

模糊詮釋資料與標示語言及其在電子文件之應用

蕭勝文* 林信成* 陳勇任**

*淡江大學資訊與圖書館學系

**淡江大學教育科技學系

wen@mail.dils.tku.edu.tw

摘要

本研究首先從模糊理論與全文標示兩個領域角度切入，探討其在電子文件檢索系統上之重要性；接著，提出模糊詮釋資料 (Fuzzy Metadata) 之概念，實際應用於電子文件檢索系統上。我們設計了兩套實驗性系統，主要研究成果有二：(1) 電子新聞管理系統，針對新聞資料內容進行軟性模糊分類，使得檢索更具彈性，亦可提供較佳回線率；針對新聞內容以 XML 格式進行全文標示，使內容更具自我描述性，更能夠達到較高精確率；(2) 圖書模糊分類檢索系統，提出模糊詮釋資料 (Fuzzy Metadata) 之概念，並將其應用於圖書館的分類與編目上，以加強書目資料的關聯性，提升檢索結果的廣度。

關鍵詞：詮釋資料、模糊理論、模糊詮釋資料、模糊分類、標示語言

Abstract

At the first, the study focuses on the importance of electronic system for documents in the view of Fuzzy Metadata and practically apply it to the searching system. We develop two experiential systems and the main research results of news processed with softly fuzzy classifying, so that it could makes the search more elastic and could provide higher recall ratio. Besides, the contents of news marked with full text in XML makes the content more self-descriptive and more precise. Classifying and cataloging system, this study proposes the concept of Fuzzy Metadata, and then applies it to library classification and catalogue for the purpose of reinforcing the connection of bibliography, such that the broadness of search results could be improved.

Keywords: Metadata, Fuzzy Theory, Fuzzy Metadata, Fuzzy Classification, XML

1. 前言

網路的發達改變了人們取得資訊的方式及習慣，以往得在浩瀚的書籍與報章之間奔波查找，現今透過網路只要在彈指之間即可取得數量龐大的資訊。然而，當大量電子文件透過網路出版、傳播

之後，如何組織這些龐大的資訊以及提供有效的檢索機制，便成為數位典藏與電子出版等相關研究領域所考慮的重要課題之一。一般而言，回現率 (Recall Rate) 和精確率 (Precision Rate) 是評估檢索效能的兩大重要指標。

因此本研究設計了兩套實驗性系統，電子新聞資料庫管理系統與圖書模糊分類檢索系統，並提出模糊詮釋資料 (Fuzzy Metadata) 之概念。

2. 電子文件模糊分類

分類 (Classification) 是知識管理 (Knowledge Management) 的基礎工程之一，各行各業在自己的專業領域中隨時隨地都在進行著分類工作。分類規則有時必須遵循共通的規範，有時則以自訂的方式為之；而分類模式又可分為明確分類模式與模糊分類模式。

2.1 明確分類模式

傳統圖書館分類模式為明確分類 (Crisp Classification)，屬於硬性分類法。主要作法是將某一冊圖書歸屬於特定類別，並給予分類號。書籍與類別的關係為明確關係 (Crisp Relation)，只有「屬於」和「不屬於」兩種可能。

假設某一圖書館館藏中共有 n 冊圖書，其中第 i 冊記為 d_i ，並以 D 表示所有圖書的集合，即 $d_i \in D, i=1,2,\dots,n$ ；現有 m 個圖書類別，第 j 類記為 c_j ，並以 C 表示所有類別的集合，亦即 $c_j \in C, j=1,2,\dots,m$ 。對於某冊圖書 d_i 與某個類別 c_j 的歸屬關係，我們可以用明確集合表示法表示如下：

$$\sigma_{ij} = \begin{cases} 0, & d_i \notin c_j \\ 1, & d_i \in c_j \end{cases} \quad (1)$$

其中， σ_{ij} 為特徵函數，若 d_i 屬於 c_j ，則 $\sigma_{ij} = 1$ ，否則 $\sigma_{ij} = 0$ 。由於 σ_{ij} 的值非 0 即 1，所以又稱為二值分類法，簡稱二分法。不論圖書館採用何種分類法 (杜威十進分類法、美國國會分類法或中國圖書分類法)，皆屬此種分類法，適用此一數學模式。

2.2 模糊分類模式

若被分類對象之主題涵蓋了好幾個學科領域，且有輕重程度之別，則明確分類模式便不適用。此時，屬於軟性分類模式的模糊分類法便可派上用場。若將 d_i 視為一個定義在論域 C 上的離散型模糊集合，且將 d_i 歸屬於某個類別 c_j 的程度以 μ_{ij} (介於 0~1) 表示，則此一模糊集合可記為：

$$d_i = \sum_{j=1}^m \mu_{ij} / c_j \quad (2)$$

亦即：

$$\begin{aligned} d_1 &= \mu_{11}/c_1 + \mu_{12}/c_2 + \dots + \mu_{1m}/c_m \\ d_2 &= \mu_{21}/c_1 + \mu_{22}/c_2 + \dots + \mu_{2m}/c_m \\ &: \\ d_n &= \mu_{n1}/c_1 + \mu_{n2}/c_2 + \dots + \mu_{nm}/c_m \end{aligned}$$

舉例而言， $d_1 = \mu_{11}/c_1 + \mu_{12}/c_2 + \dots + \mu_{1m}/c_m$ 代表 d_1 此則電子新聞在 c_1 這一類別裡的歸屬值是 μ_{11} ，在 c_2 這一類別裡的歸屬值是 μ_{12} ，在 c_m 這一類別裡的歸屬值是 μ_{1m} 。若將其歸屬函數畫成離散式模糊集合，則如圖 1 所示。圖中顯示出 d_1 新聞分別在 c_1 、 $c_2 \dots c_m$ 類別中所佔的歸屬函數值高低，愈接近於 1，表示屬於某類別的程度愈高，反之，愈接近於 0，表示屬於某類別的程度愈低。

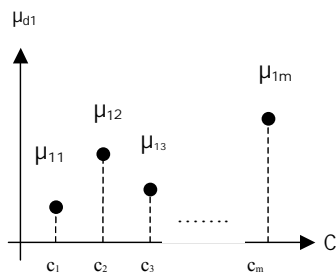


圖1 將電子文件以離散型模糊集合表示

再者，我們可將各電子新聞($d_1 \sim d_n$)在各個類別($c_1 \sim c_m$)中分屬的歸屬函數值($\mu_{11} \sim \mu_{nm}$)，以一個 $n \times m$ 的模糊關係矩陣表達如下：

$$\begin{matrix} & c_1 & c_2 & \dots & c_m \\ \begin{matrix} d_1 \\ \vdots \\ \vdots \\ \vdots \\ d_n \end{matrix} & \left(\begin{matrix} \mu_{11} & \mu_{12} & \dots & \mu_{1m} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \mu_{n1} & \mu_{n2} & \dots & \mu_{nm} \end{matrix} \right) \end{matrix}$$

圖2 電子文件與類別之模糊關係矩陣

由此可知， d_i 與 c_j 之間具備一種「模糊歸屬關係」，從語意的觀點而言可描述成「 d_i 歸屬於 c_j 的程度為 μ_{ij} 」。因此，上述的模糊關係矩陣實際上等同於下圖所示的模糊語意網路。我們可以很清楚的看出，模糊歸屬函數值在此網路中，扮演了一個將相關概念以不同的強度加以連結的重要角色。

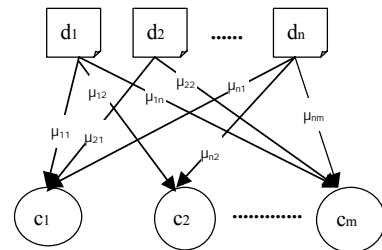


圖3 電子文件與類別之模糊語意網路

3. 模糊詮釋資料

在解釋模糊詮釋資料的意涵之前，我們先對模糊詮釋資料下一個簡單的定義：「所謂模糊詮釋資料，乃是指描述事物之模糊性質或彼此間之模糊關係的資料而言」。

綜觀目前大多數詮釋資料的發展而言，其目的大都在於「明確的」描述事物，對於各項資料之著錄皆有一定的規範以供遵循，極少具有模糊地帶，或涉及「模糊化」資訊者。然而，在真實世界中，模糊概念到處存在，許多事物的性質是模糊的、關係是模糊的，甚至人類有許多語言和思維也都是模糊的。因此，若要讓電腦可以處理這些模糊概念，便必須藉助模糊理論的原理，將這些模糊語意轉換成電腦能理解的模糊詮釋資料，以便使得所描述之事物，能更廣泛的涵蓋無法明確描述的模糊資訊。

總而言之，模糊詮釋資料使得自然直覺與電腦世界之間搭起了一座交流的橋樑，更符合人類軟性思維的本質。

4. 系統實驗結果

4.1 電子新聞管理系統

本研究所設計之電子新聞管理系統，可分為前端子系統與後端子系統兩部分，如圖 4 所示。若依功能劃分則可區分為五大部分，分別是：(1) 呈現模組 (Presentation Module)；(2) 檢索模組 (Searching Module)；(3) 編輯模組 (Editing Module)；(4) 管理模組 (Management Module)；(5) 資料庫 (Database)。

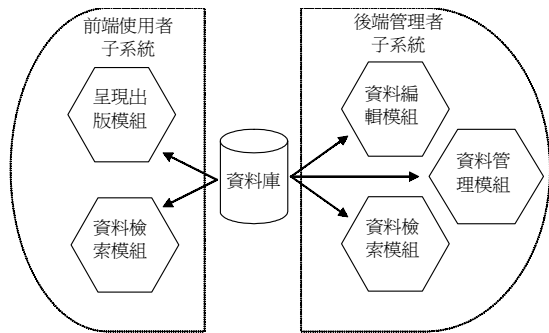


圖 4 系統功能模組

(1) 檢索模組

本模組提供新聞資料檢索功能，主要由「模糊檢索引擎」與「語意檢索引擎」所構成：

(a) 模糊檢索引擎：提供新聞分類的模糊檢索功能，使用者可決定欲查詢之新聞資料其所屬分類的歸屬函數值(0 至 1)，模糊檢索引擎針對所有新聞資料錄之歸屬函數值，在其對應的模糊集合當中依據使用者的檢索條件進行搜尋，使得其檢索的方式更具彈性，以提供較佳的回現率。

(b) 語意檢索引擎：傳統的搜尋引擎僅提供資料欄位的檢索功能，並未提供針對內容語意上之特定目標加以檢索，如人名、地名等，本模組除提供傳統的欄位與關鍵字檢索功能之外，藉由 XML 具備資料自我描述之特性，可針對新聞內容的人、事、時、地、物等內文語意加以檢索，提高檢索結果的精確率。

(2) 資料庫

我們依據上一單元所分析的模糊分類模式，可以很容易的將模糊關係矩陣轉換為關聯式資料表的模糊分類欄位，如表 1 所示，並給予歸屬函數值，使得檢索模組可藉以判別新聞所屬之模糊類別。

表 2 模糊關聯式資料表

	C_1	C_2	...	C_m
d_1	μ_{11}	μ_{12}	...	μ_{1m}
d_2	μ_{21}	μ_{22}	...	μ_{2m}
⋮	⋮	⋮	...	
d_n	μ_{n1}	μ_{n2}	...	μ_{nm}

使用者可經由資料檢索模組，輸入欲查詢之條件(可選擇搜尋模式：模糊檢索或語意檢索)，如圖 5 所示，隨後資料檢索模組至資料庫中搜尋所需資料。檢索模組的檢索工作完成後，由資料庫所取出之資料後交由資料呈現模組，依照系統所訂定之 DTD 來確認 XML 資料的合法性，隨後將呈現給使用者。呈現的 XML 資料透過 DSO 以及 DOM 的解讀，可在不更改原始新聞內容之下，給予各種排版樣式，如圖 6 所示。

圖 5 資料檢索模組



圖 6 各種排版樣式

系統實驗結果進一步說明如下：

(1) 模糊檢索結果

模糊檢索引擎的目的在於提供較佳的回現率。以本系統資料庫內之新聞『今起台灣民眾赴港可辦電子簽證節省時間金錢』為例，由於內容描述香港政府之旅遊發展局等相關單位有鑑於台港旅遊人次繁多，簽證手續繁雜，因此推出新的簽證方式以節省時效。在實質上可歸屬於『生活類』的新聞；但此篇新聞內容亦與香港政府相關，因此與『國際類』亦有若干關聯，若以一般傳統的單一類別的分類法，直接將此篇新聞歸屬於生活類，不但有欠周詳，也會造成使用者在『國際類』中找不到此篇報導的問題。因此，如果我們適度的給予此篇新聞模糊歸屬函數值，使其歸屬於『生活類』與『國際類』的程度分別為 0.9 與 0.7，則藉由模糊分類檢索功能，使用者不但可以在『生活類』中檢索出這篇新聞(如圖 7 所示)，亦可在『國際類』中檢索到(如圖 8 所示)，不但增加了檢索的彈性，也提升了回現率。

日期	主題	模糊歸屬度
2002/3/16	公事處將加密平日上山班次	1.0
2002/3/18	今起台灣民眾赴港可辦電子簽證節省時間金錢	0.9
2002/2/25	大學學潮 兩萬人未過關	0.9
2002/3/18	嘉義縣大尖山開發新景點	0.8

模糊類別: Life / 相關係數>=0.7

圖 7 檢索『生活類』新聞之結果

日期	主題	類別
2002/4/3	馬里蘭校園 瘋狂到不行	International
2002/3/18	小獵號航艦難離日本機須實基地	Society
2002/3/18	二億五千萬元 打造豪華派對	Political
2002/2/25	印度前所博物館 教你認識印度史	Finance
2002/3/18	六家合資銀行登陸 中共第二季審批	International
2002/3/18	新加坡出現大陸人留學和培訓新浪潮	International
2002/3/18	今起台灣民眾赴港可辦電子簽證節省時間金錢	Society
2002/3/18	哈佛商學書刊在大陸暢銷	Life
2002/3/18	大陸漁工兩週內解禁	Political

圖8 檢索「國際類」新聞之結果

此外，我們也在模糊分類檢索的功能中提供使用者自由選擇相關係數（即模糊歸屬函數值）之選項，以利使用者篩選新聞分類的相關度，相關係數愈高表示歸屬程度必須愈高者才會被檢索出來，如圖 9 所示。

新聞分類:

模糊分類搜尋: 相關係數: 以上

內文檢索: 類別:

圖9 模糊分類檢索功能

(2) 語意檢索結果

本系統之新聞資料在經過 XML 標示加值處理之後，使用者除可依據標題、作者、時間 ... 等資料欄位進行檢索外，也可針對新聞全文內容，依據人、事、時、地、物等語意條件進行更精確的內文語意檢索。例如，圖 10 乃是以「大學」作為關鍵字詞，選擇以不限標示的方式進行檢索之結果，共找到八篇文章；對於同樣的檢索詞，如果將內文檢索條件限制於「地」，表示使用者欲檢索之條件僅為內文語意上與「大學」相關的地方、地名或地點，而非所有與「大學」概念相關的文章，結果如圖 11 所示，找到七篇符合內文語意檢索條件之文章；再者，若檢索詞仍為「大學」，但將內文檢索條件限制於「事」，即表示使用者欲查詢之條件僅為有關大學的「事件」而非地點或其他，則如圖 12 所示，更精確的只找出兩篇在內文語意上含有大學相關事件的文章。

日期	主題	類別
2002/4/3	馬里蘭校園 瘋狂到不行	International
2002/4/3	大學生登山失蹤案 不排除謠言	Society
2002/4/3	泛藍軍有聲音 拱黃俊英選市長	Political
2002/3/18	哈佛商學書刊在大陸暢銷	Finance
2002/3/18	新加坡出現大陸人留學和培訓新浪潮	International
2002/3/18	保證留學大陸 補習班涉詐欺	Society
2002/2/25	大學學測 兩萬人未過關	Life
2002/2/25	三類官員 評價最差	Political

圖10 不限檢索標示之結果

日期	主題	類別
2002/4/3	馬里蘭校園 瘋狂到不行	International
2002/4/3	大學生登山失蹤案 不排除謠言	Society
2002/4/3	泛藍軍有聲音 拱黃俊英選市長	Political
2002/3/18	哈佛商學書刊在大陸暢銷	Finance
2002/3/18	新加坡出現大陸人留學和培訓新浪潮	International
2002/3/18	保證留學大陸 補習班涉詐欺	Society
2002/2/25	三類官員 評價最差	Political

圖11 限定檢索標示為「地」之結果

日期	主題	類別
2002/4/3	大學生登山失蹤案 不排除謠言	Society
2002/2/25	大學學測 兩萬人未過關	Life

圖12 限定檢索標示為「事」之結果

4.2 圖書模糊分類檢索系統

我們以中國圖書分類法 0~9 類為基準，從淡江大學圖書館中，每一類找出兩本圖書，共取得二十本圖書及其書目資料作為實驗用途。這些書的內容大都不只涵蓋單一領域，如「圖書館學的哲學」、「新聞心理學概論」、「土地倫理」... 等，但由於傳統上，圖書館仍是採用明確的分類與編目規則，所以每本書都僅歸屬於一個單一類別，給予一個單一類號，我們將原始的分類在圖 13 左側 A 部分標示出來，此部分顯示了圖書與類別之間的明確歸屬關係。

接著，我們經由檢閱這些書籍的內容，重新歸納其涵蓋的主題範圍；再依據前述的模糊理論概念，將這些書籍進行模糊分類，然後彈性的給予每本書籍多重類號及適當的模糊歸屬函數，圖中央 B 部分說明了模糊分類後圖書與類別之間的模糊歸屬關係。如此一來，模糊分類不但直接的在圖書與類別之間，建立了輕重程度不一的模糊關係，更間接的使得圖書與圖書彼此之間也建構起了模糊關係，如圖 13 右側 C 部分所示。

這些由模糊關係所交織出來的模糊語意網路並非無中生有，而是原本就存在於這些書籍之間，只不過因為傳統上採用明確分類法，使得編目員在分類時，不得不取主要類別而捨次要類別，以致強制切斷了書籍之間天生的關聯性，模糊分類僅是重建這些原已存在的關聯性罷了。書與書之間原本就不應被視為獨立了個體，藉由加入模糊詮釋資料，使得圖書的主題就像地圖一樣能互通相連，讀者便能經由這個主題類號找到特定的圖書，也能在另一個類號裡找到相同之書籍，如「圖書館學的哲學」，雖其正規明確的類號屬於「0 類：總類」，但模糊分

類除可歸屬於「0類：總類」外，尚可歸屬於「1類：哲學類」，所以當讀者查找「圖書館學的哲學」這本書時，除了可在「0類：總類」找到之外，也可以在「1類：哲學類」中找到，而且還可以藉由書籍間的模糊語意網路查找到相關書籍，進行延伸閱讀。例如，由圖中可知，借閱「圖書館學的哲學」的讀者有可能對哲學有興趣，因此，有可能成為其延伸閱讀的相關書籍有「哲學與新聞」、「新聞心理學概論」等；也有可能對圖書館建築有興趣，則「土地倫理」等書則較相關。此外，各書籍之關係度更可從歸屬函數判斷其輕重，數值越接近於1越相似，反之，越接近於0則越不相似。因此，假使一本書籍其內容涵蓋多個領域，其所屬類別也將分屬多重類別，同類型的相關書籍也將越廣。

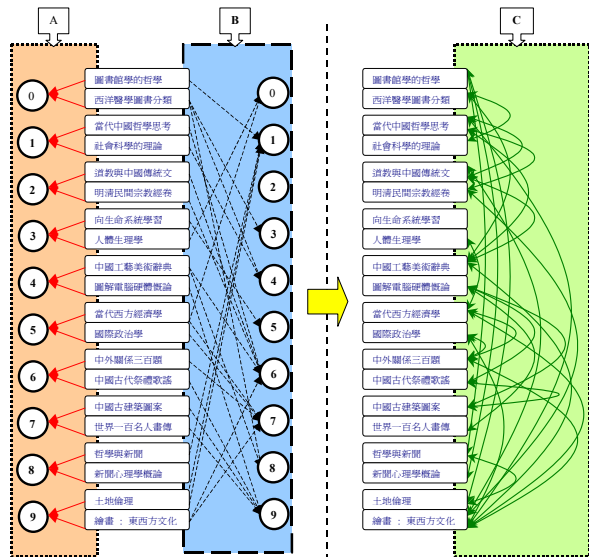


圖 13 明確分類與模糊分類關係圖

本系統主要分為 (A) 前端使用模組：提供讀者透過個人需求選擇適當檢索方式進行書目檢索，依功能分為：(1) 明確分類檢索模組 (Crisp Search Module)、(2) 模糊分類檢索模組 (Fuzzy Search Module)、(3) 相似書籍檢索模組 (Similar Search Module)；(B) 後端管理模組：提供系統管理者維護書目資料之新增、刪除、修改以及模糊分類度數選擇之動作，進入該模組前必須鍵入特定帳號與密碼提供系統驗證，才能進入進行管理動作。系統架構如圖14所示。

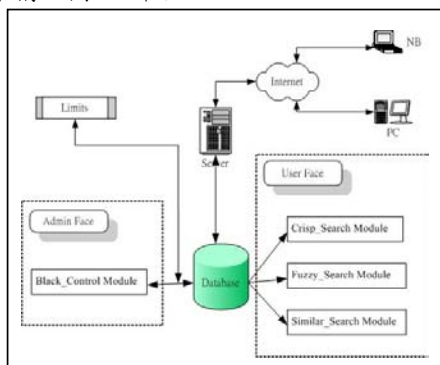


圖 14 系統架構圖

各模組之功能及實驗結果進一步說明如下：

(1) 明確分類檢索模組

本模組提供使用者鍵入關鍵字並選擇以書名、作者、出版社、出版地、出版年、館藏地、ISBN 進行比對查詢。檢索結果如圖 15 所示。



圖 15 明確分類檢索模組與結果

(2) 模糊分類檢索模組

本模組提供使用者選擇圖書模糊分類歸屬值，搭配圖書類別進行查詢，如圖 16 所示。類別部分提供使用者以下拉式方式選擇中國圖書分類法 0~9 類，模糊度部份也採用下拉表單方式供選擇其模糊歸屬函數值，以決定檢索資料時的關聯度，查詢結果之模糊度數越接近 1 之書籍，表示越接近所選擇之圖書類別。



圖 16 模糊分類檢索模組與結果

使用者選定類號與模糊度後，系統會依內部書目資料庫中的模糊詮釋資料進行比對，如核對成功系統會將書目資料依照所選擇類別之圖書模糊度數，依度數高至低排序出來。如圖 17 所示是使用者選擇類號為「總類」、模糊度為「0.1」之檢索結果，所有模糊度數高於 0.1 之總類圖書全部會呈列出來，越接近於 1 的某筆資料則會排序在最上面，以便使用者快速瀏覽。從圖中也可得知，「圖書館學的哲學」、「西洋醫學圖書分類系統析論」這二本書依內容來分類都含有「0類：總類」之類別，表示使用者可以透過此模糊檢索之功能檢索出同類型之書籍，由模糊度數來判斷「圖書館學的哲學」模糊度數 0.6，比「西洋醫學圖書分類系統析論」0.4 選接近於 1，因此屬於總類的程度較高，但兩者之間因度數相近，也代表兩者屬於同一類之書籍，如此一來，使用者更能掌握所檢索圖書的關聯性，進而提升檢索效率。

書名	作者	出版社	出版地	出版年	ISBN	索書號	館藏地	總類
圖書館學的哲學	賴鼎銘	文華	台北市	1983	9578708033	020.1	中文圖書區	0.6
西洋醫學圖書分類系統析論	劉淑萍	臺灣學生書局	臺北市	1979	957531582X	023.34	中文圖書區	0.4

圖 17 模糊分類檢索「0類：總類」之結果

圖 18 所示則為檢索類別選「哲學類」，模糊度數選「0.3」之結果，共有「人體生理學」、「社會科學的理路」、「當代中國的哲學思考」、「哲學與新聞」、「新聞心理學概論」、「圖書館學的哲學」以及「土地倫理」七本書，其中可以很清楚的看到「圖書館學的哲學」一書分屬總類與哲學類兩個領域的

程度分別為 0.6 和 0.3。

書名	作者	出版社	出版地	出版年	ISBN	書號	標地	模糊度
人體生理學	黃基樞	黎明圖書出版社	臺北市	2002	9576165156	397.6347	中文圖書區	0.03
社會科學的理論	黃鴻烈	心理出版社	臺北市	2001	9577024263	143.6365	中文圖書區	0.03
當代中國的哲學思潮	徐光壽	經濟科學出版社	北京	1994	752589598	120455	中文圖書區	0.03
哲學與新學	徐光壽	新華書店	北京	1991	720015814	890.1.6564	中文圖書區	0.03
新聞心理學概論	劉木林	北京廣播學院出版	北京	1993	7810040073	890.14.9727	中文圖書區	0.03
國家哲學的哲學	賴萬純	文華	台北市	1983	9578708033	000.1	中文圖書區	0.6 0.3
土地倫理	張元雷	富邦藝術基金會	臺北市	1999	9579775303	920	中文圖書區	0.03

圖 18 模糊分類檢索「1類：哲學類」之結果

此外，為了使對圖書分類法架構不清楚的使用者，能順利掌握圖書分類號，使檢索時所下的檢索策略更精確，我們設計了一個中國圖書分類法明細項目，供使用者參考使用，如圖 19 所示。



圖 19 中國圖書分類法詳細類表

(3) 相似書籍檢索模組

本模組提供使用者選擇圖書類別，系統會將資料庫裡的相似書目資料，依照此類別的模糊度數高低全部呈列出來。



圖 20 相似書籍檢索模組與結果

(4) 後端管理模組 (Black_Control Module)

此模組進入前必須登入特定帳號與密碼，提供管理者維護書目資料新增、刪除、修改以及模糊分類度數選擇之異動等動作。

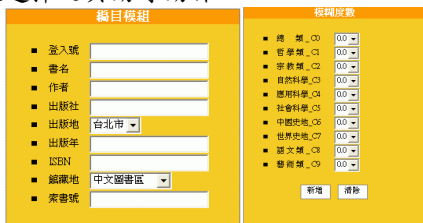


圖 21 後端管理模組

5. 結論

知識浩瀚無涯，無論使用哪一種分類系統，都無法百分之百的記載著文化資產與社會變遷的各種形式資料，設計出完美的組織架構。再加上時代科技的進步，人類知識日新月異，電子文件日漸增加、複雜，用來組織我們人類知識的分類系統，必須設法隨著知識創新快速的節奏跟上腳步。

本研究設計的二套系統，提出了結合模糊理論與全文標示之實際作法以及模糊詮釋資料的概念，主要研究成果是藉由 Fuzzy 理論的技術，以歸屬函數的高低來判別電子文件所屬類別輕重，並發

展出具有模糊搜尋的檢索功能；另一個成果是藉由 XML 進行電子文件內容的語意描述，各個進行資料處理及交換的模組皆以 XML 為基礎，系統內之所有資料亦採用 XML 格式，並發展出具有精確語意的檢索功能。由實驗結果可以清楚的看出結合模糊詮釋資料與標示語言之作法，的確可以有效的提升回現率與精確率。更可以重建知識間的關聯性，使其更加環環相扣，解決傳統分類系統的二分法問題，進而使得一般使用者可以利用模糊分類檢索系統搜尋相關資料，使查檢資料更直接、更有效率。

參考文獻

- [1] Bauer, Mary C, Dewey decimal classification : 200 schedules expanded for use, Catholic Library Association, 1988
- [2] HighWire Press 計畫，
<<http://www.highwire.org>>，(27 July .2003)
- [3] International Federation of Library Associations and Institutions : IFLA，<<http://www.ifla.org/>> (29 July 2003)
- [4] IPTC, "News Industry Text Format",
<<http://www.nitf.org/>> (20 July. 2003).
- [5] Jyh-Shing Roger Jang, Chuen-Tsai Sun, Eiji Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence, Prentice Hall (1996).
- [6] L. A. Zadeh, "Fuzzy sets," *Information and Control* 8 (1965), pp. 338-353.
- [7] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. on Syst., Man and Cybern.* SMC-3 (1973), pp. 28-44.
- [8] Library of Congress. Cataloging Policy and Support Office, *Library of Congress classification. H. Social sciences*, Washington, D.C. : Library of Congress, Cataloging Distribution Service, 1994
- [9] MARC Standards <<http://lcweb.loc.gov/marc/>> (29 July 2003)
- [10] Robert R. Korfhage, *Information Storage and Retrieval*, Wiley Computer Publishing, New York, 1997, pp. 196-199.
- [11] XMLNews.org, "XML and News Industry",
<<http://www.xmlnews.org/>> (20 July. 2003).
- [12] 林信成、彭啟峰, *Oh ! Fuzzy 模糊理論剖析*，台北：第三波，民 83。
- [13] 陳和琴、吳瑠瑤、江綉瑛編著，*圖書分類與編目臺北縣*，空中大學，民 85
- [14] 陳雪華，「網路資源與 Metadata 之發展」，*圖書館學刊*，12 期 (民 86 年)，頁 23-28。
- [15] 賴永祥，*中國圖書分類法*，增訂七版(台北市：商務，民 78 年)。
- [16] 謝瀛春、黃學碩、維習安、雷約翰、謝清俊，「新聞內容的標示-XML 之應用」，*海峽兩岸資料庫/數據庫與資訊/信息服務交流與合作論文集*，民 90.1，頁 205-212。