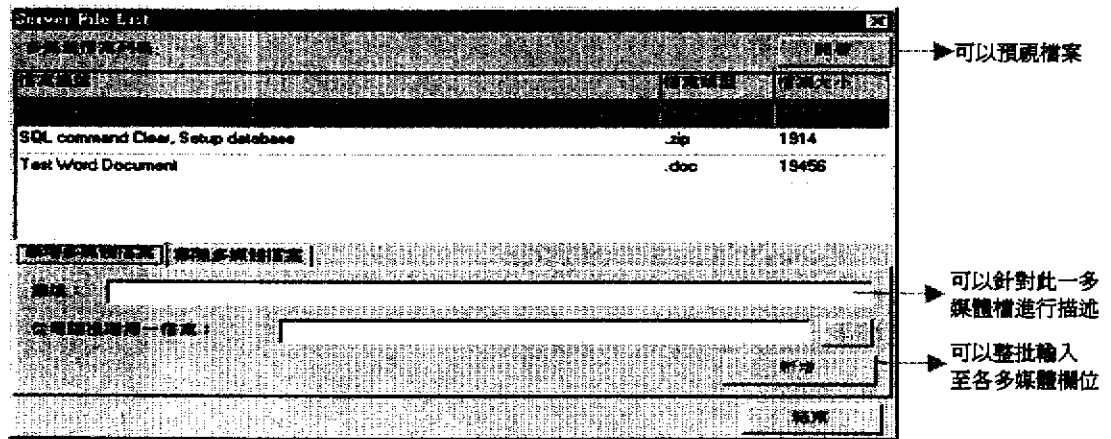
建立索引典或權威檔之流程完全同一般編目資料，系統會自動依據資料庫之 schema 執行權威控制。



圖十四：於一般編目畫面進行權威控制

5. 數位化資料的管理與描述

可對單筆多媒體檔案進行簡略編目，若有大批多媒體檔案需同時匯入，也可以採取批次匯入的功能，以節省人工作業。



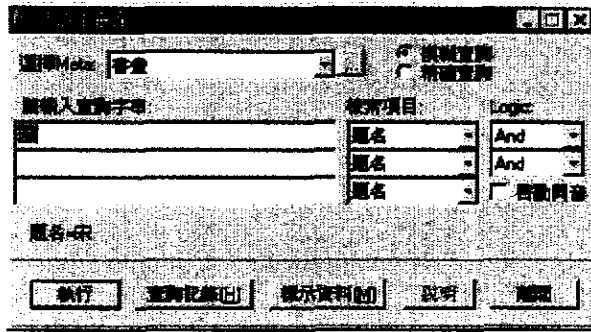
圖十五：多媒體檔案編目

6. 一般查詢功能

可以選擇一個（含）以上或所有的資料庫執行，提供精確或模糊查詢、同音、及布林邏輯運算等查詢功能。

7. 權威款目參照查詢功能

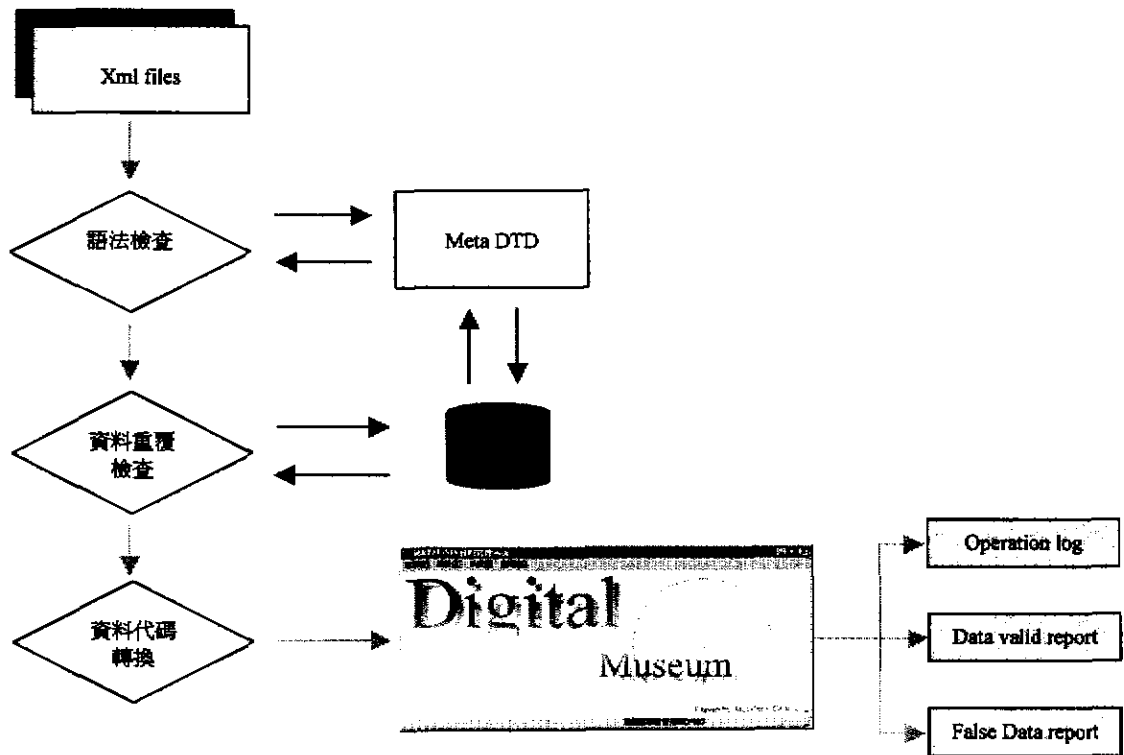
同一般查詢功能。



圖十六：查詢畫面

8. 轉入 XML 記錄

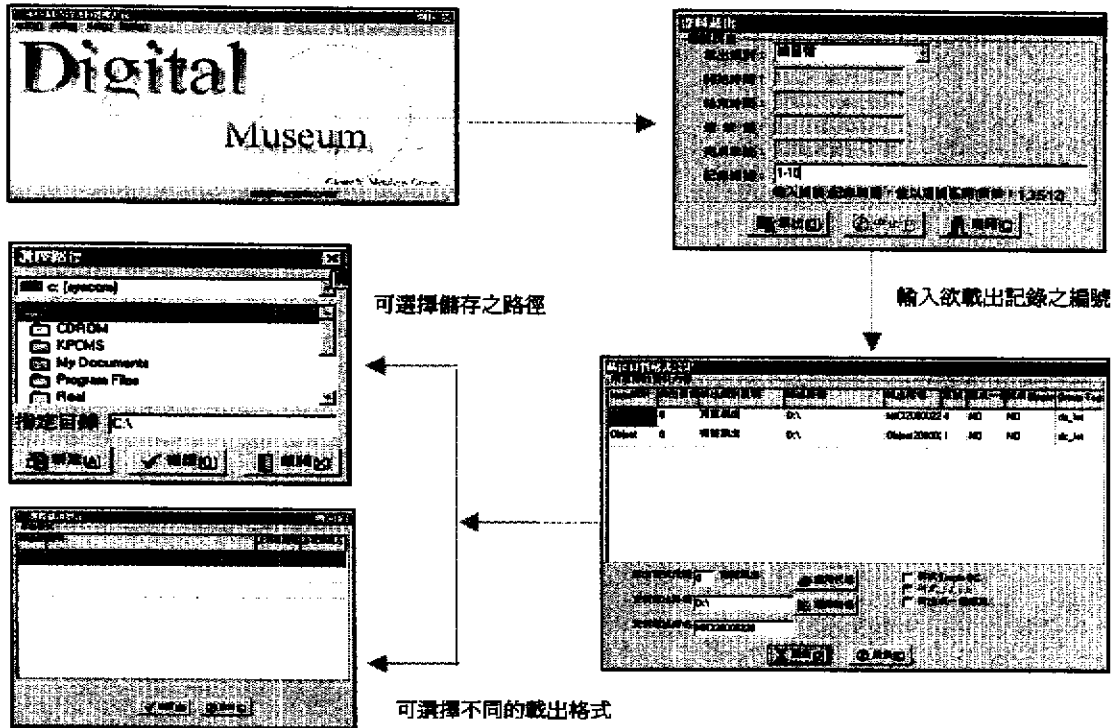
Metalogy 透過 XML 和外界做資料交換，只要系統存有該 XML DTD，即可接受該 DTD 之 XML 檔，檔案內可存一筆或一筆以上的紀錄，但不接受一個 XML 檔內有不同 DTD 格式之記錄。為避免資料重複，也可事先設定轉入之判斷條件，針對一個（含）以上的欄位，進行查核的動作。



圖十七：資料轉入流程

9. 載出 XML 記錄

可以選擇系統流水號或依指定的日期來載出 well-formed 的 XML 檔案，以供外界存取。除了可以在載出時設定相關參數外，也提供事先設定欲載出或不載出欄位之功能。此外，透過查詢功能來達到此一目的。



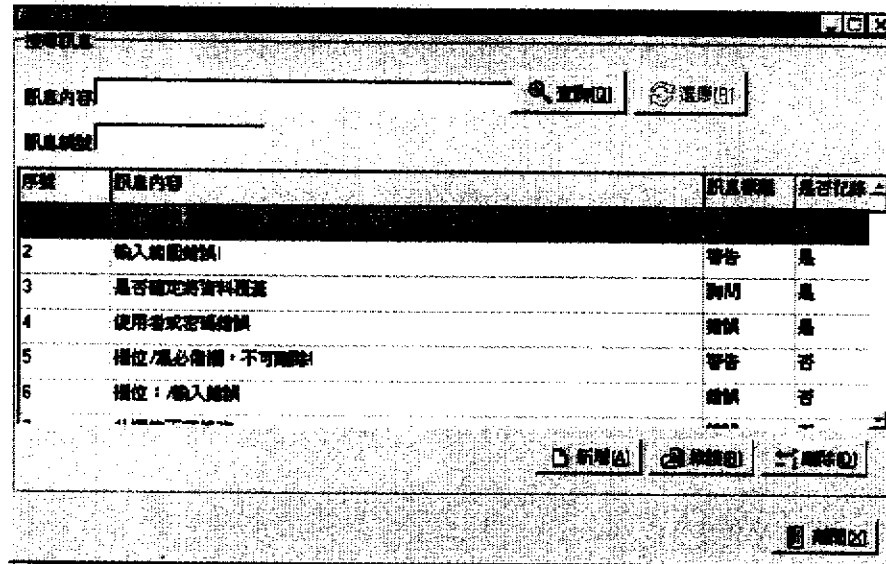
圖十八：資料載出格式

10. 使用者權限控制

提供使用者管理功能，來設定使用者之基本資料，及其操作權限。登錄系統時需輸入使用者代碼與密碼，完成確認時，系統會記錄此一使用者允許使用之權限。

11. 訊息管理

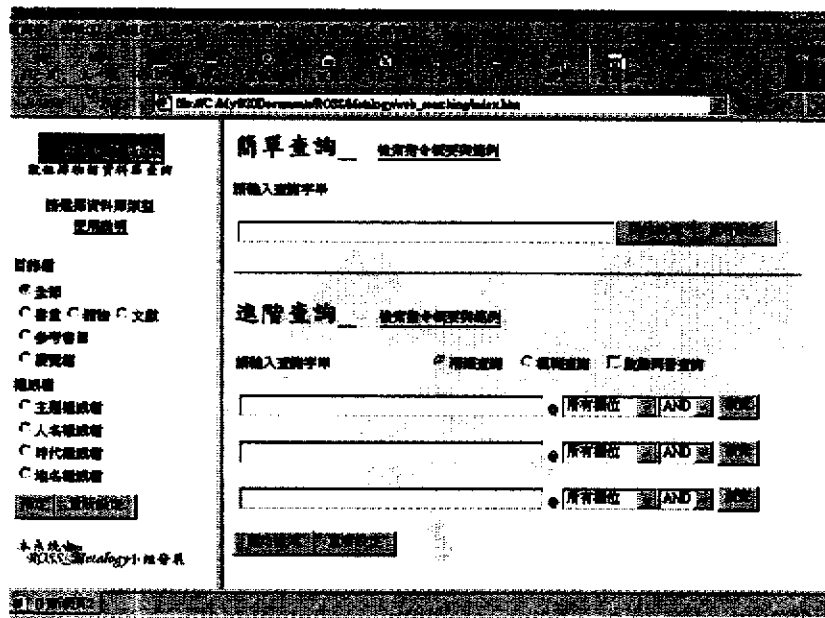
指定各錯誤訊息之說明內容、圖示、與按鈕種類。主要目的是提供系統在不須重新編譯的情況下，能夠隨時依使用者實際需要修改顯示的訊息內容，以避免部份訊息用語含糊或術語不夠專業，造成使用者誤解。



圖十九：訊息管理畫面

12. Web 查詢功能

提供與 Metalogy 系統內部相同之檢索功能。



圖二十：web 查詢畫面

(三) Metalogy 操作手冊

為了便於使用者利用本系統，我們也製作了 Metalogy 操作手冊，以圖示法詳細說明本系統的安裝、設定、使用及操作方法，相信透過本手冊，使用單位可以不假他人的幫助，輕易的就能使用自如。請參見附錄二。

(四) Metadata 及 DTD 範例

設計數位博物館/圖書館/檔案館之際，一定會針對要處理的資料設計 Metadata，而要使用 Metalogy，也一定要將 Metadata 以 XML/DTD 來表示。由於 Metadata 及 DTD 的訂定需要對使用者之資訊需求及所處理的資訊屬性深入了解，也需考慮系統的互通性，相當費時，最好能使用既有之標準。我們今年為故宮博物院的數位博物館計畫設計了幾份 Metadata 及 DTD，包括：書畫 metadata 及 DTD、器物 metadata 及 DTD、文獻 metadata 及 DTD、展覽 metadata 及 DTD、參考書目之 metadata 及 DTD、人名權威檔之 metadata 及 DTD、主題權威檔之 metadata 及 DTD、地名索引典之 metadata 及 DTD、時代索引典之 metadata 及 DTD，同時也有著錄手冊及著錄範例。

陸、結論

以XML做為Metadata的交換語法，並以XML設計數位圖書館系統，無論對各類資料的使用或系統功能擴充都非常具有彈性。目前Microsoft 的 IE 或是 Netscape 都已具有XML及XSL的處理能力，因此它在Web的普及應該是可以預期的。如果每一個資料庫系統都具備處理XML的能力，每一個網頁上放的都是XML的資料，則人類對於網路的使用將會更深入，網路蒐尋引擎的力量也會更強大。當然這一天是否真的會來臨，可能還是要看人類是否願意更無私的分享資訊而定了。茲將本計畫之貢獻分述如下：

一、對於學術研究、國家發展及其他應用方面預期之貢獻：

1. 探討處理各種電子資源之原則及適用性。
2. 分析中文數位圖書館/博物館藏品資料的特色。
3. 注意各專業學會在資料處理上的發展。
4. 建立資源著錄格式規則與著錄格式的基礎。
5. 訂定中文數位圖書館/博物館權威控制系統之規格。
6. 訂定系統權威檔之呈現模式、組織格式、瀏覽模式。
7. 訂定中文數位圖書館/博物館各子系統之互動模式。
8. 了解 XML、RDF 系統發展相關技術。
9. 訂定 XML/metadata 輸入、管理及交換系統規格。
10. 訂定 XML/metadata 輸入、管理及交換系統呈現模式、組織格式、瀏覽模式。
11. 訂定動態瀏覽查詢需求及模式。
12. 確認 XML/metadata 輸入、管理及交換系統之細部功能。
13. 建置 XML/metadata 輸入、管理及交換系統系統。

二、在國家發展及其他應用方面之貢獻有：

1. 發展詮釋資料著錄規則，並針對每一欄位做定義的制訂、語意的解釋，並提供著錄的例子，以供台灣、大陸與美國三地藏品著錄單位使用。
2. 發展詮釋資料的輸入、輸出、交換與管理的軟體，提供各類型中文數位圖書館/博物館資源使用。

三、對於參與之工作人員，所獲之訓練：

1. 行政能力的培養，如：計畫協調與聯繫，行政事務處理，會議記錄，文書檔案、工作進度控制與聯繫，與國外相關單位的書信往來，以及本計

劃網站之建立與維護。

2. 研究能力的培養，如：國內外相關標準之資料蒐集整理與研究，數位化藏品特質之調查，訪談調查與資料建檔，建立中文資訊詮釋資料格式，並翻譯相關文獻。
3. 程式撰寫能力的培養，如：XML/Metadata 程式撰寫，權威系統與著錄系統之整合，XML 整合查尋程式撰寫，建立詮釋資料交換之共通架構，訂定 DTD，以及維護軟體等。

Metadata技術是數位圖書館系統的核心，而XML是目前metadata最普遍採用的語言。Metadata的格式相當多，如EAD、GILS、FGDC、MARC、CIMI、TEI、DC等，尚有很多格式是根據上述這些格式設計的；此外，一個機構的資料類型可能有很多種，也可能會採用多種metadata 格式，這都是數位圖書館與傳統圖書館不同之處。因此，設計metadata管理系統時，不能只針對一種格式，而應該以XML為核心，能兼容並蓄各種metadata格式是比較適當的，Metalogy就是在這樣的理念下發展出來的系統，目前Metalogy是免費的，歡迎各界索取使用，並回饋意見，做為修正的參考。

註釋

1. Renato Iannella (1998). Mostly Metadata: A Bit Smarter Technology. Retrieved date from Dec. 25, 1999 from the World Wide Web: <http://www.dstc.edu.au/RDU/reports/VALA1998/>
2. 陳雪華，圖書館與網路資源（台北市：文華，民 85 年），頁 206。
3. 吳政數，「從電子檔案和元資料看未來資料著錄的發展趨勢」，在海峽兩岸圖書館事業研討會論文集，民 86 年 5 月 25-28 日，中國圖書館學會編，163-174，台北市：編者，民 86。
4. 陳昭珍，「電子圖書館資訊組織問題之探討」，在海峽兩岸圖書館事業研討會論文集，民 86 年 5 月 25-28 日，中國圖書館學會編，175-196，台北市：編者，民 86。
5. Lorcan Dempsey (1996). ROADS to Desire. Retrieved date from Dec. 25, 1999 from the World Wide Web: <http://hosted.ukoln.ac.uk/mirrored/lis-journals/dlib/dlib/july96/07dempsey.html>
6. Renato Iannella and Andrew Waugh (1997). Metadata: Enabling the Internet. Retrieved date from Dec. 25, 1999 from the World Wide Web: <http://www.dstc.edu.au/RDU/reports/CAUSE97/>
7. Pamela Drew and Jerry Ying. A Metadata Architecture for Multi-System Interoperability. Retrieved date from Dec. 25, 1999 from the World Wide Web: <http://www.nml.org/resources/misc/metadata/proceedings/drew/metaarch.html>
8. Jeff Rothenberg. Metadata to Support Data Quality and Longevity. Retrieved date from Dec. 25, 1999 from the World Wide Web: http://www.nml.org/resources/misc/metadata/rothenberg_paper/ieee.data-quality.html
9. "Z39.50" refers to the International Standard, ISO 23950: "Information Retrieval (Z39.50): Application Service Definition and Protocol Specification", and to ANSI/NISO Z39.50.
10. 陳雪華、陳昭珍、陳光華，「數位圖書館／博物館中詮釋資料之理論與實作」，圖書館學刊 13 期（民國 87 年 12 月），頁 37-60。
11. 陳亞寧等，「數位圖書博物館的資訊基礎建設：中央研究院 Metadata 的規劃設計與實施」，圖書館與資訊科學 25 卷 2 期（民國 88 年 10 月），頁 72。
12. 中央研究院資訊所，「計畫簡介：數位典藏系統技術研發」，
<<http://mars.csie.ntu.edu.tw/~dlm/plan/1st/intro6.htm>>（民國 89 年 11 月 5 日）。
13. 同註 11。
14. 陳昭珍、陳雪華、陳光華。「XML/Metadata 管理系統 - Metalogy 之設計」，

中國圖書館學會會報，65 期 (民國 85 年 12 月)，頁 1-12。

15. 陳昭珍，「打造臺灣的回憶：地方文獻數位化之理念與系統架構」，新世紀·新理念－公共圖書館發展實務研討會論文集 (台北：文建會中部辦公室主辦，民國 89 年 5 月 1 日至 3 日)，頁 75-92。
16. 國家圖書館編印，中文詮釋資料(Metadata)格式彙編 (台北市：國家圖書館，民國 89 年)。
17. 梁中平、徐千惠，「取 SGML 之長，補 HTML 之短－新一代標示語言 XML」，
<<http://www.cals.org.tw/files/cals1-4.htm>> (民國 86 年 11 月)。
18. Metalogy 系統主要發展人員包括：陳昭珍、陳雪華、陳光華、余顯強、陳嵩榮、蔡禮暉、謝順宏、歐陽慧。

柒、參考資料

- Aalberg, T. "Uniform access to collections with visual information: main subject project at NTNU, Institute for Data Technology and Information Science". *Synopsis* 28:2(Apr 1997):75-80.
- Allen, R. S. "An overview of the Federal Geographic Data Committee National Spatial Data Infrastructure, National Geospatial Data Clearinghouse, and the Digital Geospatial Metadata Standard: what will it mean for tomorrow's libraries"?. *Geography and Map Division Bulletin* 180 (Jun 1995):2-28.
- Andersen, J. "Internet and registration: a commentary". *DF Revy* 20:3(Apr 1997):75-6.
- Andresen, L. "Metadata: new key concept in Internet circles". *Bibliotekspressen* 6(Mar 1997):152-3.
- Armstrong, C. J. "Metadata, PICS and quality". *Online and CD-ROM Review* 21:4(Aug 1997): 217-22.
- Ashrafi, N & Kuilboer, J-P. "The information repository: a tool for metadata management". *Journal of Database Management* 6:2(Spring 1995):3-12.
- Bearman, D. & Duff, W. "Grounding archival description in the functional requirements for evidence". *Archivaria* 41 (Spring 1997):275-303.
- Bearman, D. "CNI/OCLC Workshop on metadata for networked images". *Archives and Museum Informatics* 10:3(1996):270-6.
- Bearman, D. "Developments in metadata management frameworks". *Archives and Museum Informatics* 10:2(1996):185-8.
- Bearman, D. "How the information revolution might affect us professionally". *Canadian Journal of Information and Library Science* 22:1(Apr 1997):38-55.
- Bearman, D. "Possible contributions of the reference model of metadata required for evidence to a reference model of metadata required for image description". *Archives and Museum Informatics* 10:3(1996):295-302.

- Bide, M. & Green, B. "Electronic tables of contents (EToCs) for serials: first steps towards standards for metadata". *Serials* 8:3(Nov 1995):257-60.
- Brodie, M. "In the company of strangers: challenges and opportunities in metadata implementation". *LASIE* 28:4(Dec 1997):40-6.
- Buttenfield, B. P. "Geographic information systems and digital libraries: issues of size and scalability. Geographic information systems and libraries: patrons, maps and spatial information". Illinois University at Urbana-Champaign (1996):69-80.
- Campbell, D. G. & Cox J. P. "Cataloguing Internet resources". *Feliciter* 43:5(May 1997):60-3.
- Caplan, P. & Guenther, R. "Metadata for Internet resources: the Dublin Core Metadata Elements Set and its mapping to USMARC". *Cataloging & Classification Quarterly* 22: 3/4(1996):43-58.
- Chen, P-W. & Cha, e. S-K. "Knowledge-based multimedia information retrieval in hyperspacng". *Telematics and Informatics* 14:1(Feb 1997):27-50.
- Cheung, W. C & Hsu, C. "The model-assisted global query system for multiple databases in distributed enterprises". *ACM Transactions on Information Systems* 14:4(Oct 1996):421-70.
- Ciganik, M. "Metainformation as a tool for information system management". *Kniznice a Informacie* 28:11(1996):441-8.
- Clemson, P. A. "An inside approach to a networked document cataloging". *Journal of Internet Cataloging*. 1:2(1997):57-64.
- Cobb, D. & Olivero, A. "The Massachusetts Electronic Atlas: an interactive Web site for access to maps and geographic data for the Commonwealth of Massachusetts". *Journal of Academic Librarianship* 23:3 (May 1997):231-5.
- Cox, R. J. "More than diplomatic: functional requirements for evidence in recordkeeping". *Records Management Journal* 7:1(Apr 1997):31-57.

- Crook, M. "Research Advisory Committee meets". OCLC Newsletter 217(Sep/Oct 1995):13-14.
- Damcevski, P.& Chatwin, D. "Facts about the figures: issues in visibility and accessibility of metadata in numerical datasets". Cataloguing Australia 21:3/4(Sep/Dec 1995):182-4.
- Dempsey, L. "Metadata". LASER Link (Spring/Summer 1997):16-17.
- Dempsey, L. "Metadata". LASER Link (Spring/Summer 1997):16-17.
- Dempsey, L. "Metadata: a UK HE perspective". British Library. Research and Innovation Report 78(1997):133-5.
- Desai, B. C. "Supporting discovery in virtual libraries". Journal of the American Society for Information Science 48:3(Mar 1997):190-204.
- Domaratz, M. "Finding and accessing spatial data in the National Spatial Data Infrastructure. Geographic information systems and libraries: patrons, maps and spatial information." Illinois University at Urbana-Champaign(1996):31-40
- Duff Will,W. "Metadata replace archival description: a commentary". Archivaria 39 (Spring 1995):33-8.
- Duff, W. "Ensuring the preservation of reliable evidence: a research project funded by the NHPRC". Archivaria 42 (Fall 96):28-45.
- Feldman,S. "Advances in digital libraries '97: the Washington IEEE forum delves into the implications of technology". Information Today 14:7(Jul/Aug 1997):12-13,55.
- Ferguson, P. "ELDIS: mapping information on development and the environment". Information Development 12:3(Sep 1996):155-62.
- Hakala, J.;Koch, T. & Husby, O. "Warwick framework and Dublin core set provide a comprehensive infrastructure for network resource description". Nordinfo Nytt 2(1996):40-8.

- Heery, R. "Review of metadata formats". Program 30:4(Oct 1996):345-73.
- Henze, V. & Schefczik, M. "Metadata. Connection between Dublin Core Set, Warwick Framework and data formats". Bibliotheksdienst 31:3(1997):413-419.
- Hurley, C. "Ambient functions: abandoned children to zoos". Archivaria 40 (Fall 1995):21-39.
- Jizba, L. "Reflections on summarizing and abstracting: implications for Internet Web documents, and standardized library cataloging databases". Journal of Internet Cataloging 1:2(1997):15-39.
- Kerr, M. "NewsAgent for Libraries". Vine 104 (1996):51-3.
- Kingwell, J. "Trends in interactive information systems for earth observation from space: Towards a global, digital image library service". Australian Special Libraries 29:2(Jun 1996):49-60.
- Lamb, P. C. "Seven products you can sell online". Digital Publishing Strategies 1: 12(Aug 1997):9.
- Lehmann, K-D. "European national libraries and the CoBRA Forum of the EU Libraries Programme". Alexandria 8:3(1996):155-66.
- Lehmann, K. D. "Difficulties with the levels. Cataloguing rules, data formats and communication interfaces". Zeitschrift für Bibliothekswesen und Bibliographie 44:3 (May/Jun 1997):229-40.
- Lester, R. "Internet hype and Web wonders". Assistant Librarian 89:7(Jul/Aug 1996):109-11.
- Lichtenberg, J. "Inching towards e-commerce". Publishers Weekly 244:50 (Dec 1997):35-7.
- Lunin, L. F. "ASIS, FID hold celebratory meeting in Chicago". Information Today 12 :11 (Dec 1995):25-6.
- Lynch, C. "A summary of the findings of CNL's NIDR research initiative. British

Library." Research and Development Department. BLRD Report 6250(1996):89-93.

Lynch, C. "Metadata: moving from planning to implementation". British Library. Research and Innovation Report 78(1997):137-9.

MacNeil, H. "Metadata strategies and archival descriptions: comparing apples to oranges". *Archivaria* 39 (Spring 1995):22-32.

Mangan, E. U. "The making of a standard". *Information Technology and Libraries* 14:2(Jun 1995):99-110.

Marsh, E. C. etc. "Looking ahead: the OCLC Office of Research". *OCLC Newsletter* 225 (Jan/Feb 1997):14-26.

Neville, L. "Internet resources: issues of access and possible solution". *Cataloguing Australia* 23:1/2(Mar/Jun 1997):23-7.

Nordinfo, S. v. & Nytt. "Ungern-Sternberg Information day on metadata in Lund". *Informationsdag om metadata i Lund* 3 (1996):13-16.

Paskin, N. "Information identifiers". *Learned Publishing* 10:2 (Apr 1997):135-56.

Perkins, M. "Why don't search engines work better?." *Information Outlook* 1:5(May 1997):40-1.

Phillips, J. T. "Metadata - information about electronic records". *Records Management Quarterly* 29:4(Oct 1995):52-5.

Pountain, D. "Of teams and components". *Byte* 22:5(May 1997):40IS15-16, 40IS18, 40IS20,40IS22.

Queau, P. "Meta-data". *IFLA Journal* 23:1(1997):20-2.

Rand, R. Y. etc. "Libraries, global change data, and information management".
Library
Hi Tech 13:1-2 (1995):26-42.

- Rochester, M. "Report on the 63rd IFLA conference, Copenhagen, August 31-September 5, 1997". *Education for Library and Information Services: Australia* 14:3(Nov 1997):62-4.
- Rogers, D. "Cataloguing Internet resources: the evolution of the Dublin Core metadata set". *Cataloguing Australia* 23:1/2(Mar/Jun 1997):17-22.
- Rotter, W. MALVINE. "A digital project of the Austrian Literature Archive". *Mitteilungen der Vereinigung Osterreichischer Bibliothekarinnen und Bibliothekare* 50:2(1997):70-2.
- Rusch-Feja, D. "Subject-oriented collection of information resources from the Internet". *Libri* 47:1(Apr 1997):1-24.
- Schram, S. G. & Wallace, W. W. "USDA Global Environmental Change Data Assessment and Integration Project". *Journal of Agricultural & Food Information* 3:2(1995):39-47.
- Shoklar, K.; Basu, C. & Shah, K. "Putting legacy data on the Web: a repository definition language". *Computer Networks and ISDN Systems* 27:6(Apr 1995):939-51.
- Simth, J. T. "Meta-analysis: the librarian as a member of an interdisciplinary research team". *Library Trends* 45:2(Fall 1996):265-79.
- Smits, J. "Digital metadata, standards for communication and preservation". *European Research Libraries Cooperation*. 6:4(1996):383-406.
- Sorensen, A. "Registration in the future: new demands, new media; keywords from the DF conference on 4 and 5 February 1997". *DF Revy* 20:3(Apr 1997):71-4.
- Sprehe, J. T. "The U.S. Census Bureau's data access and dissemination system (DADS)". *Government Information Quarterly* 14:1(1997):91-100.
- Steele, C. "Organization of knowledge on the Internet". *Information Management Report* (Sep 1996):1-6.
- Stratford, J. S. & Stratford, J. "Computerized and networked government information:

- California Environmental Resources Assessment System". *Journal of Government Information* 22:3(May/Jun 1995):267-72.
- Tammaro, A. M. "Cataloguing, cataloguing ... metacataloguing". [Catalogando, catalogando ...metacatalogando.] *Biblioteche Oggi* 15:1(Jan-Feb 1997):80-14.
- Tomimoto, H. "Systematization of network resources". *Igaku Toshokan (Journal of the Japan Medical Library Association)* 44:1(Mar 1997):54-65.
- Vidgen, R. "Temporal object modeling: diagramming conventions and design considerations". *Journal of Database Management* 8:1(Winter 1997):14-23.
- Vidgen, R. "Temporal object modelling: diagramming conventions and design considerations". *Journal of Database Management* 8:1(Winter 1997):14-23.
- Vizine-Goetz, D. "OCLC investigates using classification tools to organize Internet data". *OCLC Newsletter* 226 (Mar/Apr 1997):14-18.
- Wallace, D. A. "Managing the present: metadata as archival description". *Archivaria* 39(Spring 1995):11-21.
- Watson, B. C. "Research Advisory Committee meets, gets update on OCLC Office of Research projects". *OCLC Newsletter* 229(Sep/Oct 1997):16-18.
- Watson, B. C. "Research Advisory Committee meets". *OCLC Newsletter* 226 (Mar/Apr 1997):19-20.
- Watson, B. "Research Advisory Committee meets". *OCLC Newsletter* 223(Sep/Oct 1996): 18.
- Weibel, S. & Miller, E. "Cataloging syntax and public policy meet in PICS". *OCLC Newsletter* 227(May/Jun 1997):28-9.
- Werf-Davelaar, T. v. d. "The bibliographical description of electronic information resources. 2. *Informatie Professional* 1:7/8(1997):25-31.
- Wheatley, A. & Armstrong, C. J. "Metadata, recall, and abstracts: can abstracts ever be reliable indicators of document value?." *Aslib Proceedings* 49:8(Sep

1997):206-13.

White, M. "Intranets aren't always the solution". *Digital Publishing Strategies* 2:3(Nov 1997):6.

Wu, C-J. "Comparison of three metadata standards". *Bulletin of the Library Association of China* 57(Dec 1996):35-45.

Younger, J. A. "Resources description in the digital age". *Library Trends* 45:3(Winter 1997):462-87.

中國美術全集編輯委員會編。中國美術全集古代部分。台北市：錦繡，1993-1994。

世田谷美術館、日本放送協會、NHK 編輯。秦始皇帝時代展。東京都：日本放送協會、NHK，1994。

李學勤。東周與秦代文明。板橋市：駱駝，民 72。

陝西省博物館藏寶錄編輯委員會編。陝西省博物館藏寶錄。上海：上海文藝，1995。

蓋瑞忠編著。秦漢工藝史。台北市：台灣省立博物館，民 78。

韓復智等編著。秦漢史。台北縣蘆洲鄉：空大，民 85。

中國歷史 <http://sokam.flysaturn.com/C-Hist/ChinHist.htm>

秦俑考古隊 <http://www.bingmayong.org/>

獅子吼佛教專站 (<telnet://cbs.ntu.edu.tw>) 精華區

<http://cbs.ntu.edu.tw/sci/history/CHINESE/archaeology/chingling/>

行政院國家科學委員會補助國內專家學者出席學術會議報告

報告人姓名	陳雪華	服務機關及職稱	國立臺灣大學圖書資訊學系 教授兼系主任
會議時間地點	2001/05/29-2001/05/31 北京清華大學	本會核定補助文號	NSC 89-2750-P-002-018-Y
會議名稱	(中文) 第十二屆新信息技術國際會議 (英文) The 12 th International Conference on New Information Technology		

報告內容目次

壹、參加會議經過	03
一、背景介紹	03
二、會議目的	05
三、會議主題	06
四、會議簡要議程	07
五、參與人員	10
六、與會心得與建議	12

貳、附件

- 一、NIT-2001 PROGRAM
- 二、Information about NIT-2001 Logistics
- 三、NIT-2001 Participant's List

壹、參加會議經過

一、背景介紹

過去幾年，我們已感受到新資訊技術如何滲透進我們的生活、學習、工作及娛樂。計算技術和通訊技術的飛速發展已為教育、科研、商業及社會生活創造了新的基礎結構。其結果，圖書館和資訊中心能把它們為用戶提供的服務，特別是在全球資訊的存取、檢索和傳遞的能力上，提高到前所未有的水平。

除了在通訊網絡、電子出版、互動式多媒體技術、集成資訊系統及數位圖書館等方面的快速發展外，Internet 和 WWW 的爆炸式的發展和廣泛應用，已使圖書館員和資訊情報專家能夠提供前無先例的資訊服務。在數位圖書館領域，出現了難以置信的機遇，有可能實現對分散式的資訊系統進行全球資訊存取和資源分享。另一方面，網路空間正在成為主導媒體，使人們能創造、共用資訊和思想，而無需經過常規的資訊傳送渠道，這對圖書館和資訊中心已構成巨大挑戰。

在這新千年到來之際，下一代 Internet 將使數位通訊能力有成百上千倍地改善，更快、更好。因此，我們能夠看到的由世界各地開發出的分散式多媒體資訊的數位圖書館或資訊中心，一定會更多、更大。然而，為了有效地創建這些廣泛分佈的系統，我們必須懂得在創建時如何使它們能夠容易地互聯、檢索和相容（有可互操作性）。此外，我們還需知道在這易變的、不可控的、神經網路式的市場環境下如何保護知識產權。這樣一些高層次研究的問題，只能由來自眾多相關學科領域——圖書館、資訊科學、電腦科學、多媒體、通訊、網路、語言學、法律、社會科學等——的國際性專家們來攻克。NIT 2001 將對這些問題有更多些的關心。

從 1986 年起，NIT 已先後於曼谷、新加坡、Guadalajara、布達佩斯、香港、Puerto Rico、亞歷山大（美國佛吉尼亞）、里加（拉托維亞）、Pretoria（南非）、河內（越南）及臺北等地成功召開 11 次會議。這一系列會議，對於處在這個既激動人心又存在迷惘時期的圖書館及資訊工作者，在許多關鍵問題和爭論上取得全球共識，起到了促進作

用。值此進入新千年之際，在北京的清華大學召開新一次會議，它將同時為發達國家和發展中國家的與會者，提供一次正式和非正式討論的極好機會。相關學科領域的學者藉此可相互鼓勵和切磋。

在保持 NIT 既有目標和任務的同時，這屆會議還有一個特別之處，即會議在北京清華大學舉行，作為祝賀該校 90 周年校慶活動的一部分。期待著來自不同學科領域為討論數位圖書館而走到一起的各國(地區)的學者，審視在這新千年所面臨的巨大挑戰，因為這次會議一定會對分佈於世界各地的數位系統之間進行富有成果的知識交流和資訊交流作出貢獻，同時使那些低效、重復、支離破碎的數位系統不再發展。

二、會議目的

現代電腦技術、通訊技術和網路技術的發展，給圖書館和資訊服務事業的發展帶來了勃勃生機。20世紀90年代以來，世界各國掀起了一場數位圖書館研究與建設的熱潮，這一趨勢正在從根本上改變著人們獲取和利用資訊的途徑與方式，對科學研究、人類教育、經濟和社會發展產生巨大影響，推動著資訊革命的發展。數位圖書館是評價一個國家資訊基礎設施水平的重要標誌和本世紀高科技競爭的焦點之一。

當今世界科技進步真可謂日新月異，從某種意義上說，以Internet為代表的資訊技術革命已經把整個地球連接成爲一個有機的整體，爲人類展現了跨國界、跨社會、跨文化、跨語言的資訊、知識交流的無限可能性。數位圖書館是資訊高速公路的重要資訊資源，是未來社會的公共資訊的中心和樞紐，必將成爲21世紀知識經濟時代一種新的知識創新、發現和傳播的重要軟基礎設施，數位圖書館的建設已經成爲當前評價一個國家資訊基礎水平的重要標誌。正如各位所知，資訊素養這個概念是從圖書館檢索技能演變而來，隨著資訊技術的廣泛應用和由此帶來的人們生存方式、工作方式和學習方式的改變，資訊素養的重要性，日益凸現出來，其內涵和外沿進一步豐富和擴大。目前，Internet上的數位化資訊資源正在以無法估量的速度增長。面對分佈如此廣泛、載體類型如此衆多、數量如此龐大的資訊，如何加以有效的組織和管理，提供在水平和效益上遠遠超過傳統圖書館的高質量的、多樣化的、統一的全球化資訊服務，這正是亟待解決的問題，也正是舉辦這次高水準數位圖書館國際學術會議的宗旨所在。

三、會議主題

1. 國家資訊基礎結構與政策
2. 本世紀與下一世紀網路的應用
3. 數位化圖書館與系統：角色的改變、協作和探索
4. 合作研究活動
5. 可相互操作性和規模可改變性技術問題
6. 多形式，多語種問題和探討
7. 智慧代理
8. 元資料技術和工具
9. 數位化資訊的保存和存檔
10. 數位化圖書館的社會、經濟和法律問題
11. 資訊技術的教育與培訓
12. 數位化圖書館和聯盟發展
13. 遠端教育及分散式
14. 世界範圍分散式數位化圖書館的發展
15. 知識產權和版權問題
16. 安全問題

四、會議簡要議程

2001年5月29日(星期二)	
08:00 – 08:45 註冊	13:45 – 14:45 特邀報告 2
08:15 – 08:30 主持人會議	主持人：陳劉欽智，美國西蒙斯學院圖書館
08:50 – 10:10 開幕式	資訊研究所教授
致歡迎詞 清華大學校長王大中致詞	14:45 – 15:45 會場 1 網路發展和數位圖書館體系結構
貴賓介紹 胡東成，副校長； 陳劉欽智博士	主持人: Constance Thanos, Executive Director, DELOS, Pisa, Italy
貴賓致詞	14:45 – 15:45 會場 2
主題報告 陳劉欽智博士，美國西蒙斯學院圖書館 資訊研究所教授	內容和知識管理 主持人: Robin Williams, Manager, Almaden Research Center, San Jose, CA, USA
10:10 – 10:50 會間休息和集體拍照	15:45 – 16:00 會間休息
10:50 – 11:50 特邀報告 1 主持人：劉桂林，清華大學圖書館館長	16:00 – 17:00 會場 1 數位圖書館的發展—國家級
11:50 – 12:30 會場 1 數位圖書館不斷變化的技術和作用 主持人：Ed Fox, Professor of Computer Science, Director of Digital Library Research Lab, Virginia Tech, Blacksburg, VA, USA	主持人: Andy Dong, Director of Technical Research, SMETE, University of California at Berkeley, CA, USA
11:50 – 12:30 會場 2 數位圖書館和遠端教育 主持人：朱強，中國高等教育文獻保障系統 管理中心副主任，北京大學	16:00 – 17:00 會場 2 數位圖書館的發展—案例展示 (1) 主持人: Rush G. Miller, Director of Libraries, University of Pittsburgh, PA, USA
12:30 – 13:45 午餐	17:00 – 17:30 討論 主持人: David Raitt, Editor, The Electronic Library, The Netherlands
	17:30 – 18:15 張貼論文展示
	晚上 18:30 – 20:00 歡迎會 (胡東成副校長主持)

2001 年 5 月 30 日 (星期三)

08:20 – 08:50 註冊和張貼論文展示	12:20 – 13:50 午餐
08:50 – 09:50 特邀報告 3 主持人: 周立柱, 清華大學電腦科學與技術系主任	13:50 – 14:50 特邀報告 4 主持人: Hermann Maurer, Head of IICM, Graz University of Technology and Chief Scientist of KNOW, Graz, Austria
09:50 – 11:05 會場 1 元資料內容、問題和趨勢 主持人: Jay Jordan, President, OCLC, Columbus, OH, USA	14:50 – 15:50 會場 1 多格式、多語言、多媒體資訊檢索 主持人: Jieh Hsiang, Dean of Science, Chi-Nan University, Puli; and Professor of Computer Science, Taiwan University, Taipei
09:50 – 11:05 會場 2 數位圖書館的發展—案例展示 (2) 主持人: David Ruddy, Digital Library Program, Cornell University Library, Cornell, NY, USA	14:50 – 15:50 會場 2 檢索中的文本聚類和分類技術 主持人: 王愛華, 北京大學電腦科學與技術系
11:05 – 11:20 會間休息	15:50 – 16:10 會間休息
11:20 – 12:20 會場 1 本體論和智慧代理 主持人: Jian Qin, Assistant Professor, School of Information Studies, Syracuse University, Syracuse, NY, USA	16:10 – 17:30 專題會議 “中華記憶網” 專題專案研討會 主持人: 陳劉欽智, 周立柱, 項潔教授
	17:30 – 18:15 張貼論文展示

2001年5月31日(星期四)

08:20 - 09:00 張貼時間	12:20 - 13:45 午餐
09:00 - 10:00 特邀報告 5 主持人: 高文, 中國科技大學副校長, 中科院計算所	13:45 - 15:00 綜合討論 Q&A 主持人: 陳劉欽智 (特邀及與會代表大會交流時間)
10:00 - 10:40 會場 1 數位圖書館的發展—案例展示 (3) 主持人: Jane Poley, President, The American Distance Education Consortium (ADEC), Lincoln, NE	15:00 - 15:30 總結和閉幕式
	15:30 - 16:15 校園觀光
	16:30 - 18:30 參觀頤和園
	晚上
	18:30 - 20:30 宴會 (頤和園)
10:00 - 10:40 會場 2 電子資源開發 主持人: 楊宗英, 上海交通大學圖書館教授, 副館長	 <p>NIT Participants from all over the world "No man is an island!"</p>
10:40 - 11:00 會間休息	
11:00 - 12:20 會場 1 數位圖書館的互操作性和其他問題 主持人: Kurt Maly, Kaufman Professor and Chair, Computer Science Department, Old Dominion University, Norfolk, VA, USA	
11:00 - 12:00 會場 2 其他處理技術與檢索問題 主持人: 許卓群, 北京大學電腦科學與技術系	

五、參與人員

此次與會者包括十多位元資訊技術界和數位圖書館研究界傑出人物的、來自 20 個國家的 70 多位國外專家、教授與中國衆多高校、國家圖書館和科學院的資訊技術和圖書館界的專家雲集北京，實是一次盛會。

- 會議總召集者
陳劉欽智教授
Graduate School of Library and Information Science
Simmons College, Boston, MA 02115, USA
中國北京清華大學客座教授
- 會議籌備總召集者
胡東成教授，副校長
中國北京清華大學
- 劉桂林，清華大學圖書館館長
- Ed Fox, Professor of Computer Science, Director of Digital Library Research Lab,
Virginia Tech, Blacksburg, VA, USA
- 朱強，中國高等教育文獻保障系統管理中心副主任，北京大學
- Constance Thanos, Executive Director, DELOS, Pisa, Italy
- Robin Williams, Manager, Almaden Research Center, San Jose, CA, USA
- Andy Dong, Director of Technical Research, SMETE, University of California at
Berkeley, CA, USA
- Rush G. Miller, Director of Libraries, University of Pittsburgh, PA, USA
- David Raitt, Editor, The Electronic Library, The Netherlands
- 周立柱，清華大學電腦科學與技術系主任
- Jay Jordan, President, OCLC, Columbus, OH, USA
- David Ruddy, Digital Library Program, Cornell University Library, Cornell, NY, USA
- Jian Qin, Assistant Professor, School of Information Studies, Syracuse University,
Syracuse, NY, USA
- 楊冬青，北京大學電腦科學與技術系教授
- Hermann Maurer, Head of IICM, Graz University of Technology and Chief Scientist of
KNOW, Graz, Austria
- Jieh Hsiang, Dean of Science, Chi-Nan University, Puli; and Professor of Computer
Science, Taiwan University, Taipei
- 王愛華，北京大學電腦科學與技術系
- 高文，中國科技大學副校長，中科院計算所
- Jane Poley, President, The American Distance Education Consortium (ADEC), Lincoln,

NE

- 楊宗英，上海交通大學圖書館教授，副館長
- Kurt Maly, Kaufman Professor and Chair, Computer Science Department, Old Dominion University, Norfolk, VA, USA
- 許卓群，北京大學電腦科學與技術系

六、與會心得與建議

1. 考察活動簡介

國際數位圖書館合作研究計畫 (International Digital Library Program Project, 簡稱 IDLP), 整合台灣 (台灣大學、清華大學、中央研究院)、大陸 (北京大學、北京清華大學、上海交通大學) 和美國 (Simmons 等大學三地的研究人員, 共同合作開發數位圖書館關鍵技術。本次訪問擬考察大陸相關單位的發展情況, 並建立未來合作的模式。)

2. 考察活動經過

本次訪問由台灣大學資訊工程學系項潔教授和陳信希教授、圖書資訊學系陳雪華教授以及故宮博物院資訊中心蔡順慈主任和賴鼎聲博士等組成。報告人與資工系陳信希教授於 5 月 27 日, 搭乘澳門航空 NX611 班機, 在澳門轉 NX002 班機赴北京, 住清華大學附近北京西郊賓館。

5 月 28 日早上在故宮蔡主任的安排下, 赴北京中國歷史博物館訪問, 由李季副館長和谷長江副館長接待, 並由肖飛先生 (IT 主管) 作簡報。主要是探討中國歷史博物館文物數位化的狀況, 由於人力編制的關係, 該館目前只有兩位人員專門負責網站的維護, 整體數位博物館建製較無長遠的規畫。下午安排到北京故宮訪問, 由外事辦公室楊森先生接待。台大與故宮在國科會數位博物館計畫下, 合作設計故宮博物館, 此次訪問擬觀摩北京故宮在數位化方面的進展。由於當天並沒有安排好, 因此只座室外導覽, 數位博物館合作部分, 另外安排在 5 月 31 日早上再繼續討論。

5 月 29 日早上參與 NIT2001 國際會議的開幕式 (由 IDLP 計畫美方主持人陳劉清智教授主持)。下午訪問中國科學技術館, 由王渝生館長接待, 並由前館長李象益教授親至導覽。這個館去年才開幕, 整體很新穎, 但由於經費的關係, 科技館裡很多設備因為怕壞掉並沒有完全開放, 不過兩位前後館長的觀念算蠻開放。

5 月 30 日參與 NIT2001 國際會議並與 IDLP 大陸合作單位討論未來之發展與合作項目。5 月 31 日參觀清華大學校園以及清華大學附設的「育成中心」(學研樓)。

6月1日在NIT2001主辦單位安排下，遊覽長城及頤和園兩處。6月2日自由活動。6月3日訪問北京大學計算語言學研究所。6月3日下午搭澳航NX001班機，至澳門轉澳航NX608回台北。

3. 結語

北京大學、北京清華大學和上海交通大學，是本項數位圖書館計畫大陸地區的合作夥伴。在中文數位圖書館的研究，這三所學校扮演非常重要的角色。北京大學、北京清華大學和上海交通大學是大陸最著名的三所學府，在人力和設備上都擁有最好的競爭力。透過研究訪問，交換不同的心得，所學到的分工整合，更是難得的經驗，對國內數位圖書館的推展與整合有助益。

4. 帶回的資料

[1] 中國歷史博物館網站簡介

[2] 中國科學技術館簡介，科學隧道，中國古代科學展覽等

[3] NIT-2001 第十二屆新信息技術國際會議論文集：Global Digital Library

Development in the Millennium 專書一冊