

# 數位圖書館個人化檢索與推薦服務之設計與實作

## Design and Implementation of Personalized Retrieval and Recommendation on Digital Library

邱宏彬助理教授  
Hung Pin Chiu  
Assistant Professor  
[hpchiu@mail.nhu.edu.tw](mailto:hpchiu@mail.nhu.edu.tw)

湯鎰聰研究生  
Yi Cong Tang, Master Student  
[eric\\_tang630@hotmail.com](mailto:eric_tang630@hotmail.com)

陳揮明碩士  
Hui Ming Chen, Master  
[sam@mail.fjsh.cy.edu.tw](mailto:sam@mail.fjsh.cy.edu.tw)

南華大學資訊管理研究所

Department of Information Management, NAN HUA University

### 摘 要

近年來網際網路的盛行以及資訊數位化的技術逐漸成熟，改變了資訊和知識的產生、處理及傳播方式。資訊和知識的產生已邁入電子出版時代，而網際網路不受時空限制、便利、快速等特性，也使其成為人們取得和傳播資訊的主要管道。圖書館乃是將人類資訊和知識加以蒐集、整理、保存，並提供利用及傳播的場所，同樣深受於電子出版和網際網路的衝擊。數位圖書館的建置是利用資訊數位化以及網際網路的技術來達成傳統圖書館服務品質的升級。但是在此資料繁多且容易取得的環境下，資訊氾濫的結果往往會令使用者無法隨心所欲地尋找需要的資訊。因此，本研究希望藉由探討個人化技術與應用，架構「數位圖書館上個人化檢索與推薦系統」，運用使用者的個人興趣檔，來建構個人化圖書檢索介面與個人化推薦服務，協助使用者找到符合其自身需求的書籍與資料。本研究實際以智慧型 GAIS 搜尋引擎[6]基礎來實作此個人化搜索與推薦服務，以驗證其可行性，本研究結果可為數位圖書館智慧型服務建立進一步研究的基石。

關鍵詞：數位圖書館、搜尋引擎、個人化服務、推薦系統。

## Abstract

The information digitalization and internet technology can be used to construct a digital library to upgrade the service qualities of traditional library. However, the data are diverse and growing rapidly so that the users can't effectively retrieve their desired information under such digital environment. Therefore, designing intelligent user interface and search engine to guide the users for searching on digit library are the very important issues.

This paper proposes a personalized retrieval and recommendation mechanism to enable the intelligent searching on digital library. By using user's individual interest profile, the personalized books search interface and recommendation services are implemented to assist users to effectively find the books that they really want.

**Keywords:** Digital library、Search Engines、Personalized Service、Recommendation System

### 壹、導論

近年來網際網路的盛行以及資訊數位化的技術逐漸成熟，改變了資訊和知識的產生、處理及傳播方式。資訊和知識的產生已邁入電子出版時代，而網際網路不受時空限制、便利、快速等特性，也使其成為人們取得和傳播資訊的主要管道。因為電子出版的成熟以及網際網路的盛行，使得資訊得以快速生產和傳播，而網路上的資訊也正以驚人的速度成長，使得使用者面臨了資訊超載 (Information Overload) 的問題日趨嚴重。因此，網際網路資訊服務業者，不論是強調資訊取得廣度、指引網路初學者的入口網站 (portal) 如 Yahoo；或是著眼於專業資訊，強調資訊取得深度的垂直網站 (vertical portal) 如 Amazon、CNN 等，皆開始提供個人化服務[17]，為使用者解決資訊超載的問題[9][13][16]。

圖書館乃是將人類資訊和知識加以蒐集、整理、保存，並提供利用及傳播的場所，同樣深受於電子出版和網際網路的衝擊。圖

書館的發展，亦從傳統的實體圖書館到複合式圖書館或數位圖書館。圖書館的館藏與服務逐漸朝向電子化發展，以因應網際網路及電子出版時代的來臨。數位圖書館 (Digital Library) 的建置是利用資訊數位化以及網際網路的技術來達成傳統圖書館服務品質的升級。隨著數位圖書館的日漸成熟，使用者在取用或搜尋這些資訊資源時，亦會產生資源眾多或搜尋結果過多，所需資源不易尋找的情形，使用者同樣面臨了資訊超載的問題[9][13]。

資訊提供者主義 (Provider-ism, PVRM) [1] 和使用者主義 (User-ism, USRM) [1] 這兩種觀念之下，將使用者主義的觀念落實 在數位圖書館的做法是化被動為主動的貼心服務[10]。以圖書館使用者角度來看，PVRM 有四個特性：永久 (Permanent)、多樣 (Versatile)、大量 (Massive)、即時 (Real-Time)；而 USRM 有下列特性：聯合 (Union)、語意 (Semantic)、推薦 (Recommendable)、量身訂做

(Personal-Made)。因此如何改善圖書館的服務內容，並且進一步地替使用者解決資訊超載的問題，成為數位圖書館發展的下一目標。

圖書館的一切作為都是為了提供使用者良好的服務，協助使用者獲取資訊、運用資訊，從而產生知識[7]。但是在此資料繁多且容易取得的環境下，資訊氾濫的結果往往會令使用者無法隨心所欲地尋找需要的資訊，因此，探究如何讓使用者能正確且快速地取得所需資料，並發展一個以人為中心之智慧型使用者介面是當今數位圖書館發展中的重要議題。

目前大部分圖書館只提供「個人化資訊選粹服務」[1]，將資訊簡化並提供給使用者使用，節省使用者時間。但部分圖書館網站所提供的個人化資訊選粹服務尚未達到為使用者過濾資訊之功能。以目前的作法來看，仍然停留在以前的大量行銷的方式(mass marketing)[11]。此外，其他個人化服務功能，如個人化檢索，交通大學PIE@NCTU就有提供部分功能[13]。相較於目前熱門的網際網路上的書局，例如：亞馬遜書局(www.amazon.com)在網路上就有以下的服務：它能針對不同的使用者，提供不同的新書籍推薦。對於使用者已訂的書，能線上推薦購買該書的相關購買資訊，如先前購買者中同時購買那幾本書。所以使用者可以得到個人化的資訊服務，同時也增加使用者購買書籍的動機，提高書局的收入。

有鑑於此，本研究將既有的圖書館查詢

架構，加入個人化服務與推薦技術，採因人而異的個人化介面，提供更多元化的資訊加值服務，以提高資訊服務品質及使用者滿意度。本研究希望藉由探討個人化技術與應用，架構出「數位圖書館上個人化檢索與推薦系統」，運用使用者的檢索紀錄檔，來建構個人化圖書檢索介面與個人化推薦服務，協助使用者找到符合自身需求的書籍與資料。綜合上述，本研究將提供適用於數位圖書館之個人化檢索與推薦服務模組，實作一個數位圖書館之個人化檢索與推薦服務系統，最後再經由此系統提供更符合個人化要求的查詢結果與相關館藏推薦。

本文第貳節將說明研究方法與相關文獻探討，包括圖書館網站個人化服務、資訊檢索與資訊過濾相關技術、推薦系統技術、中國圖書分類法；接下來，第參節將描述本研究之系統架構設計與運作方式；在第四節將實作並探討本研究架構；最後是結論與未來展望。

## 貳、研究方法與相關文獻探討

本研究將收集資訊檢索、個人化服務以及推薦服務等相關文獻，以其理論基礎提出個人化服務應用在數位圖書館的模型，並加以分析探討所遇到的問題，且針對推薦系統的使用者提出一個依使用者個人興趣檔分群的方法，最後實際應用個人化技術於維護使用者個人興趣檔、產生個人化查詢及處理個人化推薦，並驗證本研究的可行性及應用性。

### 一、數位圖書館個人化服務

在服務業中，傳統的個人化是指服務人員與顧客間的社會性互動，程度上可能從冷漠且缺乏個人化，到貼心而個人化[17]。至於網際網路上的個人化服務至今尚無一致的定義。一般而言，網站個人化是指針對每一個使用者特殊的需求設計更具彈性的網站。個人化服務方式最著名的即為 Amazon 網路書店“new for you”的功能。當使用者遨遊於 Amazon 網路書店時，系統會自動紀錄個人的行程，並利用使用者的電子交易行為，歸納出個人的閱讀興趣，當使用者下次進入 Amazon 網路書店時，系統會提供個人化的書籍推薦。

根據 Jupiter Communications 對前 25 個線上零售商所做的調查，有 40% 的網站表示他們已經開始提供個人化的服務，若加上未來一年內會採用個人化技術的網站則高達 93% [27]。Dell Online 的資深經理 Bill Morris 更表示在一至二年內，個人化將是網際網路的主流[21]。個人化在網際網路發展策略中的重要性由此可見。

數位圖書館乃是擁有相關資源(包含軟體設備、網路、專業人士...等)以執行下列任務的機構：對數位形式的館藏進行挑選、組織、提供使用、解譯、傳播、保持完整性、長期保存等工作，並提供這些數位形式的館藏能為特定使用者快速且經濟地運用。

從這個定義上來看，電子圖書館與傳統圖書館的任務與功能並無不同，只不過電子

圖書館所處理的是以數位形式的資料為主。根據以上定義，Cleveland 提出了電子圖書館的幾項特性，其中有一項為：數位圖書館是傳統圖書館的另一種面貌，其館藏包含電子式(與數位式)以及印刷和其他(例如，影帶和聲音)媒材。圖書館網站雖然提供了電子化資源和服務來便利使用者使用，但使用者亦面臨了資訊超載的問題[9][13]。

在使用者面對眾多電子資源的情況下，不知如何選擇也不易尋找符合自己所需的資訊的使用者，在檢索時常常面臨檢索結果過多的情形，檢索功能和介面亦缺乏彈性無法幫助使用者真正檢索到所需要的資訊。鑑此，圖書館著重於改善資訊檢索的能力、提供使用者個別所需的資源，並落實與擴展「專題選粹服務(SDI)」[8]等相關構想，而開始發展圖書館網站的個人化服務，其中尤以北卡州立大學(North Carolina State University)的“MyLibrary@NCState”最為著名。MyLibrary@NCStat 提供使用者一個人化資訊環境，讓使用者自行設定所需的電子資源與服務內容；並提供個人化的新知通告服務。使用者依圖書分類號，輸入需要的分類號範圍，系統便會定期寄送 E-mail，通知使用者符合需求的新進館藏資源，包括：新書、期刊、線上資料庫、網路資源等。此外，系統也提供了“Message from the Librarian”的功能，由符合使用者學科類別的館員，定期更新內容，包括：最新消息之通知、建議，以協助使用者在該學科

領域中保持適當的敏銳度與新穎性[24][13]。圖書館同樣希望藉由圖書館網站提供主動、長期且個人化的資訊服務，來解決使用者資訊超載的問題，幫助使用者更容易獲取其所需的資訊，並維持使用者對於新知的敏感度；另一方面亦可進一步提昇圖書館的服務品質與形象[2]。

因此，個人化服務已成為圖書館網站設計發展及使用者服務上的焦點，美國圖書館與資訊技術協會（Library and Information Technology Association, LITA）在 1999 年的年終會議所討論出的圖書館未來七大發展趨勢中，將圖書館網站個人化服務視為首要趨勢[25]。由此可見，個人化服務已成為圖書館網站未來的發展趨勢。

圖書館網站個人化服務的意義，簡單地說，是提供使用者一個人化的圖書館網站，一個以使用者為中心，可以客製化所蒐集的圖書館資源之界面[24]。也就是讓使用者依照自己的個人需求來客製化數位圖書館，藉由這種方式可以幫助使用者經由單一的存取管道而有效率地發現、組織、以及最佳化地管理圖書館的資源和服務。簡而言之，這些服務可以幫助使用者根據個人的需求而過濾和檢索資訊，反映個人的興趣或研究的需要[21]。

有關圖書館網站如何實施個人化服務，其整體內涵，提供哪些功能，有不少學者提出相關的理念與構想[3][4][5][9][11][12]，大部分皆從個人化資訊環境或個人資訊空間來切入，並論及個人化

資訊環境的構成要件以及可提供的服務項目或功能。茲將國內有關數位圖書館個人化服務的研究及研究的功能彙整如下表 1。

表 1、數位圖書館個人化服務相關研究

國內相關研究				
研究主題 研究者與年代	個人資訊 選粹服務	個人化 瀏覽	個人化 檢索	個人化 推薦
王姿瑁, 民 89[3]	✓			
孫冠華, 民 89[11]				✓
林俊佑, 民 90[5]			✓	
楊雅雯, 民 90[9]	✓	✓	✓	
鄭玉玲, 民 92[12]		✓	✓	✓
余明哲, 民 92[4]				✓

## 二、資訊檢索與資訊過濾

### （一）資訊檢索

資訊檢索(Information Retrieval), 早在 1952 年就出現這個名詞，並從 1961 年開始被廣泛使用，一直流行到現在[23]。其目的在利用資訊科技協助儲存、管理、搜尋大量資料，以方便檢索。資訊檢索系統的主要功能如下[18]：

1. 分析文件資源的內容。
2. 辨別相關資訊，加以整理、分類。
3. 將文件內容分析結果加以儲存，以便利檢索。
4. 分析使用者查詢。
5. 將使用者的查詢比對檢索資料庫。
6. 找出相關資料，展示給使用者。
7. 根據使用者回饋進一步調整檢索系統。

資訊檢索系統依檢索方式的不同可分為全文檢索系統 ( Full-text Information Retrieval System ) 和指標式檢索系統 ( Indicative-text Information Retrieval System )。全文檢索允許檢索文件內任何字串，而指標式檢索只能針對某些特定的關鍵語句進行檢索。雖然資訊檢索系統所提供檢索的方式不同，但系統架構大致相同。

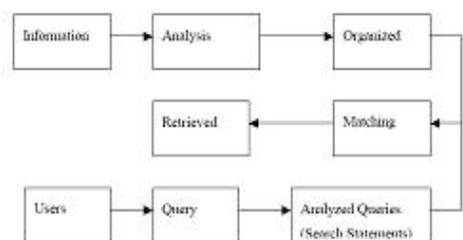


圖 1、資訊檢索系統架構

如圖 1[20]，資訊檢索主要是基於使用者的查詢條件，從較大的文件集中挑選出相關的文件子集合，以找出符合的資訊。使用者可以瀏覽系統所回應的文件集合，從中找出真正想要的資訊。

GAIS ( Global Area Information Servers ) 是一個智慧型多國語言全文檢索系統，她擁有一個彈性的搜尋及索引引擎，使得 GAIS 可以適用於各種不同領域的全文檢索系統當中[6]。GAIS 搜尋引擎 ( Search Engines ) 主要包含兩個部份：Gais 主要應用於智慧型全文檢索作業；Gaisindex 主要應用於索引資料檔建立作業。本研究中，乃是利用 gais 直接對檔案進行全文檢索的方式。未來資料

量增加時，可再加入支援 gaisindex 的功能，加快檢索的速度。GAIS 搜尋引擎具有特色如下：

1. 提供使用者一個彈性、可自定的搜尋策略與輸出結果的能力。
2. 支援功能強大的查詢語言。

## (二) 資訊過濾

資訊過濾技術 ( Information Filtering ) 是一種能從網路即時產生的大量資訊中，針對特定主題過濾出重要訊息的資訊檢索技術，與現行資源搜尋系統所提供的被動式 ( Passive ) 資訊檢索技術和一般化資訊服務 ( General Information Service ) 不同，資訊過濾強調的是主動式 ( Active ) 與個人化的資訊服務 ( Personalized Information Service ) [3]。

資訊過濾技術一般可以區分為內容式資訊過濾 ( Content-based Information Filtering, 或稱為 Feature-based Information Filtering ) 以及協力式資訊過濾 ( Collaborative Information Filtering ) 兩類 [1]。內容式資訊過濾通常利用關鍵字詞比對方式，對文件內容 ( Content ) 加以分析比對，著名的商用系統如 Verity、CompassWare、IBM InfoSage 及 InfoSeek Personalized News 等皆是採用此類技術。至於協力式資訊過濾技術，則不直接分析文件內容，改以分析使用者的個人背景、知識、興趣等，先找出興趣接近的同好，之後針對使用者的查詢主題，並從這些同好所感興趣的資訊中，再分析選取最可能相關的資

訊提供參考。相關服務包括資訊選粹服務 ( Selective Dissemination of Information , SDI )、個人化推薦 ( Personalized Recommendation )以及個人化檢索( Personal Search ) [12]。本論文所著重在於個人化推薦及個人化檢索兩項服務。

### (三) 資訊檢索與資訊過濾比較

資訊檢索是根據使用者的查詢，從大量資料中檢索出相關文件。而資訊過濾則是過濾掉不相關的文件。資訊過濾技術強調的是讓網路資訊檢索具備更主動 ( Active )、長

期 ( Long-Term ) 及個人化 ( Personalized ) 的資訊服務能力，藉以反應出傳統資訊檢索系統中的被動 ( Passive )、短期 ( Short-Term ) 及一般化 ( General ) 等資訊服務缺失，這之中又以個人化服務差異最大。因為具有資訊過濾能力的系統，會建立個人興趣檔 ( Profile )，並針對不同使用者提供不同的服務內容，當有符合的資訊被檢索出時，系統就會主動的寄送 Email 通知使用者[3]。圖 2 可充分說明資訊過濾和資訊檢索的觀念。

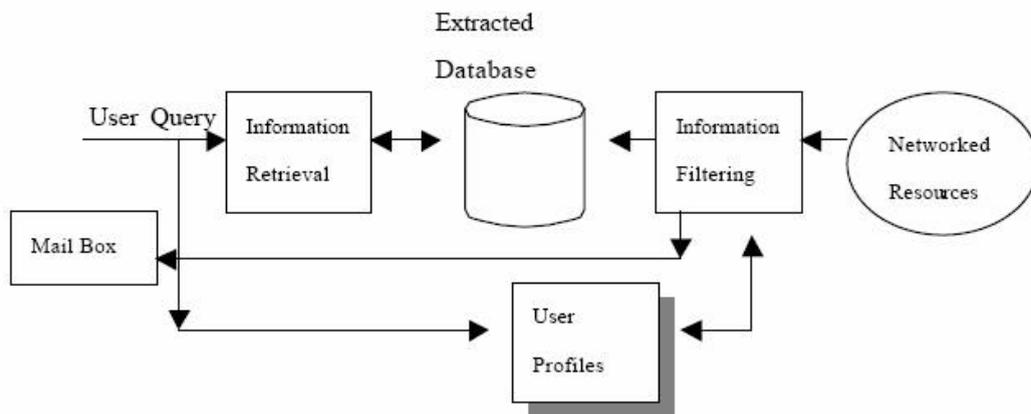


圖 2、資訊過濾和資訊檢索觀念圖

### 三、推薦系統

當資料量變得龐大且分散，使用者利用傳統關鍵字檢索搜尋資料的方式，變得相當費時且不容易聚焦。因此，為了要降低使用者在網路上的搜尋時間，提供使用者更適切的資訊，自動化推薦系統 ( Automatic Recommender System ) 便因應而生。自動化

推薦系統有許多不同的推薦方式，如 Collaborative Filtering 和 Content-Based Filtering 等方式[1]。

協力式過濾 ( Collaborative Filtering ) 的方法是一種最典型的推薦方法，自從 1992 年由 Goldberg 等人 [19] 提出後就被廣泛採用在一般的推薦系統中，甚至在

1997 年 Resnick 和 Varian 提出推薦系統 ( Recommender System ) [29] 這個名詞之前，協力式過濾就是推薦系統的代名詞。協力式過濾主要的概念是使用者利用同儕對某物品既有的評價來預期自己對該物品可能的評價。這個想法就好像當我們要做某樣決定之前，會事先詢問一些有經驗的人，以他們的意見做為決定的參考。

推薦系統目前應用最廣的領域在於電子商務上，主要用來提升消費者回店消費的機率以增加營業額。網路商店和實體商店有一個重要的不同是在於網路商店沒有店員，這固然減少了人事上的費用，可是當顧客上門時沒有店員幫忙介紹商品、推銷商品、回答顧客的問題和給與顧客建議，商店和顧客之間沒有互動，無形中就減少了很多交易的機會。面對這個問題，其中一個解決的方法就是建立一套具個人化的推薦系統，個人化的推薦系統可以取代店員部分的功能，其中最重要的功能就是推銷商品、給予顧客購買上的建議，因此推薦系統逐漸被應用在網路商店上，幫助網路商店的經營。

關於推薦方式，指的是如何找出顧客感興趣的商品並列出推薦清單。推薦方法是推薦系統的核心，方法的好壞直接影響了推薦結果的優劣。因為需要不同而發展出多種不同的推薦方法，常利用的方法有基本檢索 ( Raw Retrieval ) [5]、人工選粹 ( Manually Selected )、統計分析 ( Statistical Summaries )、屬性導向 ( Attribute-Based ) [3][12][28]、商品間的關聯性 ( Item- to-Item

Correlation ) [31]、顧客間的關聯性 ( User-to-User Correlation ) [12][28] [31]。這些方法除了可以獨立運作外，也可以互相搭配使用來提升推薦的效果。目前推薦系統在電子商務上的應用，比較偏重於個人化的功能，希望能根據每個使用者不同的興趣分別給予使用者符合興趣的推薦。個人化推薦系統主要的方法分成兩種，分別為協力式過濾和內容導向過濾。在之前提到的方法中，商品間的關聯性和使用者間的關聯性是協力式過濾推薦常用到的方法，而屬性導向推薦其實就是內容導向過濾推薦的別名。

推薦系統的概念應用在圖書館上，希望使用者之間能互相分享資訊，每個使用者都是過濾代理人 ( Filtering Agent )，可以幫忙篩選館藏，提供別人借閱書籍意見參考。而圖書館中每個使用者的借閱歷史可視為是給別人的館藏推薦清單，使用者可以參考這些推薦清單，來挑選自己有興趣的館藏，而不用從圖書館眾多的館藏中盲目地尋找。這麼多的推薦清單中，其實並不是每一份都是適合自己需要的，因為每個人的興趣或目的不同，所以會借閱不同的館藏，分別產生不同的借閱歷史和推薦清單，其中可能只有某部分清單是真正符合使用者需要的。因此推薦系統還必須事先幫使用者決定那些推薦清單是使用者需要或感興趣的，而不需要使用者每次都從上千份推薦清單中挑選。

各式各樣的推薦系統為因應各種不同的需要而產生。目前推薦系統新的研究方法是加入資料探勘的技術，其中最常用的是以

資料探勘中關聯規則探勘 ( Association Mining ) 的方法找出交易集 ( Transaction Set ) 中隱含的知識 ( Knowledge )、有用的資訊 ( Useful Information ) 和關聯規則 ( Association Rules )，例如使用者與使用者間的關係、書籍與書籍間的關係或使用者與書籍間的關係等[4][11][12]。

#### 四、其他相關技術與文獻

本研究中使用的其他相關技術或文獻，共分為三部分。第一，中國圖書分類法 [15]，於本研究中的應用是可將個人檢索後的結果依中國圖書分類法的主題分門別類；第二，相似度衡量的方法，在本研究中應用於個人興趣檔相似度的比對；第三，SCSA ( Simple Cluster Seeking Algorithm ) 演算法，在本研究中應用於將使用者依個人興趣紀錄檔的特徵聚類分群。

##### (一) 中國圖書分類法

「中國圖書分類法」的優點在於使用年代久遠，從民國初年使用迄今，廣泛應用，且分類主題詳盡、明瞭，讓人容易一眼就明白，並隨著時代不同而持續更新，不會流於過時的窠臼之中。國內現在約有百分之八十以上的圖書館在使用這種方法，本研究就地取材，依現有資源的分類規劃，加強圖書館個人化服務。

中國圖書分類法的分類方法，是採階層式分布，依種類分為十大類，每一大類以一個百位數整數代表，如 400 代表「應用科學類」，下一個十進位代表更細的分類如

410 代表醫藥、420 代表家事、...，490 代表商學等十個應用科學的分類；以 410 來說，下一個十進位又以醫學為主再細分為十類，411 代表衛生學、412 代表公共衛生，以此類推下去，階層式的表示法。在圖書館的編制當中，每一個中國圖書分類編號都會有一些書籍隸屬之。表 2 是說明圖書分類方法的例子，412 代表「公共衛生類」，則 412.4 代表「公共衛生類」中的「防疫類」，412.1、412.2、412.3、...，412.9 等子類別都是歸屬於「公共衛生類」之下。

表 2、中國圖書分類細目說明表

412	為「公共衛生」
412.4	為「公共衛生」中的「防疫」
412.48	為「公共衛生」中的「防疫」中的「其它傳染病預法」
412.481	為「公共衛生」中的「防疫」中的「其它傳染病預法」中的「隔離」
412.483	為「公共衛生」中的「防疫」中的「其它傳染病預法」中的「消毒」

##### (二) 相似度衡量

本研究中個人興趣檔是由許多型樣 ( Patterns ) 所組成，與一般量化 ( quantize ) 相似度衡量[26]方式不太相同，所以在此採用 Anandanpilla [30] 提出相似度衡量的方法。此方法是依照新舊案例 ( Case ) 間型樣 ( Patterns ) 的重疊程度 ( Overlap ) 來比對兩者相似程度，如公式(1)所示。

$$M(C_1, C_2) = n(P_0) - n(P_1) - n(P_2) \quad (1)$$

其中  $M$  代表兩個 Cases ( $C_1, C_2$ ) 的相似程度,  $P_0$  為  $C_1$  和  $C_2$  重疊的 patterns,  $P_1$  為在  $C_1$  而不在  $C_2$  的 patterns,  $P_2$  為在  $C_2$  而不在  $C_1$  的 patterns,  $n(P_i)$  為  $P_i$  的 patterns 數量。舉例來說: 假設有三個 cases 和一個 New-case, 如表 3 所示[30], 表 4 為 New-case 與三個 cases 比對重疊程度後計算出的相似程度 (Measure), 此範例中的結果可得知 New-case 與 case3 最相似。

表 3、Cases 內含的 patterns

Case Number	Patterns
1	$p1, p2, p3, p4$
2	$p1, p4, p6$
3	$p2, p4, p5, p6$
New-case	$p2, p3, p5, p6$

表 4、Cases 中 Patterns 相似度範例

Cases	$P_0$	$P_1$	$P_2$	Measure
New-case,1	$p2, p3$	$p5, p6$	$p1, p4$	-2
New-case,2	$p6$	$p2, p3, p5$	$p1, p4$	-4
New-case,3	$p2, p5, p6$	$p3$	$p4$	1

### (三) SCSA 演算法

群聚 (clustering) 是把有形或抽象的物件歸類到類似物件類別的過程; 將類似物件集成同群, 不同群物件的集合不相似, 群聚與分類 (Classification) 最大不同是, 群聚不預先知道類別標籤, 而把資料歸類成新類別[22]。例如它可透過數學方法來尋找空間物件的相似性, 而分析最終目的是將資料進行分類的工作。

群聚法廣泛地應用在許多領域, 例如模式識別, 資料分析, 和圖像處理。就市場分析領域來看, 分群基於顧客購買模式 [22]。在本研究中使用 SCSA (Simple Cluster Seeking Algorithm) 演算法[12], 將使用者依個人興趣紀錄檔的特徵聚類分群。

SCSA 演算法中, 主要的精神是以樣本之間的相似度做分群的依據, 當相似程度超過所定義的相似度門檻值 (Similar threshold) 時, 則歸類為同一分群, 以和目前 case 最相近的做為群聚重心, 本研究用公式(1)來度量相似度, 參閱圖 3 的演算法。以表 3、表 4 為例: 若將相似度門檻值定為 0, 則相似程度大於 0 歸為同一群; 相似程度小於 0 為不同群。所以表 3 中, New-case 和 case3 為同一群 ( $M(\text{New-case}, 3) = 1$ )。

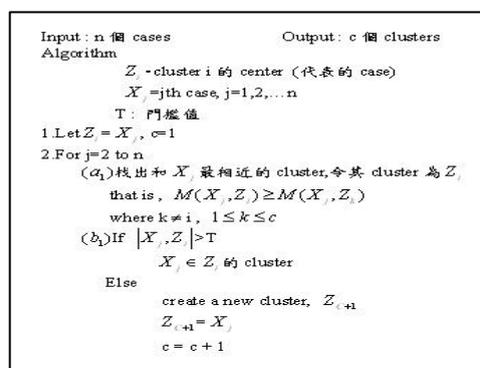


圖 3、SCSA 演算法

## 參、系統架構

### 一、個人化檢索與推薦服務系統架構

數位圖書館上個人化檢索與推薦服務系統主要分成三個模組進行: 改良式檢索模

組、個人化檢索模組和個人化推薦模組。此系統建構個人化圖書檢索介面與個人化推薦服務，協助使用者找到符合自身需求的書籍與資料，盡可能達到個人化環境的目標。系統整體架構如圖 4 所示。

系統各模組簡介如下：

(一) 改良式檢索模組：以智慧型 GAIS 搜尋引擎為基礎，快速建構現有圖書館的查詢功能，以提供更具彈性的檢索功能和查詢結果。

(二) 個人化檢索模組：利用使用者興趣檔幫助檢索，提供更符合個人化要求的查詢結果。

(三) 個人化推薦模組：利用使用者興趣檔，探勘出興趣相近的使用者，並運用群聚分群方法加速搜尋的腳步，透過此機制為使用者推薦適當的興趣檔關鍵字及書籍，盡可能達到個人化環境的目標。

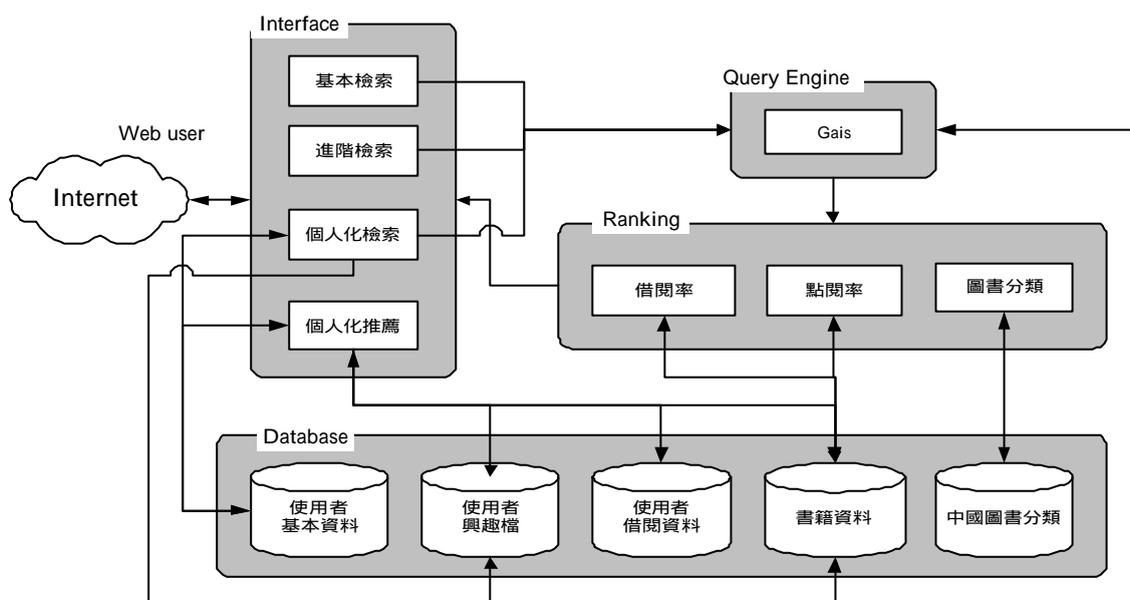


圖 4、系統整體架構圖

## 二、改良式檢索模組

目前一般的數位圖書館具有以下問題：操作界面統一、失去彈性，搜尋結果繁多導致使用者無法尋找及過濾所需的資訊

等問題[9]。在本研究中提出以智慧型 GAIS 搜尋引擎為基礎，快速建構現有圖書館的查詢功能，以提供使用者一個彈性、可自定的

搜尋策略與輸出結果的能力。在本研究中稱此模組為改良式檢索模組 ( Improved Retrieval Module, IRM )。

IRM( 參見圖 5 ) 中包含兩個檢索介面：

基本檢索及進階檢索。基本檢索提供了關鍵字同音及容錯檢索，檢索結果可以由使用者自訂呈現排名方式，有中國圖書分類法的分類主題、借閱率和點閱率。透過使用者自訂

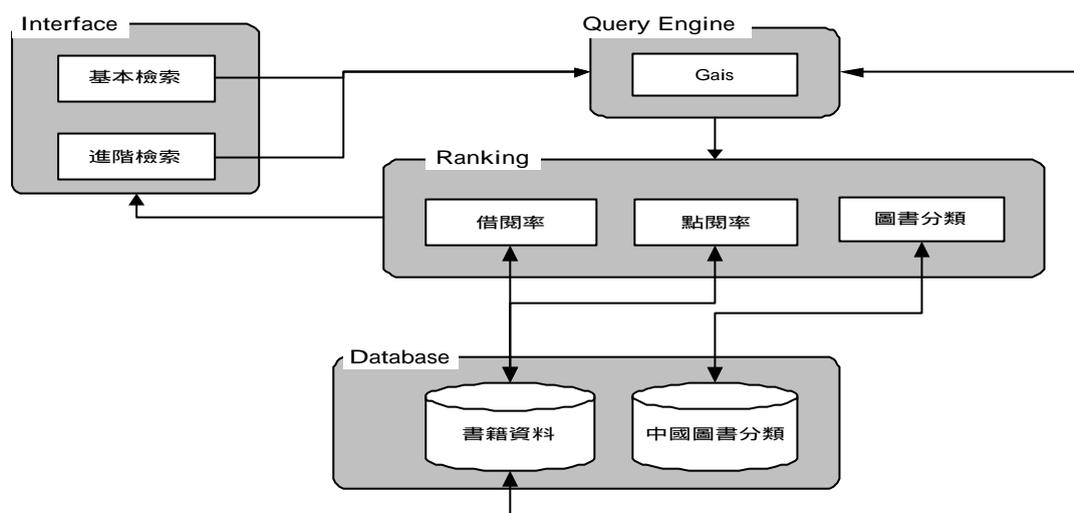


圖 5、改良式檢索模組

排名更能符合使用者需求。而進階檢索則是使用者可輸入多個查詢字串，並使用布林運算式 ( AND、OR、NOT ) 來自訂的搜尋策略。而排名方式和基本檢索相同，有中國圖書分類法的分類主題、借閱率和點閱率。

IRM 運作方式與步驟如下：

- (一) IRM 事先從書籍 ( Book ) 資料庫建立書籍索引檔 ( Book index )，提供給 GAIS 搜尋引擎檢索用。書籍索引檔範例如表 5 所示。
- (二) 使用者透過網站輸入檢索字串 ( Keywords ) 給 GAIS 搜尋引擎，

GAIS 根據檢索字串和檢索設定 ( QueryOption ) 對書籍索引檔作搜尋。舉例說明：當使用者檢索「computar」選取不分大小寫且容錯一個字元，則提供給 GAIS 搜尋引擎：Keywords= computar，QueryOption= -i -1( 其中-i 代表不分大小寫，-1 為容錯一個字元 )。

登錄號	書名	借閱次數	點閱次數	分類號
00194595	JAVA 2	8	15	312.932J3

表 5、書籍索引檔 ( Book index ) 範例

(三) 檢索到的資料存成暫存檔 (Search Results Temp) 以步驟 (二) 為例, 可以檢索到書名有 computable, computer, computational 等字串的書籍。

(四) 讀取暫存檔 (Search Results Temp), 透過排名 (Ranking) 機制, 將檢索結果呈現給使用者。IRM 運作過程如圖 6 所示。其中虛線部分代表可針對檢索結果, 再次檢索。

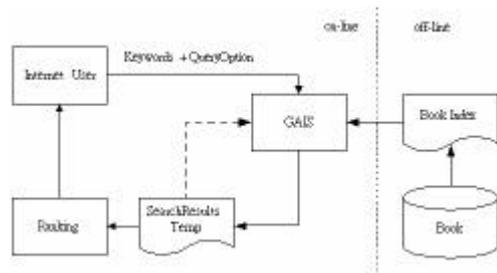


圖 6、改良式檢索模組運作圖

檢索結果排名如下：

(一) 中國圖書分類

以中國圖書分類法的分類階層為主, 將檢索到的資料依中國圖書分類法的主題分門別類, 讓使用者進階點選, 再將屬於此主題的書籍顯示出來。

(二) 借閱率

將檢索到的資料經過熱門借閱率的排序, 呈現給使用者。

(三) 點閱率

系統會記錄每本書被檢索到, 並被點閱書籍相關資料的次數, 即為書籍點閱率。此模組可以將檢索到的資料經過書籍點閱率的排序, 呈現給使用者。

### 三、個人化檢索模組

此模組會收集使用者搜尋時所用的關鍵字來判斷使用者的興趣, 並依據使用者使用之關鍵字產生系統關鍵字來代表使用者的興趣。另外當使用者查詢館藏資料時, 使用者可以選擇是否加入個人興趣檔中的關鍵字。系統會利用使用者查詢關鍵字和使用者個人興趣檔中的關鍵字, 來判斷使用者最需要的館藏, 把查詢結果先加以過濾和排序後再呈現給使用者。個人化檢索模組架構圖參見圖 7。其中排名方式和 IRM 相同, 有中國圖書分類法的分類主題、借閱率和點閱率。個人化檢索模組運作方式與 IRM 相同, 如圖 8 所示。

在此模組中, 另外一個重要的功能是個人興趣檔的建立與維護, 包含以下功能：

- (一) 收集使用者檢索時所用的關鍵字。
- (二) 使用者可以自己增加關鍵字或刪除不要的關鍵字。
- (三) 系統自動推薦關鍵字。

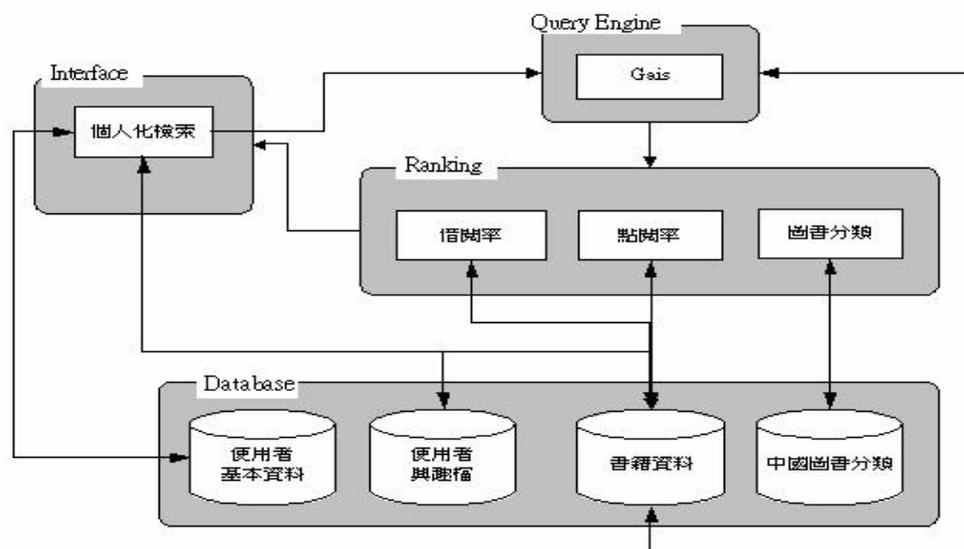


圖 7、個人化檢索模組架構圖

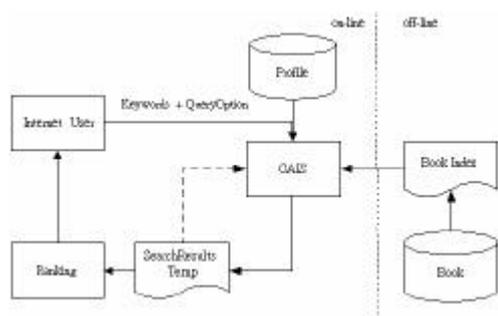


圖 8、個人化檢索模組運作圖

圖書館有新進館藏時，系統可以根據使用者的個人興趣檔來判斷哪些新進館藏是使用者可能會感興趣的，並將這些新進館藏推薦給使用者。此模組可提供關鍵字推薦、書籍

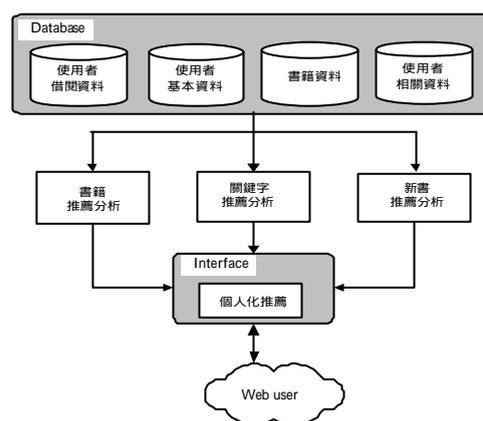


圖 9、個人化推薦模組架構圖

#### 四、個人化推薦模組

此模組（參見圖 9）主要是先分析個人興趣檔，挑選和使用者有相同興趣的其他使用者，稱之為“同好” (Friends)。同好間因為個人興趣檔類似，所以就將他們個人興趣檔的關鍵字和借閱紀錄，推薦給其他同好的使用者。另外一個重要的功能是新書推薦，當

推薦、新書推薦。我們分別將這些推薦說明如下：

#### (一) 關鍵字推薦

有許多不同的方法可以達到推薦的目的。在本研究中使用同好間喜好心理 (like minds) 的推薦方法 [15]。找出和登入使用者至少有一個相同關鍵字的其他使用者，因為他們的關鍵字很有可能也是這位使用者所感興趣的。以下說明同好關鍵字推薦過程：

1. 使用者相同关键字的同好使用者清單，首先從資料庫中執行以下 SQL 查詢：

```
select distinct(k2.username)
from keyword k1, keyword k2
where k1.username='$valid_user'
and k1.username != k2.username
and k1.kw_WORD = k2.kw_WORD
```

這個查詢中，k1 是用來存放目前這位使用者興趣檔中的關鍵字，而 k2 是用來存放其它使用者的關鍵字。現在要尋找其他使用者 (k2.username) 和目前使用者相同的關鍵字 (k1.kw\_WORD = k2.kw\_WORD)，並且那個人不是本身自己 (k1.username != k2.username)。

這個查詢會產生至少有一個關鍵字和目前使用者關鍵字相同的同好使用者清單，其中 \$valid\_user 代表目前使用者。以表 6 為例，使用者 1 和 3 至少有一個關鍵字和 \$valid\_user 的關鍵字相同。

表 6、個人興趣檔簡例

Username	keyword (個人興趣檔關鍵字)
1	Data Mining, Fuzzy, Visual Base, Java
2	Data Mining, Java
3	Fuzzy, Java, Neural, Database
\$valid_user	Fuzzy, Visual Base, Neural

2. 產生推薦關鍵字清單，我們可以使用以下的查詢找出其他的關鍵字：

```
select km_WORD
from keyword
where username in $sim_users
and km_WORD not in $user_word
group by km_WORD
having count(km_WORD) > $popularity
```

這個查詢中，變數 \$sim\_users 包含了那些有相同喜好的使用者清單，而變數 \$user\_word 則包含了目前使用者的關鍵字清單。如果 k1 已經有了我們要加入的關鍵字，則不需要推薦給使用者。最後設定一個門檻值 \$popularity 變數來做過濾，因為我們不需要推薦那些過於個人喜好的關鍵字，我們只需要推薦那些有著一定數量使用者都喜歡的關鍵字，以表 6 為例，使用者 1 和 3 推薦給 \$valid\_user 的同好關鍵字有 Data Mining, Java, Database。假設門檻值 \$popularity=1，則同好的關鍵字中超過有一個同好使用者相同擁有才被推薦。而推薦的

同好關鍵字中，其他同好使用者沒有 Data Mining, Database，所以被過濾掉。最後推薦給\$valid\_user 關鍵字只有 Java。

### (二) 書籍推薦

與關鍵字推薦方法類似，不同之處是更改成推薦書籍。找出和登入使用者至少有一個相同關鍵字的其他使用者，因為他們的其  
他書籍很有可能也是這位使用者所感興趣的。先找出和目前使用者相同關鍵字的同好使用者清單，再來產生推薦書籍清單。

### (三) 新書推薦

當圖書館有新進館藏時，系統可以根據使用者的個人興趣檔來判斷哪些新進館藏是使用者可能會感興趣的，並將這些新進館藏推薦給使用者。以下說明步驟及運作方式：

1. 新書推薦模組事先從書籍資料庫建立新書索引檔 (New Book index)，提供給 GAIS 搜尋引擎檢索用。書籍索引檔範例如表 5 所示。
2. 當使用者登入時，會將個人興趣檔中關鍵字作為檢索字串 (Keywords) 給 GAIS 搜尋引擎，GAIS 根據檢索字串 (Keywords) 對新書索引檔作搜尋。舉例說明：如果個人興趣檔中的關鍵字是 (Data Mining, Fuzzy, Java)，則提供給 GAIS 搜尋引擎：Keywords= Data Mining, Fuzzy, Java。
3. 將檢索到的資料存成暫存檔 (Search Results Temp)。以步驟 2 例子，可檢索到書名有包含 Data Mining 或 Fuzzy

或 Java 字串的新書。

4. 透過排名 (Ranking) 機制，將檢索結果呈現推薦給使用者。模組運作過程如圖 10 所示。

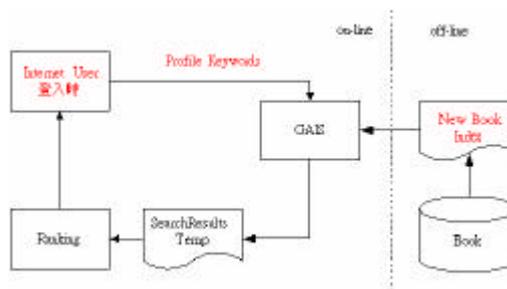


圖 10、新書推薦流程圖

## 五、改良式個人化推薦模組

在個人化系統中，所強調的是系統能夠立即的給予回應。若搜尋時間過久，會導致使用者喪失耐心，而降低顧客滿意度。以個人記錄檔資料表中有 10,000 紀錄為例，要比對登入使用者的個人記錄檔與其它一萬筆記錄(10,000)間，找出同好使用者再作推薦動作。如此一來，每有一個使用者登入系統時，系統都要重新計算一次。隨著資料表中資料記錄的增加，所需要的記憶體很龐大，搜尋時間會增加，降低系統效能，將無法符合個人化需求。

為了達到即時推薦的目的，本研究採用 SCSSA 演算法的分類方法，事先找出重疊程度 (Similar threshold) 高的樣本，歸類為同一分群 (同好使用者)，以加速個人推薦搜尋的速度，將個人化環境圖書資料推薦的系統效能提高，在本研究中稱為改良式個人化

推薦模組 ( Improved Personalize Recommendation Module, IPRM ), 參見圖 11。

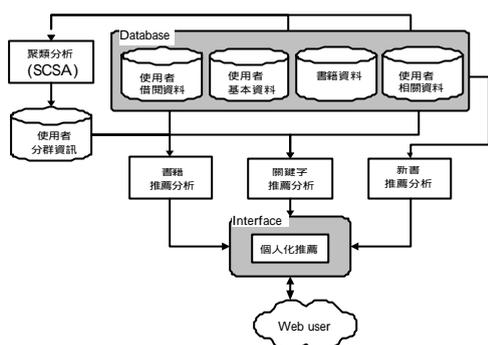


圖 11、改良式個人化推薦模組架構圖

(一) 個人興趣檔的相似度衡量

使用公式 (1) 來計算個人興趣檔的相似度。於本研究中 Case 代表個人興趣檔，Patterns 代表個人興趣檔中的關鍵字。而透過上述的方法得到的相似程度(Measure), 代表使用者間個人興趣檔相似程度。舉例說明：如表 7 所示，以公式 (1) 來計算 \$valid\_user 與其他使用者的相似度，可得到表 8。

表 7、個人興趣檔簡例

Username	keyword (個人興趣檔關鍵字)
1	Data Mining, Fuzzy, Visual Base, Java
2	Data Mining, Java, Database
3	Fuzzy, Java, Neural, Databas
\$valid_user	Fuzzy, Visual Base, Neural, Database

表 8、個人興趣相似度範例

Username	Measure
\$valid_user,1	-2
\$valid_user,2	-4
\$valid_user,3	1

(二) 使用 SCSA 演算法聚類分群

本研究採用 SCSA 演算法的快速分群能力，期望將重疊程度高的樣本合併成同一分群。透過個人興趣檔相似度衡量可得到個人興趣檔相似程度，再來使用 SCSA 演算法做重疊程度比對，就可以判斷是否歸為同一分群。如表 8，假設相似度門檻值是相似程度(Measure)為 0，則相似程度大於 0 歸為同分群；相似程度小於 0 為不同分群。所以表 8 中，\$valid\_user 和使用者 3 為同分群 (Measure=1)。

使用者資料中有 n 個使用者，則會有 n 個樣本，每一個使用者代表一個樣本，我們要計算第一個樣本點和其他樣本點重疊的部分，也就是個人興趣檔相似程度。若重疊程度大於 Similar threshold，則將之併入第一個分群，並給予一個分群編號；若小於 Similar threshold，則將之標示為另一個分群。以此類推，直到找出所有的分群為止。演算法如圖 3 所示。

(三) 推薦關鍵字與書籍

以上的處理程序都在系統背景先執行完畢，當使用者登入時，只要找出使用者與同分群的使用者，推薦他們個人興趣檔的關

鍵字和借閱書籍。

### 1. 推薦關鍵字

找出和目前使用者相同分群的使用者清單，產生推薦關鍵字清單進行推薦（與個人化推薦模組中推薦關鍵字步驟二相同）。以表 7、8 為例，使用者 3 和 \$valid\\_user 為同一分群，使用者 3 推薦給 \$valid\\_user 的同好關鍵字有 Java。假設門檻值 \$popularity=0，則最後推薦給 \$valid\\_user 關鍵字為 Java。

### 2. 推薦書籍

找出和目前使用者相同分群的使用者清單，產生推薦書籍清單進行推薦（與個人化推薦模組中推薦書籍步驟二相同）。以表 7、8 為例，使用者 3 和 \$valid\\_user 為同一分群，使用者 3 推薦給 \$valid\\_user 的書籍有 Java : how to program。假設門檻值 \$popularity=0，則最後推薦給 \$valid\\_user 書籍為 Java : how to program。

在個人化推薦模組的方法中，每有一筆新進資料進入，則須重新比對此新樣本點和樣本空間中所有點，比對的時間會超出使用者所能等待的時間，對於提供個人化服務的系統來說無法達到其效能。

IPRM 的方法，先依樣本空間中點的分布特性做好分群，分群大小不限，而是依群聚特性決定，符合實際資料分群特性。將樣本空間內分成若干個分群後，降低新進資料比對時間，提昇系統效率。

## 肆、系統實作與探討

依據本文所提出適用於數位圖書館之個人化檢索與推薦服務模組，設計一套數位圖書館個人化檢索與推薦服務系統與網站雛形，以便對本文所提出之理論基礎與系統架構作進一步驗證與討論。

系統開發環境與工具採用 Pentium M 1G MHz 個人電腦以及 Windows 2000 作為系統之開發環境，主系統採用 Apache 網頁伺服器加上 MySql 資料庫。網站部份則採用 PHP(Hypertext Preprocessor) 與 Macromedia Dreamweaver MX 作為開發工具。

### 一、改良式檢索模組

改良式檢索模組（IRM）中包含兩個檢索介面：基本檢索及進階檢索。基本檢索提供了關鍵字同音及容錯檢索，檢索結果可以由使用者自訂呈現排名方式，有中國圖書分類法的分類主題、借閱率和點閱率，如圖 12。透過使用者自訂排名更能符合使用者需求。



圖 12、改良式檢索模組一

而進階檢索則是使用者可輸入多個查詢字串，並使用布林運算式 (AND、OR、NOT) 來自訂的搜尋策略。而排名方式和基本檢索相同有中國圖書分類法的分類主題、借閱率和點閱率，如圖 13。



圖 13、改良式檢索模組二

## 二、個人化檢索模組

使用者查詢館藏資料時，使用者可以選擇是否加入個人興趣檔中的關鍵字。系統會利用使用者查詢關鍵字和使用者個人興趣檔中的關鍵字，來判斷使用者最需要的館藏，把查詢結果先加以過濾和排序後再呈現給使用者，如圖八。於圖 14 上有一個使用者資訊區塊，就是使用者可以對個人興趣檔建立與維護，包含增加刪除關鍵字。



圖 14、個人化檢索模組一

## 三、個人化推薦模組

使用者可以選擇是否加入推薦的關鍵字，如圖 15。在個人化檢索畫面上，右上角區塊為同好推薦的書籍。左下角區塊為個人化的新書推薦，如圖 16。



圖 15、個人化檢索模組二



圖 16、個人化推薦模組

檢索結果分別以中國圖書分類法、借閱率和點閱率排如圖 17、18、19。

查詢結果：共 43 筆  
中國圖書分類：

網際網路Internet  
特殊電腦方法：多媒體  
全球資訊網 (World wide web)  
電腦網路  
**電子資料庫**  
介質與存取固定 (標準)

查詢結果：共 3 筆  
中國圖書分類：

書號	書名
00130725	第三屆國際資訊管理研討會暨學術研討會論文集(附) 加贈1張中小企業
00103111	臺灣經濟的現狀、發展與未來之研究(以東大學 經濟系系刊刊)
00109420	從科學傳播觀點探討管理學與資訊科學的資料技術 行爲研究

書號	書名
00143247	網際網路的應用
00159666	網際網路的市場行銷
00159675	網際網路的應用與實務之製作(Openpage 2D Webmaster 完全設計手冊)
00151283	袖珍書的新世界(附附附附附附附附附附附附附附附附) 附附附附附附附附附附附附附附附附附附附附附附
00151954	行銷(OT)的附附附附附附附附附附附附附附附附
00143971	網際網路的客戶服務
00143956	新類型UNIX系統與附附附附附附附附附附附附附附附附
00152177	網際網路的附附附附附附附附附附附附附附附附附附附附附附
00172584	使用附附附附附附附附附附附附附附附附附附附附附附

圖 17、中國圖書分類法檢索結果

借閱率排名

書號	書名
00103032	非常JAVA 手冊
00103034	輕鬆學會 java script
00104424	java 語言入門
00109399	java 程式設計
00109392	WWW 程式設計: java Script 程式發展手冊
00103150	java 技術手冊
00109395	java 實用範例馬上活用
00109397	java 程式發展入門手冊
00159719	java 1.2-24小時自學手冊
00159817	java 物件導向程式基礎

圖 18、借閱率排名檢索結果

點閱率排名

書號	書名
00194824	精彩HTML: javascript+VBScript+ASP+CSS
00195324	CSS/HTML 網頁設計-網頁概論/HTML 基本語法/CSS 入門/軸軸、滑鼠、邊境特效/CSS與XML javaScript 整合
00103036	JAVA 魔法秘笈
00103032	非常JAVA 手冊
00102676	java Applet: 在WWW 上的應用
00103034	輕鬆學會 java script
00104424	java 語言入門
00104419	JAVA 精要
00103151	java 探險
00109399	java 程式設計

圖 19、點閱率排名檢索結果

四、模組優點

(一) 改良式檢索模組

1. 以智慧型 GAIS 搜尋引擎為基礎，提供使用者一個彈性、可自定的搜尋策略與輸出結果的能力。
2. 經過借閱率排行或點閱率排行的篩選，查詢到的書目是更加符合大眾化興趣的資料。
3. 中國圖書分類法的分類呈現方式幫助使用者查詢，提升使用者需求的正確性。

(二) 個人檢索模組

1. 在系統登入時即要求使用者帳號，密碼的確認動作，確保使用者的隱私權 (Privacy)。
2. 此模組會收集使用者搜尋時所用的關鍵字，並新增至使用者個人興趣檔。
3. 使用者可以自行維護個人興趣檔，讓系統知道使用者興趣所在。
4. 檢索時加入個人興趣檔，查詢到的書目將更加符合個人化興趣的資料。

(三) 個人化推薦模組

1. 可以由個人興趣檔的關鍵字和借閱紀錄主動推薦更多合適的資源(同好個人興趣檔的關鍵字或借閱的書籍)給有同樣興趣或有需求的使用者。
2. 將推薦的項目整合於相同的畫面，以區塊的方式呈現，讓使用者更方便瀏覽。
3. 透過推薦門檻值 \$popularity 的調整，所推薦的關鍵字或書籍會越能滿足使

用者的喜好。

4. 透過 SCSA 演算法聚類分群，加速個人推薦的搜尋速度，以達到線上即時推薦。

## 伍、結論與未來展望

在本研究中，設計與實作一「數位圖書館上個人化檢索與推薦系統」，系統主要分成三個模組進行：改良式檢索模組、個人化檢索模組和個人化推薦模組。運用使用者個人興趣檔，來建構個人化圖書檢索介面與個人化推薦服務，幫助使用者篩選圖書館的館藏，讓使用者可以更快地找到有興趣的、甚至是以前從來沒想到或沒接觸過的館藏。

經由推薦門檻值的調整，所推薦的關鍵字或書籍會越能滿足使用者所喜好。而且透過 SCSA 演算法聚類分群，可以加速個人推薦搜尋的速度，以達到線上即時推薦。

未來，可藉由此系統的使用率及非熱門書目之借閱率來了解此網站的成效如何，以作為改進的參考，讓數位化圖書館品質更加提昇。如此，個人化環境的功能將更完整。未來的研究與應用可朝著以下的方向進行：

### 一、系統評估方式

在本研究中個人興趣檔是由許多型態 (Patterns) 所組成，與一般量化 (quantize) 評估方式不太相同，所以在缺少大量實際資料和研究時間限制下，無法對本研究系統進行評估。這也是未來應著重的研究方向。

### 二、個人興趣檔中關鍵字權重值

在個人興趣檔中每個關鍵字重要性不同，所以未來可以考慮使用本體論 (Ontology)，將關鍵字加入不同權重值 (weight)。讓個人興趣檔更能代表使用者興趣。

### 三、個人化瀏覽

將使用者所需資訊整合於相同的畫面，以區塊的方式呈現，讓使用者更方便瀏覽。也就是個人化瀏覽的做法，將也是未來應著重的研究方向。

### 四、演算法的改進

由於個人化推薦應用在圖書館上所需要的記憶體很龐大，未來期望能找出一個更適合應用在圖書館上的資料探勘演算法，減少記憶體並與資料庫結合，整合出一個更適用的檢索與推薦系統。

### 五、門檻值探討

在未來，我們可以用更大的資料量來驗證本研究門檻值對於搜尋結果的影響，並深入探討不同門檻值對於不同資料類型之適用性，找出一些適用的建議門檻值。

## 參考文獻

- [1] 卜小蝶，圖書資訊檢索技術，台北，文華圖書，85年。
- [2] 卜小蝶，「淺談個人化服務技術的發展趨勢對圖書館的影響」，國立成功大學圖書館館刊第二期，87年10月。

- [3] 王姿琄,「圖書資訊系統之流通模組和 Profile Server 個人化資訊服務」,中正大學資訊工程研究所碩士論文,89年1月。
- [4] 余明哲,「圖書館個人化館藏推薦系統」,交通大學資訊科學研究所碩士論文,92年6月。
- [5] 林俊佑,「在數位圖書館多代理人系統中以本體論為基礎的內容檢索」,清華大學資訊工程系碩士論文,90年6月。
- [6] 吳哲慶,「大型搜尋引擎個人化」,中正大學資訊工程研究所碩士論文,91年1月。
- [7] 柯皓仁,「電子圖書館實施經驗談—以交通大學圖書館為例」,書苑季刊,第47期,pp.10-33,2001。
- [8] 唐秀珠,「以書目計量學方法探討專題選粹服務的發展」,台灣大學圖書資訊研究所碩士學位論文,76年6月。
- [9] 楊雅雯,「個人化數位圖書資訊環境—以PIE@NCTU為例」,交通大學資訊科學研究所碩士學位論文,90年6月。
- [10] 湯春枝,「從個人化服務行銷的理念談交通大學個人化數位圖書資訊服務PIE@NCTU系統」,國立成功大學圖書館館刊,第九期,91年4月。
- [11] 孫冠華,「圖書館新書推薦之個人化服務方法」,中山大學資訊管理研究所碩士論文,89年7月。
- [12] 鄭玉玲,「運用資料探勘技術實作數位圖書館上個人化之檢索與推薦服務-以南華大學圖書館為例」,南華大學資訊管理學系碩士論文,92年6月。
- [13] 施毓琦,「大學圖書館網站個人化服務之使用者需求研究」,臺灣大學圖書資訊學研究所碩士論文,91年6月。
- [15] 賴永祥,中國圖書分類法,增訂八版,台北,文華圖書管理資訊股份有限公司,90年。
- [15] L. Welling, and L. Thomson, PHP & MySQL 完全攻略,台北,上奇科技,90年。
- [16] A. Gambles, “The HeadLing personal information environment: Evaluation phase one,” D-Lib Magazine, Vol.7, Num.3, June.2002.
- [17] B. Mittal and W. Lassar, “The role of personalization in service encounters,” Journal of Retailing, Vol. 14, No. 1, pp. 95-109, 1996.
- [18] C. Faloutsos and D.W. Oard, “A Survey of Information Retrieval and Filtering Methods,” Technical report, University of Maryland, 1995.
- [19] D. Goldberg, et al., “Using collaborative filtering to weave an information tapestry,” Communications of the ACM, Vol.35, No.12, December 1992.
- [20] G.G. Chowdhury, Introduction to modern information retrieval, London: Library Association Publishing, p1, 1999.

- [21] J. E. Frook, "Future trend: getting personal with customers," *Internet Week*, Nov.1998.
- [21] J.Ghaphery, M.Kesselman, and S. B.Watstein, "Personalized information clients: Short answers to simple questions about "my library" services," *Reference Services Review*, Vol.29, No.4, pp.276 - 281, 2001.
- [22] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, California, USA, 2000.
- [23] K. Jones and P. Willet, *Readings in Information Retrieval*, Morgan Kaufmann Publishers ,CA, USA , pp.1-7, 1997.
- [24] K. Morgan and T. Reade, "Pioneering portals:MyLibrary@NCState.Informatio-n Technology and Libraries," Vol.19, No.4, pp.191-198, 2000.
- [25] LITA Top Tech Trends, "LITA Top Tech Trends January 1999 ALA Midwinter Meeting", Retrieved June 2, 2002.
- [26] Liu., X., Entropy, "distance measure and similarity measure of fuzzy sets and their relations." *Fuzzy Sets and Systems*, pp.308-318, 1992.
- [27] R.D. Hof, "Now it' s your web," *Business Week*, pp. 68-74, Oct. 1998.
- [28] R.D. Lawrence, et al, "Personalization of SupermarketProduct Recommendations," *Data Mining and Knowledge Discovery*, Vol. 5, No. 1-2, pp. 11-32, 2001.
- [29] P. Resnick and H.Varian,"Recommender Systems", introduction to special section of *Communications of the ACM*, Vol.40, No.3, March 1997.
- [30] T. Anandanpillai and T.A. Barta, "A case-based reasoning system for housing discrimination law", *Expert System with Applications*, pp.315-324, 1999.
- [31] W.Y. Lin, S.A. Alvarez, and C. Ruiz, "Efficient Adaptive-Support Association Rule Mining for Recommender Systems," *Data Mining and Knowledge Discovery*, Vol.6, No.1, pp. 83-105, 2002.