

架構在WWW與Z39.50上的近似  
自然語言OPAC檢索系統

A Quasi-Natural Language Searchable OPAC System  
Based on WWW Technology and Z39.50 Protocol

曾元顯 Yuen-hsien Tseng

輔仁大學圖書資訊學系副教授

Associate Professor, Dept. of Library and Information Science,

Fu Jen Catholic University

E-mail: tseng@blue.lins.fju.edu.tw

【摘要】

近幾年圖書館的OPAC系統服務建設，雖可讓讀者時空無礙的連線查詢，但不同的圖書館採用的系統並不一致，使用者必須面對不同的操作畫面與使用上的些許差異，而查詢時也必須一一連線才得以進行。為改善此種缺點，本文透過系統實作方式，探討在WWW架構下，利用免費的軟體開發符合Z39.50資訊檢索標準的OPAC系統，並使其具備近似自然語言查詢功能的過程，並說明其中可能遭遇的困難或隱含的問題，以作為其他單位的參考。希望本文的內容，對推廣Z39.50資訊檢索標準的運用與服務，有所助益。

【Abstract】

Recent development in OPAC systems have made search and retrieval of library catalogs anywhere, anytime possible. However, due to the discrepancies among different systems, users have to face different interfaces and usage. Furthermore, searching among several OPAC systems requires serial connection to one after another. To overcome this problem, this article discusses an implementation of an OPAC system based on the WWW and Z39.50 protocols. The implemented OPAC system allow multi-database search and natural language-like queries, thus reducing the tedious work and difficulties in query formulation. Possible problems encountered in developing such a system are discussed. It is hoped that the

implementation experiences will help other libraries in establishing a similar Z39.50 service.

關鍵詞：全球資訊網；Z39.50資訊檢索標準；線上公用目錄系統；近似自然語言檢索；多資料庫查詢

Keywords：World Wide Web; Z39.50 information retrieval standard; On-line public access catalog system; Quasi-natural language retrieval; Multi-database search

## 壹、前言

隨著網際網路的蓬勃發展，各地圖書館除加強自動化系統的建設外，也紛紛進行網路化的服務。在圖書館的網路服務中，跟大多數讀者最直接相關的為「線上公用目錄」（OPAC, On-line Public Access Catalog）系統。過去的OPAC系統都以遠端登錄（telnet）方式提供使用，然而遠端登錄有諸多缺點（註1），目前則大多採用WWW（World Wide Web）的圖形視窗環境，提供使用者一致性的簡易操作介面。然而這樣的改進，還有未盡完善的地方。對於希望查詢多個圖書館以便找到完整資料的讀者而言（即「多資料庫查詢」需求），雖可透過網路連上任何一個OPAC系統，可是不同的圖書館採用的系統並不一致，使用者必須面對不同的操作畫面與使用上些許的差異，而查詢時也必須一一連線才得以進行。

針對此種情況，目前有幾種解決方法。一是個別廠商將自己的系統串連起來，讓讀者連上任一圖書館即可

查詢其他屬於同一系統的館藏。（註2）然而此種個別廠商提出解決方案的作法，並不容易推行到跨廠商之間的合作。另一種方法是透過meta-search engine（集中式檢索引擎）將使用者的查詢條件送給數個OPAC系統，再將各個系統的查詢結果合併、整理後再傳回。目前國內有中正大學圖書館製作這樣的書目檢索系統（註3）（註4）。此方法可讓使用者只需連上此網站，即可用相同條件查詢多個OPAC系統的資料，而同時各圖書館的OPAC系統無須負擔任何更動或維護的工作，即可配合此種使用方式。然而，由於使用者都集中連線到此網站來，就整個查詢過程而言，此網站變成查詢流程的瓶頸。另外，各個OPAC系統一致、相同的欄位，才有被meta-search engine接納，進而提供利用的可能，對於個別圖書館逐漸豐富的其他類型館藏，如政府資訊、地理資料、數位圖書資源等，則較難支援。因此，圖書館本身還是需要另謀辦法，才能提供使用者介面一致欄位完整的多資料庫查詢。

相對於集中式的查詢，另一種方

式為分散式查詢。美國1988年即制訂一項現今稱為Z39.50的資訊檢索標準，以支援網路環境下的分散式資料庫查詢，讓使用者以相同的介面與使用方式，查詢不同主機、不同系統的資料庫(註5)(註6)。目前，Z39.50已成為一項國際標準，在歐洲、美國、澳洲等地都有相關的計畫，進行OPAC系統或數位化圖書館系統的整合。

Z39.50是一種主從架構的網路通訊協定(protocol)。其特點是客戶端(client)與伺服器端(server)的功能獨立，而得以提供使用者介面最大的彈性。另一特點是其為應用層(application layer)協定，因此與使用的電腦平台、作業系統及檢索引擎無關，容許系統之間具備高度的互通性(interoperability)。另外，此協定支援多種型態的資料格式與檢索點，讓客戶端易於整合並連結不同的資料庫與伺服器。相較於WWW的HTTP(HyperText Transfer Protocol, 超媒體傳輸協定)此種不記錄使用者狀態(stateless)的協定，Z39.50記錄使用者連線查詢的狀態(stateful)，因此適合資訊檢索的應用，進行多資料庫查詢時，可以提供一致性的檢索介面，免除重新學習的困擾。

與HTTP的主從架構一樣，Z39.50使用時也需要客戶端軟體才能連上Z39.50伺服器。然而WWW因為可以廣泛運用於任一單位、個人，作為出版、文宣、佈告、甚至與使用者互動等用途，加上免費伺服器與瀏覽器軟體的普及，以及維護、操作上的簡易，使得目前

幾乎所有的電腦系統都能連上全球資訊網。相對的，Z39.50的軟體較不普遍，一般人並不熟悉。由於牽涉到要求使用者下載、安裝、使用Z39.50軟體，因此，想要推廣使用者以Z39.50協定檢索資料並非易事。較佳的解決方案，是透過既有的WWW瀏覽器來檢索Z39.50伺服器的資料庫。這只需要資料庫維護單位或其他相關單位，在WWW伺服器端提供連上Z39.50伺服器的通道(gateway)軟體即可，使用者端無須作任何更動或修正。

雖然Z39.50具備上述的優點，然而大部份檢索系統都已提供的Ranked List Query功能(對中文檢索而言，可概略稱為相關程度排序查詢、模糊搜尋、近似自然語言查詢、或近似字串查詢等)，則極少軟體支援。就目前蒐集到的資料所知，僅美國NIST機構發展的ZPRISE軟體，可在Sun Solaris作業系統下，提供Z39.50協定的相關程度排序查詢。(註7)

本文中，我們將利用免費的Z39.50軟體，以系統實作的方式，建構一套架構在WWW上，符合Z39.50標準，且具備近似自然語言查詢功能的OPAC系統，並探討其中可能遭遇的困難或隱含的問題，以作為其他單位的參考。希望本文的內容，對推廣Z39.50資訊檢索標準的運用與服務，有所助益。本文後續的內容安排如下：下一節將先介紹Z39.50提供的服務項目及其重要的概念、報導Z39.50相關計畫的實施成效、並就查詢流程、維護方式與運用彈性

等方面與meta-searchengine系統作比較；第三節以免費的Z39.50軟體為例，描述結合WWW與Z39.50檢索系統的架構與實作過程；第四節則列舉實例，展示其結果；最後說明我們對Z39.50未來發展的看法。

## 貳、Z39.50 簡介與回顧

### Z39.50 發展沿革

在1980年代初時，有關資訊檢索的通訊協定即已開始制訂。美國於1988年提出 Z39.50 第一個版本，此版本後來為WAIS協定的基礎，但目前已與其後的版本不相容。1992年 Z39.50 第二版推出，制訂七項資訊檢索服務，有多家廠商與非營利單位實作相關的軟體。1995年推出第三版，制訂的檢索服務擴增為 11 項，惟至今完全符合第三版標準的 Z39.50 軟體並不多見，不過此版本與國際標準組織的查詢與檢索協定（ISO 10162/10163 Search and Retrieval protocol）一致，兩者具備可互相運作的特性，即互通性（interoperability）。所以 Z39.50 目前是國際標準，我國也已採用為資訊檢索協定標準CNS 13461。（註8）

### Z39.50 提供的服務協定

使用者以 Z39.50 協定作資料檢索時，事實上是由客戶端電腦經網路發出各項符合 Z39.50 協定的服務請求，由伺服器端電腦依照此標準一一回應。其規範的服務項目在第三版中有下列

十一項：

- 一、啓始（Initialization）：規範客戶端連上伺服器、伺服器回應的訊息與過程。
- 二、存取控制（Access Control）：連上伺服器時，所做的身份、密碼確認。
- 三、解釋(Explain)：客戶端可發出請求，請伺服器解釋其可供利用的資料庫及相關訊息，如資料庫名稱與使用說明、支援的資料格式與屬性等關於資料庫的metadata。
- 四、瀏覽／掃描（Browse／Scan）：客戶端可請求伺服器列出可供檢索的詞彙，如人名、題名等，據以瞭解資料庫中的索引詞彙或控制詞彙。
- 五、查詢（Search）：使用者的查詢條件，經客戶端軟體處理包裝後提出查詢，伺服器回應查詢摘要（所得筆數）與部份結果，並保留此結果供後續的請求利用。使用者可就某一查詢結果予以命名，便於和後續的查詢作交集或聯集的運用。此服務可供利用的查詢型態與功能將於稍後介紹。
- 六、會計／資源控制（Accounting Resource Control）：客戶端中斷伺服器的處理、或要求回應檢索計費狀況等服務，讓使用者在計費的檢索服務系統內瞭解其資源使用狀況。
- 七、排序（Sort）：客戶端可請求伺服器就某欄位（如題名、年代）或

某條件將查詢結果排序以便於檢視。

八、擷取 (Retrieve)：此部份包含展現 (present) 與分段 (segment) 服務。客戶端可就某些記錄以特定範圍、欄位或格式請求展現資料，當回應的資料非常大時 (如圖形檔案)，可請求以分段方式傳回。

九、延伸服務 (Extended Services)：使用者可由此項目，利用到沒有正式規範在Z39.50協定內的其他服務，如館際互借、定期查詢、資料庫更新等功能。

十、結果集刪除 (Result-Set Delete)：用以刪除伺服器儲存的查詢結果，降低伺服器的負擔。

十一、結束 (Termination)：兩端電腦結束查詢會期，斷線離開的服務。

### 支援的記錄格式

Z39.50支援多種資料庫記錄格式，使查詢結果可依照指定的格式正確顯示。較常見的格式有下列幾項：

- 一、MARC (MACHine Readable Catalog)：書目記錄機讀格式。
- 二、OPAC (Online Public Access Catalog)：線上公用目錄格式，包含館藏流通等訊息。
- 三、SUTRS (Simple Unstructured Text Record Syntax)：簡單非結構文字記錄格式，可運用於一般的純文字檔。

四、HTML (Hyper Text Markup Language)：超文件標示語言格式。

五、GRS-1 (Generic Record Syntax)：一般記錄格式，適用於階層式標示記錄，即類似SGML、HTML之類的文件。

### 屬性集 (attribute set)

Z39.50可提供的查詢條件，可以說是由伺服器端支援或選用的屬性集所決定。屬性集內列出可用的屬性元素 (attribute element)，其中每個元素包含一對屬性型態 (attribute type) 及其屬性值 (attribute value)。屬性集來自對資料庫記錄查詢的需求，不同領域可能需求不盡相同。因此屬性集可就個別領域的需求而另行制訂，或擴充原有的部份產生另一個屬性集。目前較常討論到的屬性集有下面幾種：

- 一、Bib-1 (bibliographic attribute set)：書目性記錄屬性集。
  - 二、STAS (Scientific and Technical Attribute Set)：科學與技術屬性集，乃擴充Bib-1而來。
  - 三、GILS (Government Information Locator Service)：政府資訊定位服務，為檢索政府資訊所制訂的屬性集。
  - 四、CIMI (Consortium for the Computer Interchange of Museum Information)：為檢索、交換博物館資料所制訂的屬性集，其一部份也是延伸Bib-1而來。
- 其中Bib-1幾乎為所有的Z39.50軟體

所支援。Bib-1 有Use、Relation、Position、Structure、Truncation、Completeness等六個屬性型態。其中Use是用來描述所要查詢的欄位，例如Use的屬性值為author時表示要查詢作者欄位；Relation描述欄位比較如何運算，例如以「小於」方式查詢年代欄位；Position描述查詢的位置，如本文中任意位置；Truncation指定作何種切截比對，如左切截、右切截、不作切截等等。稍後的範例中，有這些屬性型態的用法，讀者會更清楚他們在查詢中的作用。

### 查詢型態

不同的查詢型態有不同的語法，Z39.50一共支援下列六種查詢型態：

- 一、Type-0：此為私用型態，只要伺服器端與客戶端事先知會、互相知曉，即可用標準外的特別語法進行查詢。
- 二、Type-1：此為Z39.50最主要的型態，以後置語法（Reverse Polish Notation, RPN）來表達查詢詞彙及其屬性條件。
- 三、Type-2：ISO 8777 型態的查詢語法。
- 四、Type-100：即共同命令語言（Common Command Language, CCL）語法。
- 五、Type-101：Type-1 的擴增語法，允許鄰近字串的查詢，以及對檢索結果以不同的屬性作限制查詢。
- 六、Type-102：即Ranked List Query，

用以支援多數系統已經具備將結果按相關或近似程度排序的查詢，如模糊搜尋、近似自然語言查詢等。

基本上客戶端可以指定上列任一型態查詢伺服器，伺服器再據以剖析查詢條件，作適當的查詢處理。其中Type-1 為基本的查詢語法，任何伺服器都必須支援。

介紹到此，我們可以舉幾個以Bib-1 為屬性集並以 Type-1 型態為語法的查詢範例。部份客戶端軟體在初次使用作設定時，會要求使用者將查詢條件的屬性型態及其屬性值以編號表示。例如，查詢條件：「水資源（1, 62, 2, 3, 3, 4, 1, 5, 100）」，表示要查詢的欄位為摘要（第一個1代表use屬性型態，其值為62在Bib-1中代表摘要），而查詢時要以字串相等的方式作比對（2代表 relation 屬性，其值3代表相等），並查詢該欄位中任意位置（3代表 position，其值3代表任意位置），此查詢字串的結構為片語（4代表structure，其值1代表片語），另外不作切截查詢（5代表 truncation，其值100代表不作左或右切截），最後完整屬性（6為completeness）沒有指定。另舉一個用到多個查詢詞彙的例子，假若查詢條件為「OR(AND (information [1, 4], retrieval [1, 4]), system [1, 62])」，則表示要在題名中（4在Bib-1中代表題名）查詢 information 以及 retrieval 或是在摘要中查詢含system的記錄。

當然，上述查詢條件過於繁瑣，

一般使用者並不熟悉，所以大多數客戶端軟體都提供視窗元件讓使用者以點選方式輸入查詢條件，使查詢介面簡單易用。

### 相關計畫

近年來，在歐、美、澳等地有相當多運用Z39.50資訊檢索標準的計畫。（註9）這些計畫主要在連結各個圖書館不同的OPAC系統，使系統之間具備互通的特性，以便利使用者的利用。例如英國的M25 Link計畫，乃用以連結英國M2聯盟38個圖書館的OPAC系統，這些圖書館使用包括UNICORN、INNOPAC、TALIS、HORIZON、LIBERTAS在內的數家不同廠商系統。此計畫自1998年起進行規模較小的整合，為期兩年，伺這兩年的先導型計畫完成後，自2000年起進入全面整合的第二階段，為期也是兩年。（註10）另外，英國還有Bradford大學的BOPAC研究計畫；蘇格蘭的CAIRNS（Co-operative Academic Information Retrieval Network for Scotland）計畫。美國有Iowa州22個圖書館參與的SILO（State of Iowa Libraries Online）計畫。加拿大有虛擬聯合目錄計畫（Virtual Canadian Union Catalogue Project）。德國有DBV-OSI、義大利有SIBYLLA、法國有Aquarelle、整個歐洲有Europagate、以及澳洲有ZedWeb計畫。在數位圖書館方面，有美國政府資訊系統（GILS）、美國地理資訊系統（FGDC）、電子博物館（CIMI）等也都以Z39.50作為跨資料庫

的檢索標準。

### Z39.50 成效評估

美國Iowa州的SILO計畫（State of Iowa Libraries Online）曾就Z39.50連結各圖書館的成效做過評估。（註11）此計畫共有22所圖書館參與，而各館所採用的自動化系統不同廠商合計6家。其評估方案包括問卷、訪談與軟體功能評估，總共回收20份圖書館問卷及4份廠商問卷。評估的重點在瞭解網路上建立Z39.50服務的難易、Z39.50服務的優點與缺點、廠商Z39.50應用的情形、以及Z39.50與WWW的結合使用效果。茲列舉重要結果如下：

- 一、各圖書館採購Z39.50伺服器的最主要考量是與既有圖書館系統的相容性，其次為價格、廠商熟悉度與技術支援服務。
- 二、Z39.50對圖書館的重要性方面，有過半數圖書館認為非常重要，而「能透過網際網路查詢存取資料」、「單一使用介面查詢多個資料庫」被認為是最重要的兩個項目。
- 三、就Z39.50服務的優點方面，最常被列舉的項目有：共同的檢索介面、可查詢他館、以額外方式查詢該館、軟體的互通性等四項。
- 四、就Z39.50服務的缺點方面，被列舉最多的有：廠商系統的不一致性、缺乏館藏流通訊息（即沒有顯示藉閱記錄、也沒有預約功能）、客戶端軟體安裝維護問題（有新

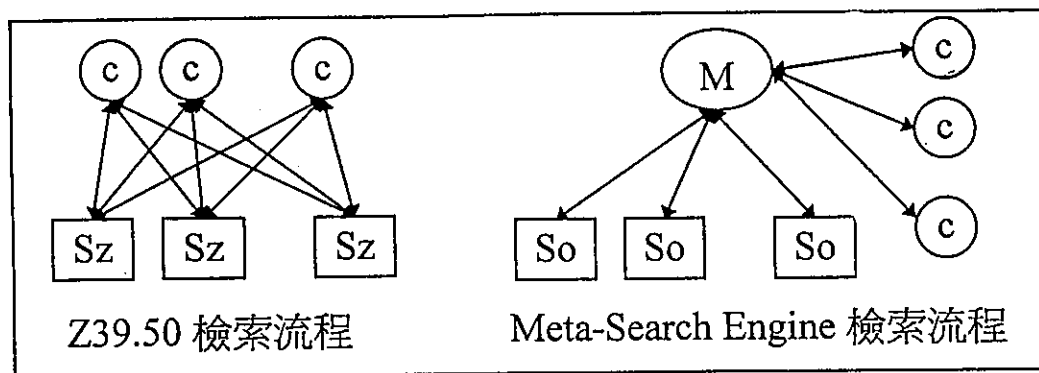
館加入需個別重新設定客戶端軟體)、檢索功能不足缺乏原有系統完整功能、反應慢、說明文件不足、缺乏對 Z39.50 的瞭解。

五、對 SILO 計畫的滿意度方面，約半數圖書館認為計畫非常成功，而大多數認為成功的因素有：單一使用介面查詢多館資料，展現各館互通運作能力；經由網路查詢存取，各館同氣連枝宛如單一圖書館；館際互借成功。

### Z39.50 與 Meta-Search Engine 系統之比較

Z39.50協定支援網路環境下分散各處、不同性質資料庫的查詢，而meta-search engine也具備類似的作用，兩者資訊檢索的流程如圖一所示。在Z39.50方面，客戶端可選擇連上一個至多個Z39.50伺服器，而以Z39.50規範的功能查詢資料庫，至於能否同一時間進行多資料庫查詢取決於客戶端軟體有否支援此項功能。在meta-search engine方

面，一般都建構在WWW上讓使用者連上此檢索引擎，進而選擇想要查詢的資料庫(或OPAC)，下達查詢條件，再由此檢索引擎將查詢條件做適當的轉換，分別傳給選定的資料庫進行檢索。之後檢索引擎蒐集各個查詢結果，經過整理後(如輸出格式的處理)再傳回給使用者。此兩種系統在查詢模式、維護難易及資料支援能力方面的比較如表一所示。由此表可知，Z39.50不會造成查詢流程的瓶頸，支援的資料格式也比較豐富，並且具備擴充的彈性，應該是網路環境下較佳的檢索模式。可惜 Z39.50 的軟體不夠普及及易用，目前實作的軟體大部份也都只提供 Z39.50 最基本的服務，如起始、查詢、擷取、結束等數項。其安裝維護上的困難，使國內大部份的圖書館尚未提供此項服務。綜合而言，Meta-search engine使用上較為簡易，但不易擴充規模；而 Z39.50 使用彈性大，但現況是軟體不夠普及、且查詢介面通常都不及原有(OPAC)系統豐富。



圖一：檢索流程比較圖。(其中c表客戶端、Sz表Z39.50伺服器、So表OPAC查詢主機、M為Meta-search engine伺服器)



表一：Z39.50與Meta-Search Engine比較表

	Z39.50系統	Meta-Search Engine系統
查詢	完全分散處理；	「集中、分散、集中」的處理模式；
維護	客戶端需知道伺服器的主機位址、埠號(port)、資料庫名稱才能連線查詢； 伺服器端需個別安裝、維護；	完全由meta-search端維護，使用者、伺服器端不需額外費力；
彈性	支援豐富的資料格式與查詢欄位；	查詢欄位有限（需各資料庫端一致）； 各種資料格式支援不易（需瞭解資料庫端的資料格式）；

### 參、系統架構

過去數年，國外已有許多家廠商推出Z39.50相關的軟體（註12）（註13）。在免費軟體方面，主要以客戶端軟體較多，伺服器端軟體較少，而能夠結合本身原有資料庫功能的伺服器軟體更少。就目前蒐集到的資料所知，免費的工具軟體有ZedKit（for Unix）（註14）及YAT toolkit（註15），可讓使用者發展出自己所需的Z39.50 伺服器或客戶端軟體。合法現成的免費Z39.50伺服器軟體有CNIDR的Isite（註16）及NIST的ZPRISE，其中ZPRISE雖可支援相關回饋、排序查詢，但只有Sun Solaris作業系統的版本，移植性低。因此可運用的現成伺服器軟體僅有Isite系統了。

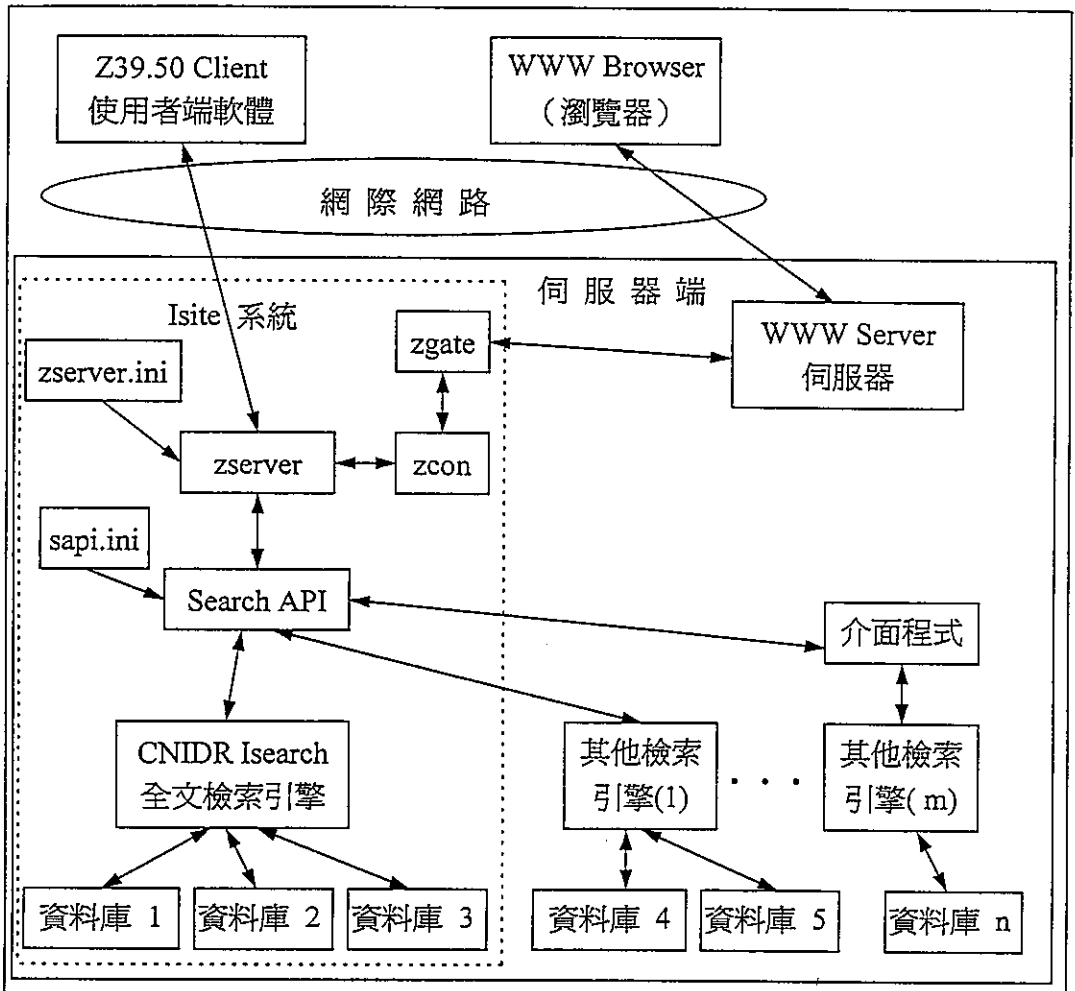
Isite資訊系統是由美國網路資訊發掘與檢索中心（CNIDR, Clearinghouse

for Networked Information Discovery and Retrieval）發展，內容包含一套Isearch英文的全文檢索引擎以及Z39.50相關的軟體。此系統可在大部份的Unix及Windows NT 作業系統上執行，目前的正式版本為2.01版，可從CNIDR的FTP站上取得：[ftp://ftp.cnidr.org/pub/NIDR\\_tools/Isite/](ftp://ftp.cnidr.org/pub/NIDR_tools/Isite/)。

Isite資訊系統架構如圖二所示。圖二顯示使用者可直接以 Z39.50 client 軟體連上 Z39.50 伺服器，也可以透過較為普及的 WWW 瀏覽器連線。如果是透過瀏覽器連線，WWW 伺服器這邊必須有一個 Z39.50 連結介面。此一連線的動作會讓使用者透過 Isite 的zgate 與zcon程式連上 Z39.50 伺服器（即zserver程式），其中 zgate 是一支 CGI（Common Gateway Interface）程式，它會針對不同的使用者查詢會期（session）而啟動不同的 zcon，由 zcon

再連上zserver伺服器以服務不同的查詢會期。所以相對應於圖二的架構圖，此連線動作的實際流程如圖三所示。使用者的連線由zgate來管理，使用者端首次連線時，zgate會回應使用者端一個隱藏變數：「session-ID」，變數內存放用以辨識查詢會期的資訊，使用者端後續的檢索連線都會將此「session-ID」的內容傳回伺服器端。

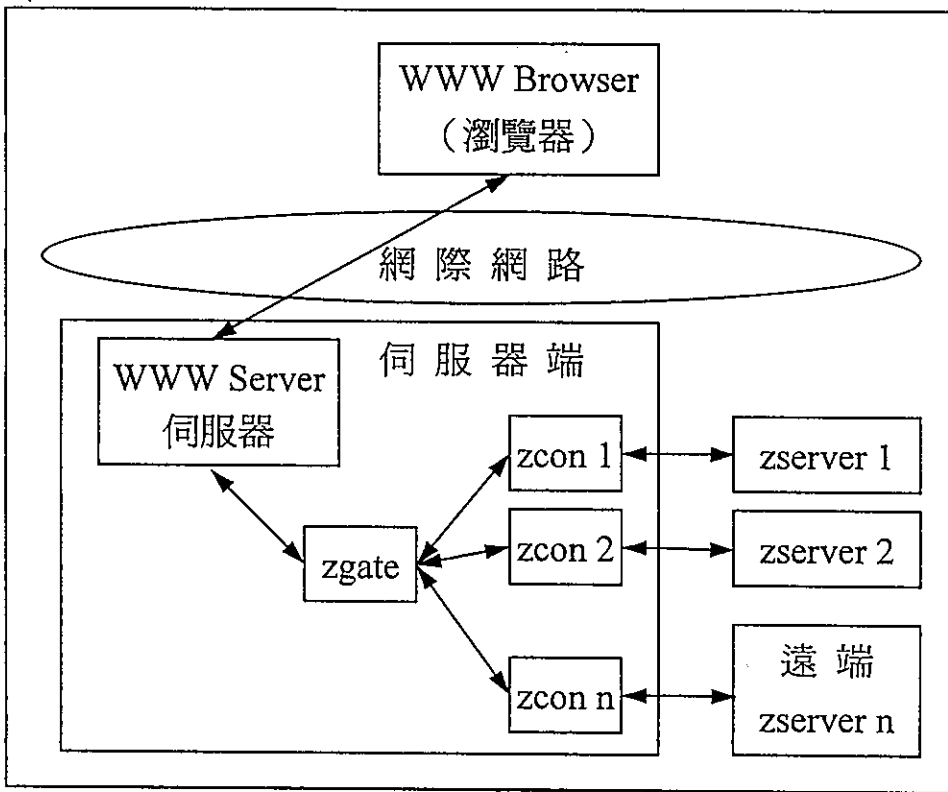
藉此隱藏變數，zgate 可以辨認使用者的連線，若是新的連線則啟動一個新的zcon程式，若已發生過則叫用已經啟動的zcon。而zcon的功用則宛如一個Z39.50客戶端，用以連上zserver伺服器，此伺服器可以在同一部主機，也可以在遠端的另一部機器上。這就是為什麼我們可以在某個網站透過Z39.50檢索標準，來查詢不同主機資料庫的原因。



圖二：Isite 資訊系統架構圖

至於zserver伺服器可連接哪些資料庫則非常彈性。從圖二可知，此資料庫可以是Isite的資料庫，也可以是安裝人員加裝的其他資料庫。Isite系統內附的Isearch全文檢索程式，可用以檢索HTML、SGML以及SUTRS無結構性文字檔案等多種格式文件。使用者也可以將已有的資料庫連上（如書目資料），使其具備 Z39.50 檢索標準的功能。透過 Isite 制訂的檢索應用程式介面SAPI（Search Application Programming

Interface），只要使用者採用的檢索引擎其輸出與輸入符合此介面的規定，再加上對兩個組態檔（zserver.ini 與 sapi.ini）適當的設定，即可成功連接zserver伺服器。但是由於一般的檢索引擎並不是專為Isite系統而設計，所以其輸出與輸入不會符合SAPI的規定，這時我們就必須自行撰寫檢索引擎的介面程式，將其輸出與輸入進一步處理，並轉成必要的格式，才能正確的連接Isite的伺服器。



圖三：Isite系統HTTP連接Z39.50的系統架構圖

使用者的查詢條件可透過瀏覽器傳給WWW伺服器，再透過 zgate 與 zcon 兩支介面程式連上 zserver 伺服器，zserver 透過SAPI介面呼叫檢索引擎，檢索引擎查詢完資料後將結果回報給zserver伺服器，zserver 透過 zcon 與 zgate 程式將結果傳給WWW伺服器，再由WWW伺服器傳資料給使用者。圖二中虛線內的程式與檔案都由Isite系統提供，所以除了前述組態檔案的設定及介面程式的撰寫工作外，讀者只要準備WWW伺服器、自建的資料庫以及其檢索引擎即可。

Isite 定義的 SAPI 是相當簡單的軟體介面。假若我們有自己的資料庫要連上Isite，此資料庫的名稱假設為OpacDB，那麼我們要在組態檔中加入一些訊息讓 zserver 知道。作法是在組態檔案 zserver.ini 中加入（或修改）這一行：

```
DBList = OpacDB, OtherDB1, Other DB2
```

亦即把要宣告的資料庫名稱寫在DBList這一行，若資料庫超過一個以上則以逗點「,」分開。另外在 sapi.ini 檔案中加入底下數行宣告：

```
[OpacDB]
Type = SCRIPT
Location=/usr/bin/Search
GetFull=/usr/bin/GetFull
Results=/tmp/results
```

其中Location的值代表我們檢索引擎的介面程式其所在路徑與名稱，在此例中名稱為Search；GetFull的值為使用者就某筆檢索結果擷取其完整記錄

所需呼叫的程式，此例中為GetFull；而Results則是檢索引擎的介面程式存放查詢結果的檔案名稱。當使用者的查詢透過 Z39.50 送進來時，zserver 會以下面命令列的方式呼較我們的介面程式：

```
/usr/bin/Search /tmp/results.<pid>
QueryTerm[attr1,value1]
```

這時我們的介面程式就可從命令列的第二個參數讀到檢索詞及其屬性條件，其格式如下：QueryTerm[attr1,value1]。介面程式 Search 可就此條件呼叫檢索引擎，待獲得查詢結果時，把結果寫到第一個參數指定的檔案內，此例中其檔名為/tmp/results.<pid>（在此<pid>表示 process 的編號，乃系統自動給定無須理會）。另外，將查詢結果寫到檔案時必須符合下列格式，zserver才能正確解讀：

```
[Default]
HitCount=2
Diagnostic=0
Separator=##separator string - your choice
##
[Data]
Key-for-record1
Record data for record number 1
##separator string - your choice##
Key-for-record2
Record data for record number 2
```

此範例格式中，查詢結果摘要列在 [default] 群組下，查詢結果則列在 [data] 群組下。查詢結果有兩筆記錄，每筆記錄寫成兩行，並以自訂的記錄

區隔字串分開。記錄的第一行為該筆資料的關鍵值，如 BRN 或 ISBN 等，用以讓使用者透過該關鍵值取得此筆資料的完整內容。記錄的第二行為該筆資料的簡易訊息，如書名、作者名、出版年等，作為讓使用者下達查詢條件後即可瀏覽檢視的查詢結果。使用者瀏覽檢索結果時，若想就某一筆記錄察看其詳細內容，可透過瀏覽器的點選操作，將此需求傳進伺服器。同樣的 zserver 會以下面命令列的方式呼叫我們的 GetFull 程式：

```
/usr/bin/GetFull/tmp/<tmpname> Key-for-record-n
```

這時 GetFull 程式以命令列上傳來的第二個參數（此例中為Key-for-record-n）查詢完整記錄，將結果寫入第一個參數所指定的檔案內（此例中為/tmp/<tmpname>），zserver便可正確接收，並繼續傳回給使用者檢視了。

#### 肆、Z39.50 檢索範例展示

以客戶端軟體連接一個Z39.50伺服器時，需要知道它的主機位址（IP或Domain Name address）、埠號（port）及提供的資料庫名稱，必要時還需要其他資料如記錄展現語法等。跟WWW情況類似的是，使用者常常不清楚網路上有哪些站臺可以連線，以及連線時所需的參數，如埠號、資料庫名稱等。因此有些客戶端軟體可以由維護人員預先設定好連線的組態，使用時只要點選預設的資料庫即可。但部份軟體設計缺乏彈性，往往只能連往預

設的資料庫，無法隨時指定其他資料庫進行連線。

國內圖書館目前提供Z39.50服務的有國家圖書館、省立台中圖書館、中央研究院、臺大、政大及輔大。其中輔大用的是免費的Isite系統（註17），省立台中圖書館採用的是商用Dynix系統的 Z39.50 伺服器以及工研院電通所開發的 Z39.50 客戶端軟體（註18），而其餘圖書館皆採商用INNOPAC系統的 Z39.50 伺服器與客戶端軟體。表二、表三列出國內 Z39.50 伺服器連線參數的一覽表及客戶端軟體連線網址。所有這些系統，其客戶端軟體的連線組態除了由維護人員設定外，使用者不能任意連往他處。因此，在使用上較不具彈性。此外，國內的 Z39.50 服務還有一些問題。以INNOPAC為例，其 Z39.50 客戶端軟體有兩種，一種是終端模擬的 Telnet 介面，另一種是 WWW 介面。經測試，其終端模擬介面除中文 Big-5 與 CCCII 碼的部份轉換問題外（註19），大部份皆可以正常運作，而且可用相同條件作多資料庫查詢。然而在其WWW介面方面，一次僅能查詢一個資料庫，且系統雖能正確回應輸入的英文詞彙，但輸入中文詞彙時，如「水資源」、「資訊檢索」等，卻發生「Un-supported search」、「Truncated words too short」等不正常情況。

如前所述，我們希望發展出一套系統，既符合 Z39.50 標準而具備與其他系統互通的特性，也同時能提供先

進的近似字串查詢功能而便利使用者表達查詢條件以及加快檢索結果的瀏覽與檢視。圖四所示為利用 Isite 系統，以及我們自己開發的Crystal檢索引擎（註20），透過SAPI介面程式的撰寫，完成的輔大 Z39.50 近似自然語言書目檢索系統。由於輔大的系統架構在WWW上，因此使用者只要連上http://

xlib.fju.edu.tw/Isite 網頁，選取「連線」後即可看到如圖四的畫面。使用者可在此畫面中輸入任意自由詞彙，系統會以模糊搜尋方式比對題名及作者欄位中的資料，並將結果按近似程度排列。圖五所示為查詢「電視廣告對兒童的影響」的範例。

表二：國內Z39.50伺服器連線參數一覽表

圖 書 館	伺 服 器 位 址	埠號	資 料 庫	記 錄 語 法
國 家 圖 書 館	nbnet.ncl.edu.tw	210	innopac	SUTRS*
省立台中圖書館	lib.ptl.edu.tw	210	ptlchinese	未知**
中 央 研 究 院	las.sinica.edu.tw	210	innopac	SUTRS*
台 大	tulips.ntu.edu.tw	210	innopac	SUTRS*
政 大	jenda.lib.nccu.edu.tw	210	innopac	SUTRS*
輔 大	xlib.fju.edu.tw	210	libl	HTML*

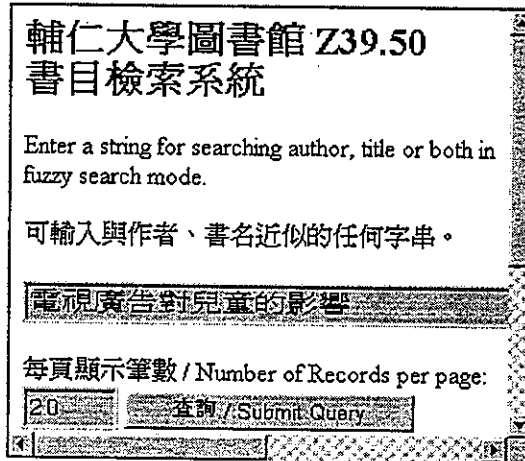
註：\*經測試，這些系統用其他記錄語法也大多正常，這裡所列為建議使用的記錄語法。

\*\*省中圖的系統經常連不上，難以測得其記錄語法。

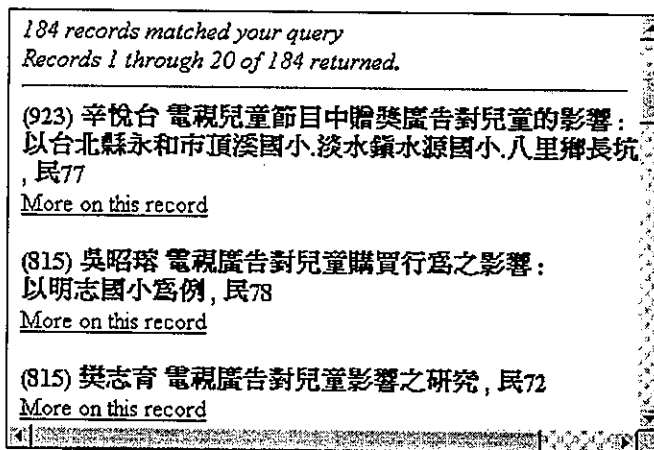
表三：國內Z39.50客戶端軟體連線網址

圖 書 館	Z39.50 的 WWW 網 址
國 家 圖 書 館	http://tulips.ntu.edu.tw : 211/z39/ncl telnet://nbnet.cl.edu.tw*
省立台中圖書館	http://search.ptl.edu.tw/z3095/ptl_c.html
中 央 研 究 院	http://las.sinica.edu.tw : 211/z39/ascin
台 大	http://tulips.ntu.edu.tw : 211/z39/ntu telnet://tulips.ntu.edu.tw*
政 大	http://tulips.ntu.edu.tw : 211/z39/nccu/innopac telnet://jenda.lib.nccu.edu.tw(user:library)*
輔 大	http://xlib.fju.edu.tw/Isite/ http://xlib.fju.edu.tw/Isite/multidb.html*

\*註：標示\*者表示可作多資料庫查詢



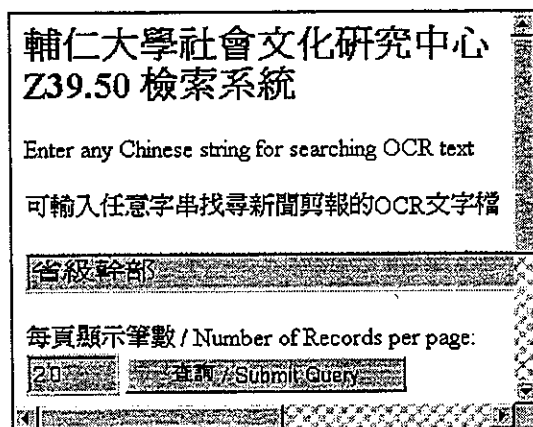
圖四：輔大Z39.50的WWW查詢介面



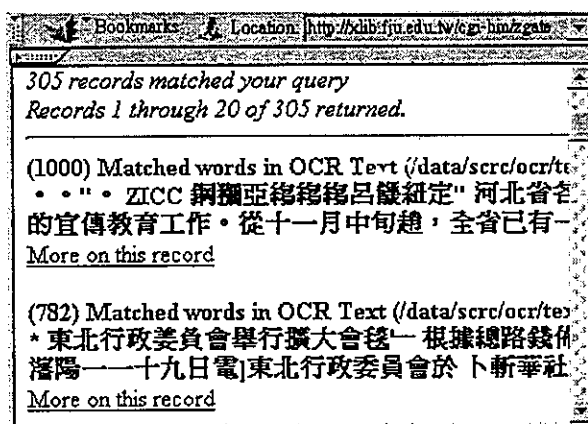
圖五：查詢範例「電視廣告對兒童的影響」檢索結果

除了文字媒體外，Z39.50也可運用於其他多媒體資料庫。圖六展示的是我們以輔仁大學中國社會文化研究中心過去四五十年來中國大陸的剪報資料，經OCR（Optical Character Recognition）文字辨識作成的全文檢索影像資料庫（註21）（註22）（註23）。同樣也是以Isite系統使其符合Z39.50檢索標準。

使用者可輸入任意查詢字串，系統以模糊搜尋比對資料庫的OCR文字，並按近似程度列出結果，如圖七所示。雖然檢索結果列出辨識錯誤的文字，然而使用者可透過Z39.50協定選取特定記錄，而得到該筆記錄完整正確的全文影像，如圖八所示。

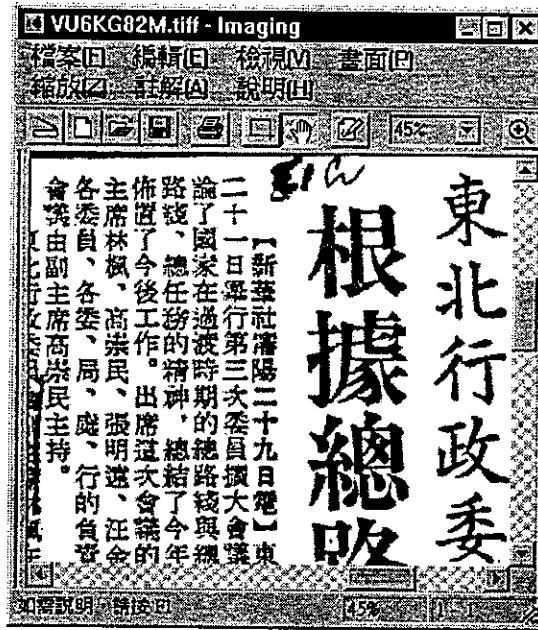


圖六：OCR文件Z39.50檢索系統



圖七：OCR文件檢索結果





圖八：檢索結果全文影像（點選圖七第二筆結果的More on this record）

天主教輔仁大學圖書館 Z39.50  
多資料庫查詢系統

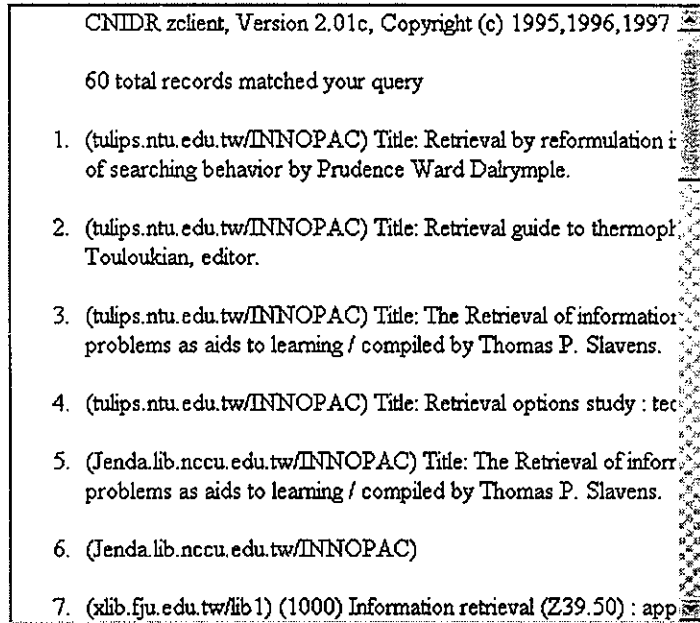
Field	Operator	Text Input
Title	contains	retrieval
Title	contains	

<input type="checkbox"/> 國家圖書館	<input type="checkbox"/> 台中省立圖書館
<input checked="" type="checkbox"/> 臺灣大學	<input checked="" type="checkbox"/> 政治大學
<input type="checkbox"/> 中央研究院	<input checked="" type="checkbox"/> 輔仁大學

Maximum Number of Responses to View: 30

圖九：多資料庫查詢介面



圖十：多資料庫查詢結果

國內的Z39.50檢索系統，除了INNOPAC系統的終端模擬Telnet介面的客戶端軟體具備多資料庫查詢功能外，其他WWW介面的客戶端軟體都無此功能。雖然Isite系統裡有此設計，但經我們測試，其多資料庫查詢功能也是在終端模擬介面下的zclient程式才有作用，其WWW的介面即不正常。為此，我們特別發展了一套CGI軟體，呼叫Isite的zclient程式，讓使用者也可以透過WWW進行多資料庫查詢。圖九為此系統的檢索介面，圖十是以「retrieval」一詞同時查詢臺大、政大與輔大的結果。從這個例子可以看出Z39.50提供的

互通性功能：雖然輔大與臺大、政大用的是不同的系統，但經Z39.50協定的連結，運用適當的軟體工具，很快就可以將他們整合在一起，讓使用者感覺到他們猶如是同一系統了。

## 伍、結論

本文介紹Z39.50資訊檢索標準協定的基本概念、運用範圍、相關計畫與實施成效，並以免費軟體Isite系統為例，說明結合Z39.50服務的OPAC系統實作過程，文中並以實例說明實作結果。雖然我們實作出可實際運用的系統，而且為國內唯一具備排序查詢的Z39.50

服務，同時能與其他圖書館的OPAC系統連結，但我們發現 Z39.50 的運用與發展還有幾處困難的地方：

一、Z39.50的相關軟體不夠成熟，不能安裝即用，說明文件不足。與迅速普及的WWW軟體比較，WWW免費的Apache伺服器佔全球裝置量五成以上，不僅各種電腦平台都有支援，而且維護簡易、安裝即用。然而開發實作 Z39.50 軟體的廠商比開發WWW軟體的廠商還多，但免費的 Z39.50 軟體不僅稀少，製作、成熟度都不精良。有的 Z39.50 系統以軟體工具方式呈現，使用者還要自行配合開發軟體。另外有的軟體如Isite系統，雖有現成的伺服器、客戶端軟體、以及通道程式（WWW gateway），但經我們將近兩年來對各個版本持續的安裝、測試，仍然不能平順的連結不同系統的資料庫，還要自行開發程式，並修改Isite的錯誤。（註24）

二、大部份Z39.50軟體大多僅提供連線、查詢、展示、結束等基本簡易的功能，尙未支援全部服務。圖書館等資訊服務單位若投資Z39.50的服務，雖可獲得SILO計畫類似的成效，但似乎沒有預期的那麼理想。尤其廠商系統的不一致性、檢索功能不及原有系統完整、軟體安裝維護問題、系統反應慢效率差等問題，是最主要的缺失。以輔大的系統為例，Z39.50的檢索

服務建置完成已一年多，但幾乎沒有使用者利用。一項服務若實際使用次數太少，應當沒有繼續存在的必要，可以刪除取消以減輕維護系統的負擔。

三、HTTP, HTML, CGI等WWW網路技術的競爭。Z39.50 協定比 HTTP 等 WWW 相關協定還早制訂，但Z39.50能夠做到的服務，WWW等技術也都能模擬做到，而且這些協定都比較簡單，很快就廣為人知，迅速普及。Z39.50 協定相較之下，則是規格嚴謹、技術詞彙艱深、複雜度高、不易實作。

四、資訊服務單位對系統的互通性(interooperability) 需求不強烈。多數單位在規畫新系統時，主要目標在服務本單位的讀者。對於與其他單位系統的互通性、或其他單位讀者檢索本單位資料庫的需求，其優先順序常落於本單位系統的檢索功能、使用者介面、資訊組織、分析與管理之後。Z39.50 的互通性功能需求不強，在推廣上便不容易，廠商或研發單位便不輕易投入，品質優良的軟體也就不易出現。其結果將造成各個資訊服務單位必須各自投入 Z39.50 服務的建置。

以上所述主要是國內 Z39.50 運用與發展上的困難，畢竟國外建置Z39.50服務的單位非常多，軟體的移植與運用也相對較為簡單。未來若繼續投入相關計畫的研究與軟體的發展，對於

資料格式複雜、需分散維護與互通性兼顧的（公共）資料庫而言，應當能發揮 Z39.50 的預期效益。過去 MARC 使書目資料能夠讓圖書館分享，現在 Z39.50 則使檢索設備（包含軟、硬體、資料庫以及檢索介面）能夠讓使用者分享。

### 註釋

- 註 1：曾元顯，「架構在 WWW 上的分散式線上公用目錄系統」，海峽兩岸圖書館事業研討會（臺北市：國家圖書館，民國 86 年 5 月 26-28 日），頁 263-277。
- 註 2：請參見清華大學圖書館的 OPAC 系統，<<http://www.lib.nthu.edu.tw/>>。
- 註 3：中正大學圖書館，<<http://lib.ccu.edu.tw/>>。
- 註 4：林大彰，「圖書虛擬聯合目錄實作」，（碩士論文，國立中正大學資訊工程研究所，民國 87 年 7 月）。
- 註 5：Juha Hakala, "Z39.50-1995: Information Retrieval Protocol-- An Introduction to the Standard and It's Usage," Automation Unit of Finish Research Libraries, Helsinki University Library, 18 Jan. 1996, <<http://renki.helsinki.fi/z3950/z3950pr.html>>。
- 註 6：John A. Kunze & R. P. C. Rodgers, "Z39.50 in a Nutshell," <<http://www.nlm.nih.gov/pubs/staffpubs/rodgers/z39.50/z39.50.html>>。
- 註 7："Guide to Z39.50/PRISE 2.0: Installation, Use, & Modification," <<http://www-nlpir.nist.gov/~over/zp2/main.html>, Jan. 7, 1997 >。
- 註 8：陳昭珍，「CNS 13461 資訊檢索服務與協定標準」，臺灣史料數位化研討會（民國 87 年 5 月 27 日）。
- 註 9："Maintenance Agency page for International Standard Z39.50," <<http://lcweb.loc.gov/z3950/agency/>> (13 Nov. 1998)。
- 註 10："M25Link Project," <<http://www.M25lib.ac.uk/M25link>>。
- 註 11："An Evaluation of Z39.50 within the SILO Project," <<http://www.silo.lib.ia.us/bluang.html>>(1996)。
- 註 12："Z39.50 Software," <<http://lcweb.loc.gov/z3950/agency/projects/software.html>>。
- 註 13："Z39.50 Client and Web Gateway Surveys," <<http://www.dstc.edu.au/RDU/reports/zreviews/>> (4 Sep. 1996)。
- 註 14："ZedKit for UNIX," <<http://www.crxnet.com/downloads.html>>。
- 註 15："YAZ Toolkit," <<http://www.indexdata.dk/>>。
- 註 16："CNIDR resources for information discovery and retrieval," <<http://www.nlm.nih.gov/pubs/staffpubs/rodgers/z39.50/z39.50.html>>。

- //www.cnidr.org/ir/ir.html>.
- 註17：曾元顯，「架構一個 WWW 上的 Z39.50 伺服器」，中國圖書館學會會訊 5 卷 1 期 (104) (民國86年年3月31日)。
- 註18：賴忠勤，「談在全球資訊網 (WWW) 以 Z39.50 協定檢索書目資料庫系統之規畫」，書苑季刊 43 期，頁 42-53。
- 註19：鄭文彬，「Z39.50 協定客戶端之研究與實作」，(碩士論文，國立中正大學資訊工程研究所，民國87年7月)。
- 註20：曾元顯，「模糊搜尋、相關詞提示與相關詞回饋在 OPAC 系統中的成效評估」，中國圖書館學會會報 61 期 (民國87年12月)，頁 103-125。
- 註21：Yuen-Hsien Tseng, "Fault-Tolerant Information Retrieval for Noisy OCR text," Proceedings of the 1st Conference on Computer Vision, Graphics, and Image Processin (Taipei, 12-14 Aug. 1998) : 445-452.
- 註22：Yuen-Hsien Tseng, "Toward a Digital Library of Newspaper Clippings," 第四屆國際資訊管理研究暨實務研討會 (臺北，民國87年11月20-21日)，頁 18-25。
- 註23：Yuen-Hsien Tseng, "An Approach to Retrieval of OCR Degraded Text," to appear in 圖書館學刊 (National Taiwan University)。
- 註24：曾元顯，「架構一個 WWW 上的 Z39.50 伺服器(二)」，中國圖書館學會會訊 5 卷 4 期 (107) (民國86年12月1日)。