

行政院國家科學委員會專題研究計畫 成果報告

Web 2.0 社群資源分享系統之智慧型搜尋與個人化推薦的研究

研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 100-2221-E-343-007-
執行期間：100年08月01日至101年07月31日
執行單位：南華大學資訊管理學系

計畫主持人：邱宏彬

計畫參與人員：碩士班研究生-兼任助理人員：范楨祺
大專生-兼任助理人員：李宗育

公開資訊：本計畫可公開查詢

中華民國 101 年 10 月 11 日

中文摘要：社群資源分享系統是目前 Web 2.0 平台上非常受歡迎的一種應用，它允許使用者上傳他們的資源(圖片、影音、註解、書籍、文章等)，並且自由的加註任何標籤(Tags)在資源上。如何了解使用者的需求，提供友善的搜尋服務模式，讓使用者充分運用 Web 2.0 網站的資源，是 Web 2.0 應用成功的重要關鍵。本研究以旅遊共享資訊為研究對象，嘗試提出一個結合使用者意圖(Intension)、領域知識本體(Ontology)、與大眾分類法(Folksonomy)的智慧型知識搜尋與推薦機制，有效的輔助使用者清楚地表達其搜尋目的與需求，以進行有效率的資訊搜尋。本研究採用 YAHOO 旅遊網頁做為實驗資料來源，以建立旅遊知識本體以及大眾分類等資料庫。實驗結果顯示結合使用者意圖、領域知識本體、與 Folksonomy 在搜尋與排序推薦上確實能帶來效益。

中文關鍵詞：Web 2.0 社群資源分享系統、使用者意圖、領域知識本體、大眾分類法、智慧型搜尋

英文摘要：Social resource sharing systems are web-based systems that allow users to upload their resources (such as pictures, audio and video, books, documents etc.), and to label them with arbitrary words, so-called tags. How to provide the friendly search services, let the users fully utilize the shared resources of the Web 2.0 websites, will be the successful key of Web 2.0 applications. This study attempts to integrate the user intension, ontology, and folksonomy to propose intelligent searching and recommendation mechanisms by which the user searching goals can be described effectively. This study examines the system performance using the dataset from the YAHOO traveling website. The experimental results show the searching profits based on the user intention, the domain ontology, and the Folksonomy.

英文關鍵詞：Web 2.0 social resource sharing systems、User intension、Ontology、Folksonomy、Intelligent searching

一、前言

網際網路的發展已經超過十年，網路環境的成熟與資訊科技之進步，使得上網人數遽增，造成現今網際網路已進入一個以使用者為中心的 Web 2.0 新時代。Web 2.0 的本質就是集結眾人智慧所發展出的一種網路應用服務 [1-3]。Web 2.0 的一體三面包括內容、工具/服務、使用者，只重視任一面都不能成為 Web 2.0 成功的應用者 [4-5]。如何真正了解使用者的需求、應用 Web 2.0 的技術、工具及內容，提供符合使用者期待的資訊服務模式，以便使用者充分運用 Web 2.0 網站的資源，便成為 Web 2.0 環境中一項非常重要的議題。

隨著網際網路的蓬勃發展，利用網路搜尋資訊已成為時代趨勢。在目前通用的搜尋方法中，都以單一或多個關鍵字做為搜尋重點，雖然回應速度很快，但往往查得大批無關的資料，尚需一一檢視，花費許多時間才能得到想要之資訊 [6-7]。而且有時候給予適當的關鍵字，對於使用者而言並非一件容易的事。例如，當使用者想了解目前有那些較熱門的旅遊休閒活動或景點時，就很難決定精確而有效的關鍵字。

只利用關鍵字也無法有效表達使用者的需求和意圖。舉例而言：「何時適合賞櫻花？」、「到那裡賞櫻花？」、「如何從臺北到阿里山賞櫻花？」，都可以「賞櫻花」當關鍵字，但目的並不相同，其搜索「意圖」分別為「何時」「何地」「如何」。在不同的意圖下，所找到的資料也應不相同。因此，要如何描述自己的意圖來尋找答案，便成為重要的研究目標 [6-7]。再者，不同領域的使用者，看到相同的關鍵字各有不同的解釋，比方說「鹽水」，在地理專業領域上屬於「地名」，但在一般人看，就是加了鹽的水。如果在找資料的時候，事先考慮意圖與領域的不同，給予不同的關鍵字，那麼對使用者進行資料檢索時將非常有幫助 [6-7]。

標籤的作法，將描述內容主題的權利下放到一般的使用者，形成大眾分類 (Folksonomy) [3, 9-10, 14-16]，使用者找資料不再完全依賴知識內容分類法 (Taxonomy)，可以使用貼近民眾的用語來描述存放的內容。提供標籤工具的網站，更利用視覺效果呈現標籤雲 (Tag cloud)，傳達各個標籤在網站中被運用的情形 [1, 2]。另外，圖像通常欠缺文字說明，但在搜尋時，卻多是以字找圖，沒有字，就找不到圖。因此，透過標籤，圖就有了文字描述，圖像檢索系統也就能提供文字查詢功能；而且，這些關鍵字也反映使用者觀點，是建立圖像索引時，絕佳的入口詞彙 (Entry Term) 來源 [8]。

二、研究目的

社群資源分享系統 (Social Resource Sharing Systems) 是目前 Web 2.0 平台上非常受歡迎的一種應用，它允許使用者上傳他們的資源 (Resources) (網頁、圖片、影音、註解、書籍、文章等)，並且自由的加註任何標籤 (Tags) 在資源上 [14]。標籤多為簡短且足以代表該資源的關鍵字，也稱之為社會化標籤 (Social tagging)。現今網際網路有越來越多的標籤應用平台，幾乎已經成為

Web2.0 網站的標準功能。雖然這些 Web2.0 網站提供了一個好的社群平台，然而目前仍缺乏智慧型搜尋與個人化主動推薦的機制，來幫助使用者在廣大的共享資源中，有效率的找到符合使用者需求的資源。在週休二日、政府積極推動觀光產業，以及民眾逐漸重視休閒品質的趨勢下，休閒旅遊已蔚為風潮，同時，有越來越多關於旅遊經驗分享的文章，張貼在網站上。因此，本研究將以旅遊資訊為研究對象，設計能夠了解使用者意圖與偏好的智慧型搜尋與個人化推薦機制。經由搜尋與推薦系統的幫助，使用者可得到滿足個人需求的資訊服務，減少搜尋資訊成本，增加效率，讓使用者充分運用 Web 2.0 網站的共享資源。

在社群資源分享系統中，使用者貢獻了豐富的內容，這些共享的資源以標籤(Tags)的方式自由地加註、管理和分享。大量且經常變動的共享資源，蘊含著豐富的使用者經驗和知識。然而，現存的索引方式，並沒有充分的利用使用者的意圖和偏好，以引導使用者發現滿足其需求的資訊。本研究將整合使用者意圖、領域知識本體[12]、與 Folksonomy 等方法，以利用 Web2.0 強調「群體智慧」與「使用者經驗」的特性，在社群資源分享系統中，建構以使用者為中心的智慧型知識搜尋與推薦機制。讓使用者表達其意圖，並且引導使用者在大量的共享資源中，有效率的搜尋到有用的知識與經驗。同時，依據使用者偏好與行為，適時推薦給使用者所需的潛在資訊。我們將先以現代人生活中皆有的旅遊景點知識為研究對象，利用 YAHOO 旅遊網頁做為實驗的資料來源，來設計和實作此一個人化知識搜尋與推薦機制，以探討其可能的問題與解決方法，其成果將可望廣泛的應用到其他 Web2.0 社群資源分享系統上。

三、文獻探討

喻英雲 [6] 提出一種新的結合本體論與意圖之搜尋模式，使用者以自然語言輸入問題描述之後，搜尋引擎即可自動分析自然語言之語意，進而尋得符合使用者意圖之相關資料。該研究以數學領域本體為研究範圍，其搜尋模式共分為五個步驟，分別為建構數學領域本體、建構 5W1H 意圖句型庫、問句型態分析、意圖轉換、關鍵字擷取與擴展，最後將關鍵字組置於現有之搜尋引擎進行搜尋。實驗結果顯示，結合領域本體與 5W1H 意圖進行搜尋之方式，確實可提高搜尋資料之準確率並縮短查詢時間。賴嘉民 [7] 以相似的作法，應用在醫學領域知識本體上，協助搜尋所需醫學知識領域資料。然而，完全自動分析自然語言之語意，是一件很困難的工作，因此，這些方法還是在雛型階段，有許多技術性問題與演算法，須要進一步克服。再者，目前這些方法是針對傳統 Web 1.0 的搜尋引擎來設計，並未利用到 Web 2.0 群體智慧分享及協同創造的相關理念與工具。另外，領域知識與經驗是由系統設計者建立的知識本體決定，使用者沒有參與的空間，換言之，沒有利用到使用者分享的知識與經驗。

由於網路的普及，旅遊業者大都有建置本身的網站，以提供旅遊行程資訊。然而，目前並無完善的網站，整合各家旅遊業者所提供的行程訊息，以致消費者通常要花費很多時間去搜尋、比較

各家旅遊業者所推出的旅遊行程。邱英華、陳志豪等人 [11] 利用知識本體(Ontology)與語意網路服務(Semantic Web Services)技術來整合各家旅遊業者所提供的旅遊行程訊息，開發出以知識本體為基礎之旅遊行程語意網路服務系統，以達到網路服務自動化搜尋與調用的能力，讓網路服務更為人性化。同樣的，本方法並未利用到 Web 2.0 經驗分享及評比、推薦、評論等知識，以引導消費者找到有用的資訊。領域知識只由系統設計者建立的知識本體決定，沒有利用到消費者分享的經驗。

卜小蝶 (2007) [9] 提出 Folksonomy 具有以下特性：(1) 知識的分類是社會性的，與社群有關，是由社群成員來發展分類系統。(2) 使用者即產生標籤者。(3) 多使用自然語言。(4) 無嚴格的階層架構。(5) 重回饋，關鍵字是變動的。(6) 多採分堆 (Categorization) 作法，而不一定依照分類 (Classification) 方式。分堆是隨意主觀，而分類則多半具有嚴格的知識階層定義。Folksonomy 由使用者製作，雖然可以節省時間，但是標籤品質較不一致為缺點。傳統分類法則較為嚴謹，需要花較多時間建置 [3]。在[3]中也提到在 Taxonomy 內容中適時調整並加入 Folksonomy 的機制，可以讓 Taxonomy 具有更親和力的對應。Jessamyn West (2007) [15] 也說明結合 Taxonomy 與 Folksonomy 多元面向的標籤，會使得資訊更容易的被搜尋到，兩者皆有其不可抹滅價值。

四、研究方法

目前休閒旅遊已蔚為風潮，使用者可以在旅遊網頁上搜尋詳細的旅遊資訊，也可以在網站上張貼旅遊經驗分享的文章。YAHOO 旅遊網頁上將旅遊景點分門別類，同時有詳細的旅遊資訊，包括圖片、簡介、位置、官方網站等資料。登錄的會員也可以在網站上發表旅遊的經驗和評論，同時以自行定義的標籤(tags)加以歸類。例如”阿里山森林遊樂區”的標籤被標示為”日出、吉野櫻、阿里山、櫻花季、賞櫻、櫻花、春天、小火車”。但是，幫助使用者在龐大的共享旅遊知識資料庫中找到有用的文章，並且根據使用者的偏好，適時的推薦符合使用者潛在需求的資訊，並不是一件容易的事。本研究提出一個結合使用者意圖、領域知識本體[12]、與 Folksonomy[14-16]的智慧型知識搜尋與推薦機制，讓使用者以自然語言輸入需求描述，同時明確表達其意圖。藉由意圖的指定，將使得從自然語言的分析中，自動擷取出關鍵字變得較容易，而且方便進行階段性的搜尋。領域知識本體與標籤 Folksonomy，所建立的輪廓檔資訊，則進一步輔助推展出相關的標籤關鍵字。之後，將這些標籤關鍵字組置於 Web2.0 搜尋引擎上進行搜尋。搜尋的結果再以輪廓檔資訊進行主動式推薦，如此，可望幫助使用者充分運用 Web 2.0 社群網站的分享資源。圖 4-1 為本系統之架構圖，本系統主要分為「改良中文斷詞系統」以及「旅遊標籤匹配系統」兩個部分。

本研究希望藉由搜尋標籤以及斷詞技術、本體論、關聯法則來建立旅遊標籤匹配系統，讓使用者有更好的搜尋結果。改良中文斷詞系統為了解析使用者的語意，從使用者的問句中擷取出意圖以及適合的關鍵字；建置旅遊本體目的在於解決同義字、地理階層的問題；使用關聯法則[13]可以讓我們找到標籤之間所隱含的規則，進而匹配關聯關鍵字給使用者。

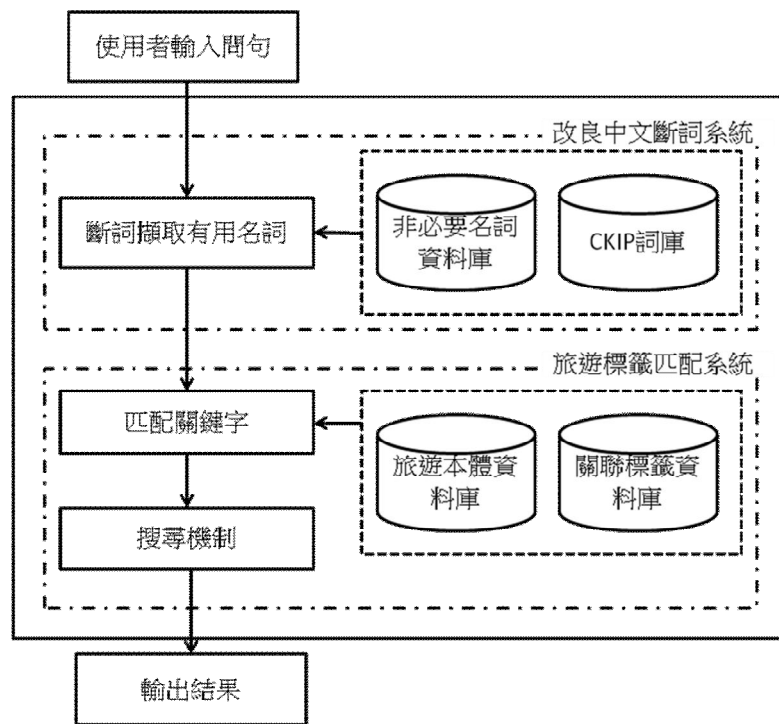


圖 4-1 系統架構圖

改良中文斷詞系統之架構如圖 4-2。中研院的中文斷詞系統[17]能將句子分詞並且標上對應的詞性。自然語言能表達意圖的通常是名詞以及動詞。本研究針對旅遊標籤進行搜尋，動機是要找旅遊景點資訊；因此名詞最能表達使用者想找的東西，而且一般關鍵字大多以名詞為主，因此我們僅保留一句話中的名詞。但是名詞又可以細分成很多種，不是所有名詞都具代表性，例如代名詞 (Nh)、位置詞 (Ncd)、時間詞 (Nd) 這些詞在查詢旅遊上未必有用，但是，可能可以做為表達意圖的參考資訊。舉例來說「南部哪裡可以賞櫻花」此問句經過中文斷詞系統標記詞性之結果為「南部 (Ncd) 哪裡 (Ncd) 可以 (D) 賞 (VC) 櫻花 (Na)」，其中屬於名詞的有「南部 (Ncd) 哪裡 (Ncd) 櫻花 (Na)」，但是由「哪裡 (Ncd)」一詞，可知使用者的意圖是在查詢旅遊景點的位置。因此「南部哪裡可以賞櫻花」此問句經過改良中文斷詞系統後所得之關鍵字組為「南部、櫻花」。旅遊標籤匹配系統如圖 4-3 所示。本研究將每個步驟所產生的關鍵或標籤字組來進行搜尋，藉此比較傳統的關鍵字搜尋、標籤搜尋、以及搭配本體論與大眾分類法之標籤搜尋之成效。

由於本體建立的工程非常浩大[12]，因此本研究限制為台灣本島地區之景點。關鍵字限制為繁體中文。旅遊本體的分類架構以奇摩旅遊網頁的地區分類方式為主。標籤的來源為奇摩旅遊網頁中的旅遊網誌。選擇奇摩旅遊網誌為資料來源之原因在於奇摩旅遊網誌比較單純，內容大多以旅遊心得為主。目前沒有現成的旅遊標籤本體可用，因此，本研究擷取奇摩旅遊的旅遊網誌為資料來源，並以下列方式建立旅遊標籤本體：

- 根據大眾分類法將奇摩旅遊文章的標籤與旅遊的所在地做分類，例如賞櫻、櫻花是阿里山國家風景區的標籤。

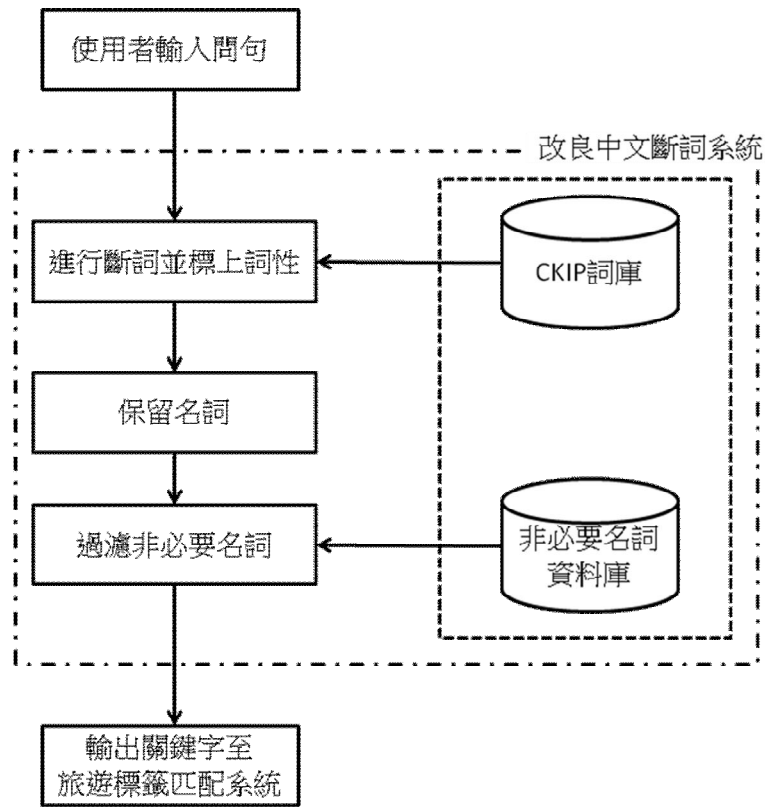


圖 4-2 改良中文斷詞系統架構圖

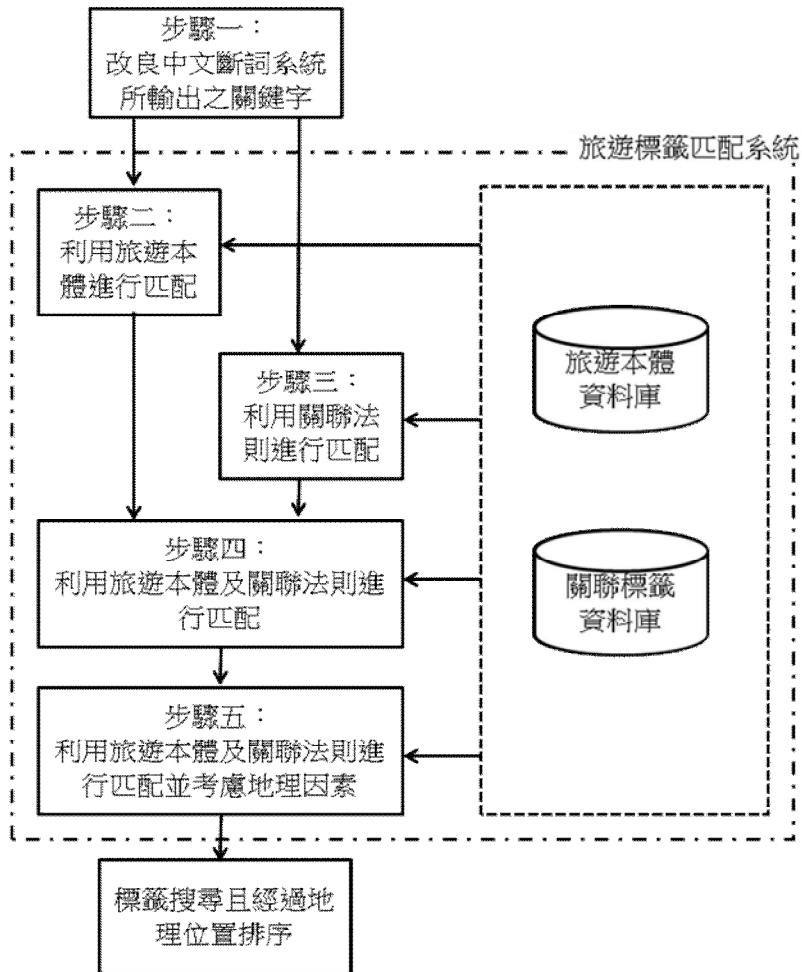


圖 4-3 旅遊標籤匹配系統架構圖

- 根據奇摩旅遊的地理分類，定義出地理名詞之間的階層關係以及繼承關係，例如阿里山國家風景區隸屬於嘉義縣市，其阿里山國家風景區的標籤賞櫻、櫻花同時也屬於嘉義縣市的標籤。
- 根據旅遊本體建立同義字資料表，例如南部、台灣南部皆為南台灣的意思。

當旅遊標籤本體建立完成後，將會有大量的標籤在資料庫中。在這些標籤中可能隱含著我們想要的資訊或是規則，因此我們可以利用資料探勘的關聯法則模型[13]來分析標籤，找出標籤之間的關聯性。利用關聯法則我們可以得知哪些標籤同時出現的頻率較高而視為有關連性。以本研究所蒐集的標籤為例，春天與賞櫻同時出現的機率非常高，因此推論此兩標籤有關連性存在。

當完成改良中文斷詞系統、旅遊標籤本體、以及標籤關聯資料表之後，接著利用標籤匹配系統來做關鍵字的修正、擴展與匹配。例如，輸入關鍵字南部、櫻花至標籤匹配系統，經過第一步驟修正後的標籤字組會是南台灣、櫻花、南部；接著針對這三個標籤，透過第二步驟與標籤關聯資料表比對找出有關連的標籤。其擴展結果為賞櫻、春天、武陵、天元宮、武陵農場、紅粉佳人、淡水、吉野櫻、南台灣、櫻花、南部。透過關聯法則找出的標籤中，很顯然的字數過多，而且出現很多不屬於南台灣地區的標籤，因此，透過第三步驟，利用地理位置資訊來篩選標籤，其結果為賞櫻、春天、吉野櫻、櫻花、南台灣、南部。

本研究將針對奇摩旅遊網誌的標籤進行搜尋，以符合標籤數愈多者優先列出。如果以這六個標籤去搜尋的話，可能會找到一篇被標上賞櫻、春天、吉野櫻、櫻花標籤的日本旅遊網誌，其原因在於其網誌的標籤符合數多，但是這並不是使用者所需的。使用者會特地使用地理名詞，代表使用者有明確的地理目標，因此，最後再依旅遊網誌的地理位置做排序。例如，使用者要找有關台灣南部的櫻花網誌，那搜尋結果列在前面的一定是台灣南部且跟櫻花有關的文章，鄰近南部的中部與東部網誌則接在後面，接下來就是北部及離島的文章，排在最後的是國外旅遊網誌。

五、結果與討論（含結論與建議）

在資訊檢索系統中，其實驗結果主要以準確率（Precision）及召回率（Recall）做為評估的基本指標，其公式如下所示：

$$Recall = \frac{a}{a+d}$$

$$Precision = \frac{a}{a+b}$$

其中 a 為符合使用者需求且有搜尋到的網頁；b 為不符合使用者需求且有搜尋到的網頁；c 為不符合使用者需求且沒有搜尋到的網頁；d 為符合使用者需求且沒有搜尋到的網頁。

根據 iProspect 在 2006 年的一份名為「搜尋引擎使用者使用行為研究（Search Engine User

Behavior Study)」調查中指出：在鍵入搜尋關鍵字後，有 62% 的搜尋引擎使用者，在前 10 筆找到使用者所需要的資訊或網頁；有高達 90% 的使用者，在前 30 筆得到所需要的資訊或網頁。因此本研究將針對傳統的關鍵字搜尋、標籤搜尋以及三種標籤擴展查詢來進行實驗，並且分析搜尋結果之前 10、30、50 筆資料。

本研究假設，如果使用者的關鍵字中含有地理位置，網誌的地點一定要符合使用者所下的地點才算符合使用者需求。判斷方式則以人工的方式來判斷符合的網頁，並計算前 10、30、50 筆資料之準確率為實驗結果。另外，由於本論文是搜尋奇摩旅遊網誌，總資料量過於龐大，無法得到 d 之值，因此不使用召回率。首先，本系統針對使用者所輸入的問句進行斷詞後擷取關鍵字，並將關鍵字出至搜尋引擎，所得到的結果當做對照組。接著，將關鍵字做為標籤，輸入至匹配系統，並將每步驟所產生之標籤字組，分別為旅遊本體擴展、關聯法則擴展以旅遊本體及關聯法則擴展，輸出至搜尋引擎，並且比較三種擴展方式之搜尋結果。

本實驗以下列兩句問句為實驗對象，分別為「問句一：南部哪裡可以賞櫻花」，以及「問句二：秋天哪裡可以賞楓葉」，此二問句之語意文法及改良中文斷詞系統之擷取關鍵字如表 5-1。

表 5-1 問句文法與關鍵字

	問句	語意文法	擷取關鍵字
問句一	南部哪裡可以賞櫻花	南部 (Ncd) 哪裡 (Ncd) 可以 (D) 賞 (VC) 櫻花 (Na)	南部、櫻花
問句二	秋天哪裡可以賞楓葉	秋天 (Nd) 哪裡 (Ncd) 可以 (D) 賞 (VC) 楓葉 (Na)	秋天、楓葉

問句一的每個步驟所對應之關鍵字與標籤如圖 5-1。表 5-2 為問句一之精確度的實驗結果。本實驗之實驗一至六乃使用本系統的標籤搜尋，對照實驗則是傳統的關鍵字搜尋。因此將本系統實驗結果最好的實驗六與對照實驗比較可以看出本系統能夠有效的提升準確率。從實驗一至五中可以發現其準確率皆有所提升，由此可知藉由本體論、關聯法則以及結合兩者的擴展查詢對搜尋準確率的提升皆有效果。

表 5-3 為問句二之精確度的實驗結果。由問句二之實驗結果可以得知，在某些情況下關聯法則對於搜尋之精準度是有提升效果的。像實驗三的關聯法則擴展查詢中，將關鍵字組擴展為「紅葉、賞楓、楓葉、秋天」，從驗數據可以看出準確率確實有所提升，而且比起實驗一多了兩個關鍵字，讓使用者搜尋到的結果更多元。

本研究使用本體論以及關聯法則的擴展查詢方式針對標籤去做搜尋，其結果普遍比傳統的關鍵字搜尋好。其原因在於目前搜尋引擎皆以使用者輸入的關鍵字來進行標題、內文的搜尋，因此，會有搜尋結果不精確以及資料量龐大的問題。透過標籤搜尋，可以能夠改善搜尋精確度的問題，因

為標籤是藉由使用者經過自己經驗為網路資源所賦予的，因此可以代表其資源之意義，比起傳統的關鍵字較符合使用者需求。然而，光是標籤搜尋對精確度的提升仍然有限，而且無法解決搜尋結果的資料量龐大問題，因此，本研究透過本體論以及關聯法則的查詢擴展方式來來改善這兩個問題，從各個實驗結果可以看出本體論搭配關聯法則能提升準確率，並且勝過傳統的關鍵字搜尋。在 2011 年 6 月 3 日的奇摩新聞中得知，搜尋引擎三巨頭打算攜手合作，打算統一標籤格式來增加搜尋結果排序的準確度，這更肯定了標籤對搜尋是有極大的幫助。

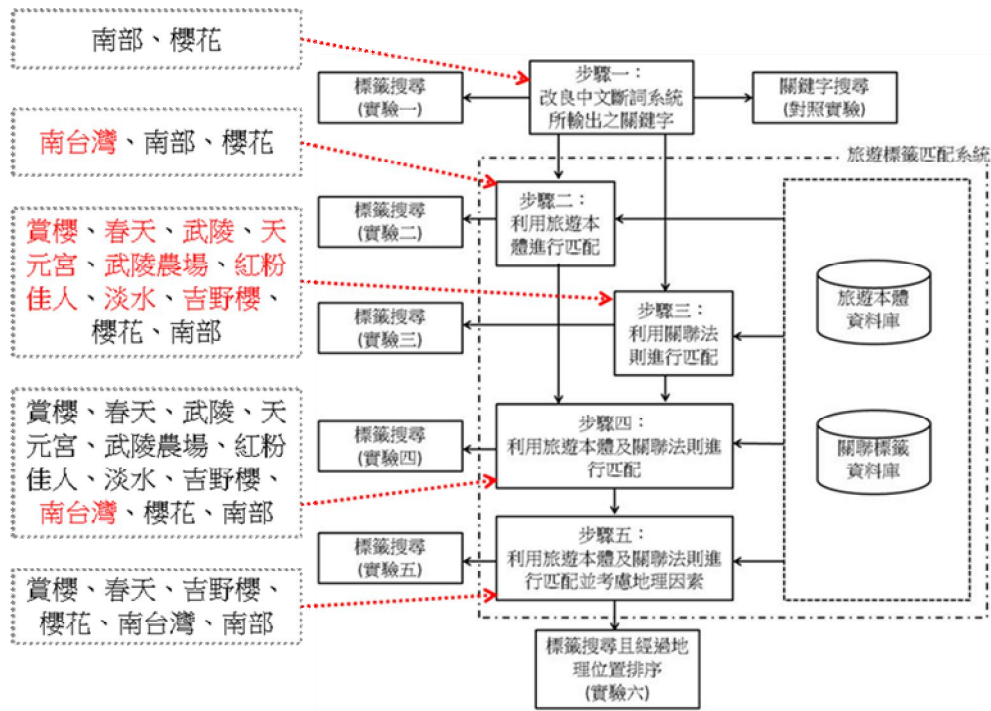


圖 5-1 問句一之實驗流程圖

表 5-2 問句一之精確度

資料筆數	實驗一	實驗二	實驗三	實驗四	實驗五	實驗六	對照實驗
前 10 筆	0%	0%	10%	10%	20%	100%	30%
前 30 筆	7%	7%	10%	10%	10%	77%	20%
前 50 筆	8%	8%	14%	14%	16%	46%	12%

表 5-3 問句二之實驗數據

資料筆數	實驗一	實驗二	實驗三	實驗四	實驗五	實驗六	對照實驗
前 10 筆	30%	30%	50%	50%	50%	100%	20%
前 30 筆	40%	40%	50%	50%	50%	93%	27%
前 50 筆	40%	40%	38%	38%	38%	56%	26%

本研究之實驗結果是以人工判斷準確率的方式來呈現結果，不過判斷者只有一個人，如果可

以讓更多人使用本研究之擴展查詢系統，並以問卷的方式調查使用者使用後之實驗結果，相信此研究結果會更加有說服力。本實驗之準確率是以布林值的方式來區分資料是否符合使用者需求，如果可以計算標籤的詞頻或是其它的加權方式來驗證搜尋的結果，應該會更恰當。另外，本研究是以本體論及關聯法則自動匹配新的標籤，使用者沒有自己選擇標籤的機會，如果以推薦的方式讓使用者可以選擇擴展的標籤，應該會找到更加符合使用者需求的資料。

藉由本實驗所蒐集的資料得知，除了地理位置資訊(Where)對使用者查詢旅遊資訊有很大幫助外，還發現了自然景觀、季節(When)的標籤數量較為龐大，以這兩種類型的標籤建立本體，對使用者將會有更多的搜尋選擇。本研究以手動的方式建立本體，相當的花費時間，如果能夠搭配人工智慧的技術自動建立本體，那本體的領域將可以動態的擴增，甚至可以將圖片、影片等網路資源建立標籤本體，如此一來，使用者意圖將不再只是 Where，而有更多的意圖可以研究。

在本研究整理資料時發現，有特定幾家旅行社會藉由標籤來打廣告。在未來除了現有的關鍵字行銷外，標籤行銷應該也會流行起來，尤其搜尋引擎三巨頭有意要整合網頁的標籤，相信未來的標籤不會像現在雜亂無章，標籤搜尋會更加的實用。

整體而言，本研究完成的工作項目如下所述：(a)建立旅遊領域知識之本體資料庫：我們參考 YAHOO 旅遊網站上的旅遊資訊的分類，同時收集使用者分享的經驗與標籤，以建立旅遊知識之本體；(b)建立大眾分類資料庫：本研究收集使用者分享的標籤，以建立大眾分類資料庫所需的資料。同時針對標籤資料，進行關聯法則分析(Association rules mining)，以儲存高度相關的旅遊標籤組；(c)利用自然語言句型分析模組，擷取出適當的意圖關鍵字與領域關鍵字；(d)完成整個系統的整合和測試，並且透過實驗設計與實驗評估指標，進行搜尋與推薦資料的分析與驗證，探討結合使用者意圖、領域知識本體、與 Folksonomy 在搜尋上所能帶來的效益。本計畫相當龐大，由於研究人力的限制，主要的成果集中在有效的搜尋與結果的排序。未來可以利用使用者過去使用標籤的個人偏好行為，進行個人化推薦的協同過濾方法，以及探討能夠適時推薦具有高度類似特性之文章的內容過濾方法。

參考文獻

- [1] 蘇健華，「WEB 2.0 發展介紹」，資訊社會研究(13)，頁 1-124，2007 年 7 月。
- [2] 湯代祿、韓建俊、連振興，「網際網路變革 - Web 2.0 理念與設計」，松崗電腦圖書有限公司，2007 年 11 月。
- [3] 陳威叡，「Web 2.0 概念下 Folksonomy 導入網路服務之初探—以網路書籤為例」，2008 數位科技與創新管理研討會，頁 72-83.
- [4] 陳品均，「Web 2.0 應用服務策略行動之研究—以 Yahoo!、Google、MSN 為例」，國立臺灣大學商學研究所碩士論文，民國 95 年 7 月。

- [5] 溫達茂、鄭麗敏，「從圖書館與資訊服務的角度看 Web 2.0 的應用」，Web2.0 與圖書館學術研討會論文，民國 97 年。
- [6] 喻英雲，「以本體論結合意圖為基之智慧型搜尋引擎之研究」，南華大學資訊管理學系碩士論文，民國 97 年。
- [7] 賴嘉民，「以 5W1H 為基之醫學本體的搜尋引擎之研究」，南華大學資訊管理學系碩士論文，民國 98 年。
- [8] 卜小蝶，「Web 2.0 與 Library 2.0 簡介」，國立成功大學圖書館館刊，第 16 期，民國 96 年，頁 8-14。
- [9] 卜小蝶，「Folksonomy 的發展與應用」，國立成功大學圖書館館刊，第 16 期，民國 96 年，頁 1-7。
- [10] 林正偉，「Web2.0 環境下以大眾分類法為基礎之知識收藏推薦系統」，國立高雄第一科技大學資訊管理學系碩士論文，民國 97 年。
- [11] 邱英華、陳志豪、梁家松，「以知識本體為基礎建構旅遊行程語意網路服務系統」，資訊管理研究，第 9 期，2009，頁 67-93。
- [12] Guarino, N., "Understanding, building and using ontologies," *International journal of human and computer studies*, 46(3/4), 1997, pp. 219-310.
- [13] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, 2007.
- [14] Jäschke, R., et al., 2008, "Discovering shared conceptualizations in folksonomies", *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 6, No. 1, pp.38-53, February 2008.
- [15] Jessamyn West, "Technology Connection Subject Headings 2.0: Folksonomies and Tags," *Library Media Connection*, 2007, pp.58-59.
- [16] Laura, Gordon-Murnane, "Social Bookmarking, Folksonomies, and Web 2.0 Tools," *Search*, Vol.14, 2006, pp.26-38.
- [17] 中央研究院資訊科學所 CKIP 中文斷詞系統。 <http://ckipsvr.iis.sinica.edu.tw/>。

國科會補助計畫衍生研發成果推廣資料表

日期:2012/10/10

國科會補助計畫	計畫名稱: Web 2.0社群資源分享系統之智慧型搜尋與個人化推薦的研究
	計畫主持人: 邱宏彬
	計畫編號: 100-2221-E-343-007- 學門領域: 資訊系統管理
無研發成果推廣資料	

100 年度專題研究計畫研究成果彙整表

計畫主持人：邱宏彬		計畫編號：100-2221-E-343-007-					
計畫名稱：Web 2.0 社群資源分享系統之智慧型搜尋與個人化推薦的研究							
成果項目		量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）	
		實際已達成數（被接受或已發表）	預期總達成數（含實際已達成數）	本計畫實際貢獻百分比			
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	1	1	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		章/本
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

<p style="text-align: center;">其他成果</p> <p>(無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)</p>	<p>未來準備將研究成果整理成論文，發表在研討會和學術期刊上</p>
---	------------------------------------

	成果項目	量化	名稱或內容性質簡述
科 教 處 計 畫 加 填 項 目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與(閱聽)人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表 未發表之文稿 撰寫中 無

專利： 已獲得 申請中 無

技轉： 已技轉 洽談中 無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

本研究採用 YAHOO 旅遊網頁做為實驗資料來源，以擷取實驗所需資料，建立旅遊知識本體以及大眾分類等資料庫，進行實驗的分析與驗證，探討結合使用者意圖、領域知識本體、與 Folksonomy 在搜尋與推薦上所能帶來的效益。本研究所提出的方法可實際應用在各式社群資源分享網站中，如 Del.icio.us 推薦書籤、Flickr 推薦相片、YouTube 推薦影片、部落格入口網推薦部落格、Amazon 推薦商品、企業內部運用標籤來標記文件的知識管理系統等，應用將十分廣泛。本計畫相當龐大，由於研究人力的限制，主要的成果集中在有效的搜尋與結果的排序，儘管如此，本研究成果已經相當程度符合原計畫的預期目標。未來將進一步利用使用者過去使用標籤的個人偏好行為，進行個人化推薦的協同過濾方法，以及探討能夠適時推薦具有高度類似特性之文章的內容過濾方法。我們將整理本研究目前的成果，以便發表在研討會與學術期刊中。