

# 國立交通大學

資訊學院 數位圖書資訊學程

碩士論文

圖書館期刊使用統計之系統設計與實作：

以博碩士論文之參考文獻為例

Design and Implementation for Statistical Citation  
Analysis of Journal Papers: A Study of References  
of Dissertations and Theses

研究生：王冬祥

指導教授：黃明居 教授

中華民國一〇〇年一月

圖書館期刊使用統計之系統設計與實作：

以博碩士論文之參考文獻為例

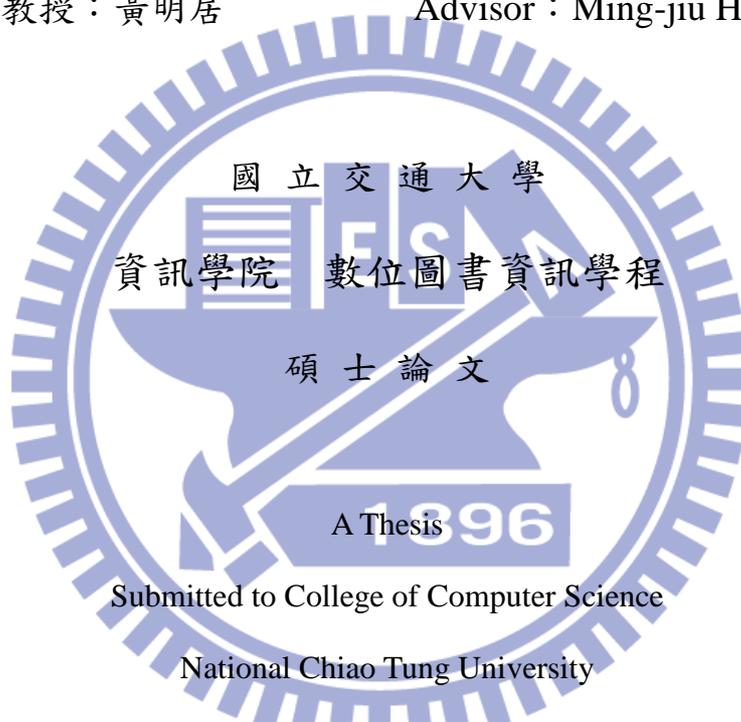
Design and Implementation for Statistical Citation Analysis of  
Journal Papers: A Study of References of Dissertations and Theses

研究生：王冬祥

Student：Tung-hsiang Wang

指導教授：黃明居

Advisor：Ming-jiu Hwang



Submitted to College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of Master of Science

in

Digital Library

January 2011

Hsinchu, Taiwan, Republic of China

中華民國一〇〇年一月

# 圖書館期刊使用統計之系統設計與實作： 以博碩士論文之參考文獻為例

學生：王冬祥

指導教授：黃明居博士

國立交通大學 資訊學院 數位圖書資訊學程碩士班

## 摘 要

探討期刊資源被使用的情況，一直以來都是圖書資訊學領域的一項重要研究，圖書館也利用各種方式想瞭解期刊資源的使用情況。針對校內的研究產出，利用論文的參考文獻來進行統計分析，是一種以使用者為導向，反應實際研究及使用需求有效的方法。但進行相關的分析研究必須仰賴大量的人工方式，蒐集、整理、過濾出相關的資訊，是此類研究中很大的負擔。

本研究以參考文獻計量為基礎，設計建置一套期刊使用統計系統(System for Citation Analysis of Journal Papers, SCAJ)。系統區分為六大模組：系統模組、刊名資料庫模組、引文模組、統計期刊清單模組、共享模組、以及統計作業模組。使用者可將參考文獻資料匯入，以進行分析作業，系統即可產出文獻類型，以及西文期刊的引用次數統計結果。此外，使用者亦可匯入期刊清單，與參考文獻統計結果進行進一步的統計，利用自訂的權重進行統計，最後產出可供參考的統計數據，協助圖書館做為評估期刊使用情況的參考。

SCAJ 可廣泛應用於參考文獻資料的統計工作，協助解決參考文獻著錄凌亂問題，並減少人工作業時間。用於圖書館的期刊資源使用狀況時，可利用使用者的論文產出，統計論文的期刊引用，回推圖書館期刊資源的使用狀況，協助研究圖書館期刊使用率的問題。

關鍵詞：期刊使用統計系統、文獻計量、期刊使用率

# Design and Implementation for Statistical Citation Analysis of Journal Papers: A Study of References of Dissertations and Theses

Student : Tung-hsiang Wang

Advisors : Dr. Ming-jiu Hwang

Degree Program of Computer Science

National Chiao Tung University

## ABSTRACT

The analysis of journal usage has always been an important research topic in Library and Information Science. Various methods have been used by the libraries to understand the usage of journal papers. Citation analysis of references of research publications is a user-oriented method which may efficiently show the actual demand of journal resource. However, it used to require a lot of time and effort to collect, organize, and analyze the information, which resulted in a heavy burden on this kind of research.

Based on the metrological analysis of the references of dissertations and theses, a System of Citation Analysis of Journal Papers (SCAJ) is proposed. The system is divided into 6 modules: system module, module of journal titles, citation module, module of statistic inventories of journals, partaking module, and statistic module. Through SCAJ, users can get the information about the categories of literatures and citation times of English journal papers. Furthermore, users can enter journal lists and self-set weighting index to do extending citation analysis. The final outcome is statistics that are worth referring to and can be used to evaluate the usage of journals by the libraries.

SCAJ solves the problem caused by confusing and disordered references and reduces the cost of labor and time. The utility rate and citation frequency of journal papers and can be precisely estimated by using SCAJ to analyze the references of dissertations and these.

Key Word : System for Citation Analysis of Journal papers (SCAJ), Bibliometrics, Utility rate of Journal papers

## 誌 謝

終於完成論文了，在進研究所的這些日子以來，總是工作、學業蠟燭二頭燒，但總算完成學業了，當學生的日子也告一段落。

論文能順利完成，首先要感謝我的指導教授 黃明居老師，感謝您在這段時間的悉心指導，以及包容我有時候因為工作而無法將心力用在學業上，讓我在論文研究的過程中，學習到該領域相關的知識，最重要的是整個研究過程中的自我挑戰。還有口試時柯皓仁教授及張東森教授的指導及寶貴意見，讓我的論文能更完善。

感謝對一路走來幫助以及關心過我的人，心裡真的懷著無限的感激。最後感謝我最重要的家人，我的父母親，以及在求學的這段時間跟我一起步入紅毯的老婆詩婉，感謝你們一直在背後支持我、鼓勵我，謝謝。



# 目 錄

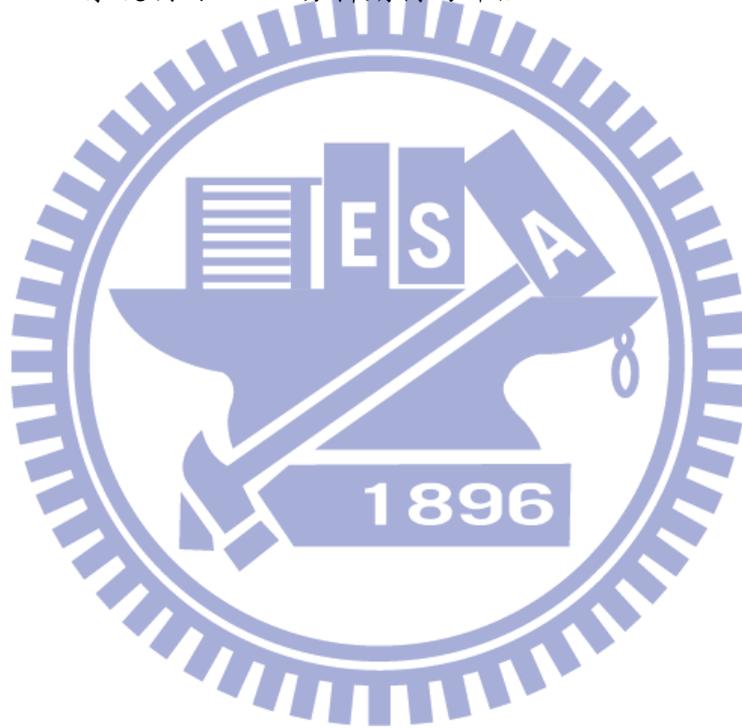
第一章 緒論.....	1
1.1 研究背景及動機.....	1
1.2 研究目的.....	2
1.3 研究範圍與限制.....	3
1.4 研究流程.....	4
第二章 文獻回顧.....	5
2.1 期刊使用分析.....	5
2.1.1 期刊分析方式.....	6
2.1.2 期刊分析系統.....	9
2.2 學術論文引文格式.....	13
2.3 相關資訊技術.....	15
2.3.1 近似字串比對.....	15
2.3.2 Web 服務.....	15
2.3.3 ASP.NET.....	17
第三章 系統分析與設計.....	19
3.1 系統需求分析.....	19
3.2 設計概念.....	22
3.3 系統設計.....	23
3.3.1 系統實體架構.....	24
3.3.2 系統作業流程.....	26
3.3.3 系統模組及功能.....	34
3.3.4 參考文獻分析功能.....	38
3.3.5 WOS Web 服務.....	43
第四章 系統實作與展示.....	49
4.1 系統環境建置.....	49
4.2 系統展示與功能介紹.....	51
4.2.1 系統模組.....	51
4.2.2 刊名資料庫模組.....	52
4.2.3 引文模組.....	55
4.2.4 統計期刊清單模組.....	61
4.2.5 共享模組.....	63
4.2.6 統計作業模組.....	63
第五章 系統評估.....	72
5.1 需求完成度評估.....	72
5.2 資料分析評估.....	73
5.2.1 系統評估資料.....	73

5.2.2 西文期刊分析評估.....	74
5.2.3 WOS 分析評估 .....	80
第六章 結論與建議.....	83
6.1 研究結論.....	83
6.2 研究建議.....	84
6.2.1 研究結果建議.....	84
6.2.2 未來研究建議.....	85
參考書目.....	86



## 表目錄

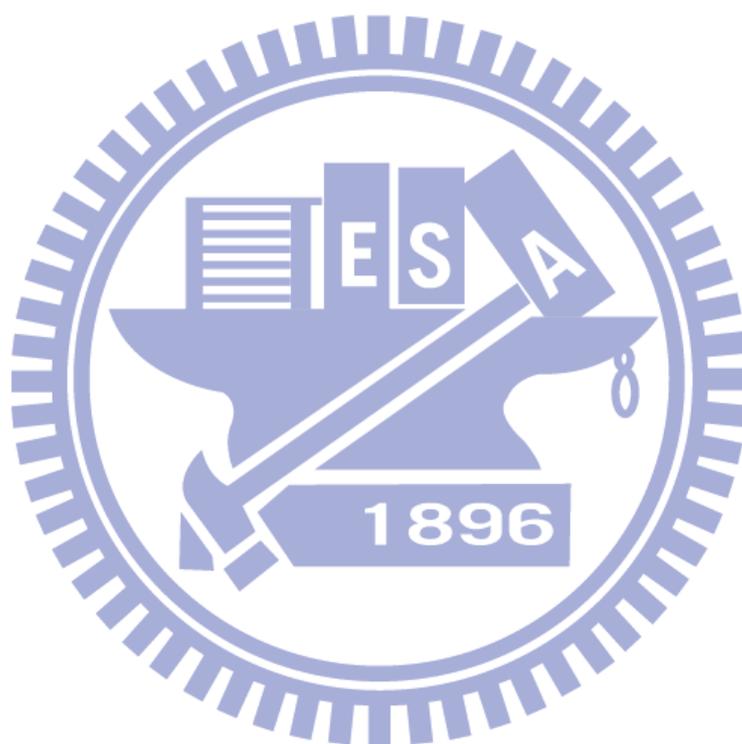
表 3-1	系統模組使用權限 .....	37
表 3-2	WOS Web 服務 searchRetrieve 參數說明 .....	45
表 3-3	WOS Web 服務回傳 XML 內容 .....	46
表 4-1	上傳引文資料儲存比較 .....	55
表 5-1	需求完成度評估 .....	72
表 5-2	系統評估資料 .....	74
表 5-3	西文期刊分析正確率 .....	75
表 5-4	西文期刊分析比率 .....	77
表 5-5	西文期刊未成功分析的分佈 .....	79
表 5-6	系統利用 WOS 分析期刊的評估 .....	81



## 圖目錄

圖 1-1	研究流程 .....	4
圖 2-1	研究文獻關係 .....	5
圖 2-2	Web Services 架構 .....	17
圖 3-1	系統需求面向 .....	19
圖 3-2	系統設計概念 .....	23
圖 3-3	系統架構 .....	24
圖 3-4	系統實體架構 .....	25
圖 3-5	系統作業流程 .....	26
圖 3-6	刊名資料庫作業示意 .....	27
圖 3-7	刊名資料庫作業流程 .....	28
圖 3-8	引文資料作業流程 .....	29
圖 3-9	引文資料校正作業流程 .....	30
圖 3-10	期刊清單作業流程 .....	31
圖 3-11	排名比較統計作業流程 .....	33
圖 3-12	系統模組 .....	35
圖 3-13	系統功能 .....	36
圖 3-14	參考文獻分析處理流程 .....	42
圖 3-15	存取 WOS Web Service 網路架構 .....	44
圖 4-1	系統環境建置架構 .....	50
圖 4-2	系統權限管理 .....	51
圖 4-3	刊名清單上傳功能 .....	52
圖 4-4	刊名資料庫管理功能 .....	53
圖 4-5	新增刊名驗證功能 .....	54
圖 4-6	引文資料上傳功能 .....	58
圖 4-7	引文資料校正 .....	59
圖 4-8	引文資料校正作業 .....	60
圖 4-9	統計期刊清單功能 .....	62
圖 4-10	共享模組功能 .....	63
圖 4-11	引文統計報表選擇 .....	64
圖 4-12	引用次數統計結果 .....	65
圖 4-13	分類統計結果 .....	66
圖 4-14	合併統計選擇 .....	67
圖 4-15	合併統計共享資料選擇 .....	68
圖 4-16	合併統計報表結果 .....	69
圖 4-17	排名比較功能 .....	70
圖 4-18	排名比較結果 .....	71

圖 5-1	西文期刊分析比率 .....	78
圖 5-2	西文期刊未成功分析的類型比率 .....	80
圖 5-3	WOS 刊名程序比對結果 .....	82
圖 5-4	WOS 篇名程序比對結果 .....	82



# 第一章 緒論

本研究期能利用資訊技術，設計出一套期刊使用統計系統（System for Citation Analysis of Journal Papers, SCAJ），以幫助研究人員以及圖書館在論文的參考文獻計量作業時能減少人工作業，以較節省人力、時間的方式進行相關統計及應用。本章就以研究的動機及背景、目的、以及流程，分述如下。

## 1.1 研究背景及動機

圖書館購置之學術期刊在各類學術資源中，常是研究時的重要參考來源，它提供了最新的研究發展、研究成果及未來發展等等的新知、新理論，讓學者在研究時借鑒參考文獻來獲取新知，對所研究加以印證，並從中啟發靈感，產生新的理論，也是學者在發表研究成果及學術表現的重要平台，由此可知，期刊在學術領域的重要程度。然而圖書館在資料的採購經費，期刊占了相當大的比例，以美國大學圖書館為例，每年採購資料預算的 30% 至 45% 用於購買期刊，有些學術性的專門圖書館之期刊費，甚至高達 80% 至 90%。由於價格昂貴，佔用了圖書館書刊經費的極大比例，圖書館在決定購買與否時通常需要了解其使用狀況，並考量其成本效益，以便把經費用在購置讀者最需要、最常用的資料上，使有限的資源能得到最有效的運用（詹麗萍，民 95）。

探討期刊資源被使用的情況，一直以來都是圖書資訊學領域的一項重要研究。研究的結果除了可以讓圖書館做為刪訂期刊以及管理期刊的參考依據，還可以讓學生、教師以及研究人員做為研究以及投稿時的參考。然而，期刊也並非一成不變，學術的研究會隨著時間改變，產生新的研究重點、新的研究領域，甚至一門新興的學科，連帶影響期刊的停刊、創刊、合併、分輯、改名等現象不時出現，同時期刊在出版的質量與水準也會隨時產生變化，所以，圖書館研究期刊資源的使用，是必須不斷持續下去，才能因應期刊的變化。

就大學圖書館而言，研究期刊使用的方法不外乎為問卷調查、使用調查等使用分析方法。期刊使用分析雖然可以反映期刊使用情形，但是其中多數方法多半容易受到圖書館管理方式、調查過程使用者的配合度、調查時間的長短及時間點等因素的影響，使得分析結果不易分析與解釋，反而容易失去公正性與可靠性。意見調查法的優點在調查問題可以不只侷限於目前既有的館藏，並可得知讀者的直接回饋。但是此法必需配合設計相當完善的問卷內容，而且只能調查出受訪者的期望和意見，及其所能想起之問題，造成統計結論與真實情況可能因此有所差距（石美玉，民92）。

然而，針對校內的研究產出，例如教師的研究論文及博碩士學位論文等，利用論文的參考文獻來進行統計分析，是一種以使用者為導向，反映實際研究及使用需求有效的方法。雖然目前國內各大學圖書館大都已建置機構典藏系統或者是博碩士論文數位典藏系統，但在論文的參考文獻部份，仍僅止於將論文的所有參考文獻資料當成一筆文字資料，存在一個資料庫欄位內，也就是對系統來說，參考文獻就是一筆字串，並無任何的統計價值，加上部份的作者並未依規定或常見的學術論文引文格式進行著錄，造成參考文獻的資料極為凌亂。雖然圖書館擁有典藏系統的參考文獻資料，但要利用它來進行相關的分析研究卻困難重重，無法利用電腦處理凌亂的著錄格式，必須仰賴大量的人工方式，蒐集、整理、過濾出相關的資訊，是此類研究中很大的負擔，要透過館員或研究人員持續的建立及更新研究更屬不易。因此，如何設計與開發一個實用性高的系統，輔助館員或研究者瞭解圖書館期刊使用之情形，並解決上述問題，為本研究最重要的動機。

## 1.2 研究目的

基於上節研究背景及動機的探討，本研究將利用參考文獻計量為基礎，設計出一套可協助研究人員或圖書館，進行論文參考文獻研究時較易於分析的系統，得以改善因沒有適用的系統或工具協助，導致需使用人工的方式做資料的分析、

比對、以及統計的困擾，甚至因需要耗費大量人力的狀況下，就不得不縮小研究規模，以達到人力可進行的範圍。

總結上述說明，本研究之具體目標如下：

- 一、設計與實作，參考文獻計量為基礎的期刊使用統計系統 (SCAJ)，讓研究人員或圖書館人員可以自由應用。
- 二、透過系統輔助，針對校內的研究產出，回推瞭解圖書館期刊使用率問題。
- 三、利用系統做為參考文獻計量的輔助工具，並且解決文獻著錄凌亂的問題。
- 四、透過系統做為輔助工具，來達到減少人力、時間，以及擴大研究範圍。
- 五、透過系統產出的數據，做為期刊評估或後續研究的參考。

### 1.3 研究範圍與限制

本研究主要是以設計與實作期刊使用統計系統 (SCAJ) 為目的，故其範圍及限制如下：

#### 一、以引文資料為分析內容

主要是以匯入的引文資料來分析其來源期刊，並計算期刊的引用次數，以做為分析期刊的量化指標。所以，本系統僅處理參考文獻的引文資料。

#### 二、匯入分析資料格式

匯入上傳的引文資料檔案，僅以單純引文資料之 Text 檔格式或者 XML 格式儲存的博碩士學位論文引文資料二種格式為主。

#### 三、引文資料類型分類

本研究所指之被引用文獻資料類型的項目分為，西文期刊、會議論文、網路資源、中文期刊、圖書、博碩士論文、研究報告、以及其它。

#### 四、以西文期刊為分析內容

由於西文期刊價格非常昂貴，分析西文期刊的使用情況為各研究單位的重要需求，故此研究之範圍，還是以分析文獻參考來源中的西文期刊為主。

## 1.4 研究流程

本研究目標在設計 SCAJ 系統，並實作建置完之後進行系統的測試評估，最後針對系統設計與實作過程，以及最後測試評估結果提出結論以及建議。詳細之研究流程如圖 1-1 研究流程。

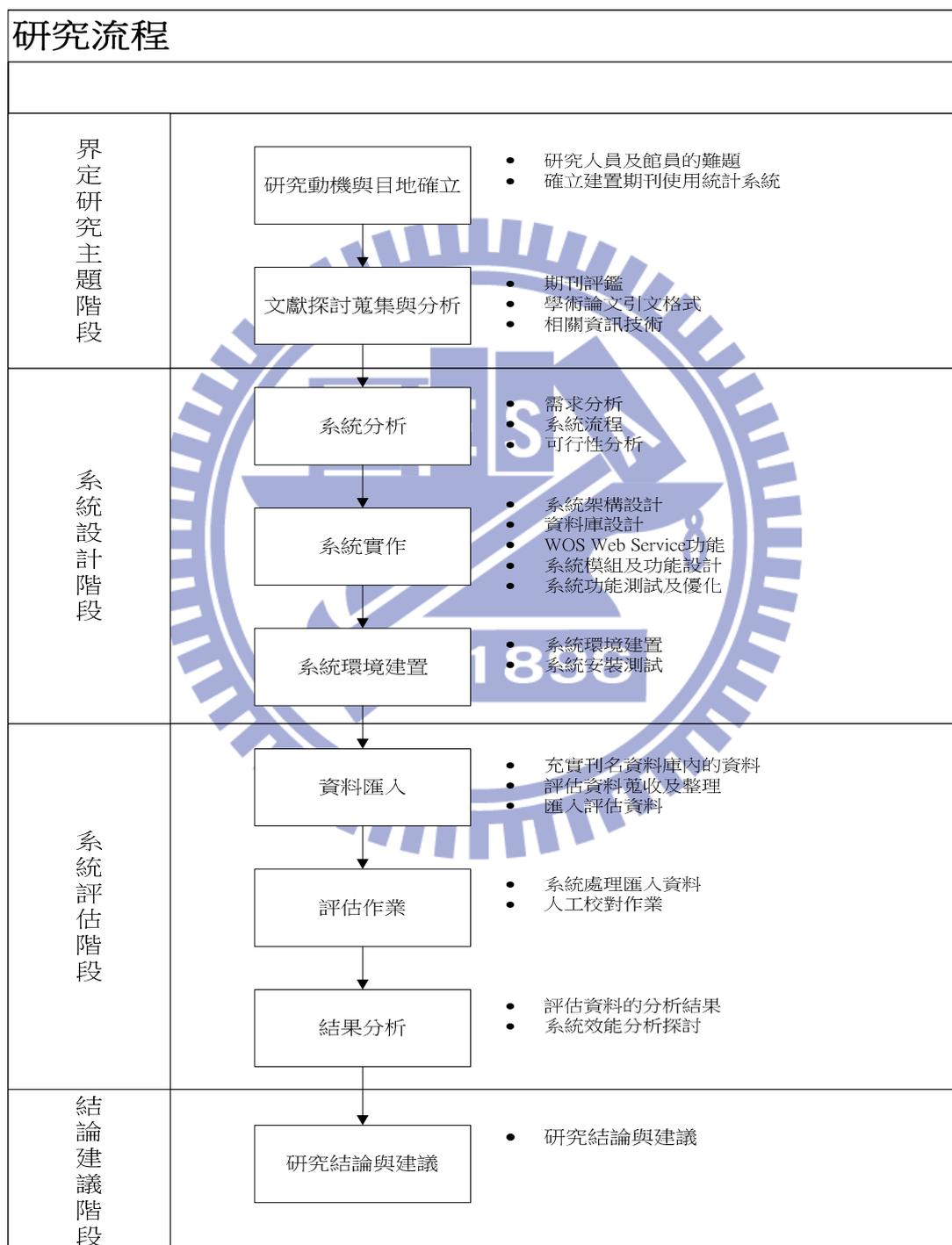


圖 1-1 研究流程

## 第二章 文獻回顧

依據研究背景、動機、以及研究目的，進行期刊分析系統及相關資訊技術、引文分析、期刊分析、以及引文格式分析，四個面向的相關文獻進行探討。如圖 2-1 研究文獻關係所示，針對系統設計實作的相關知識，以及資訊技術，期刊的分析方式、常見用於分析期刊的系統，學術論文的引文格式以及系統相關的資訊技術進行回顧介紹。

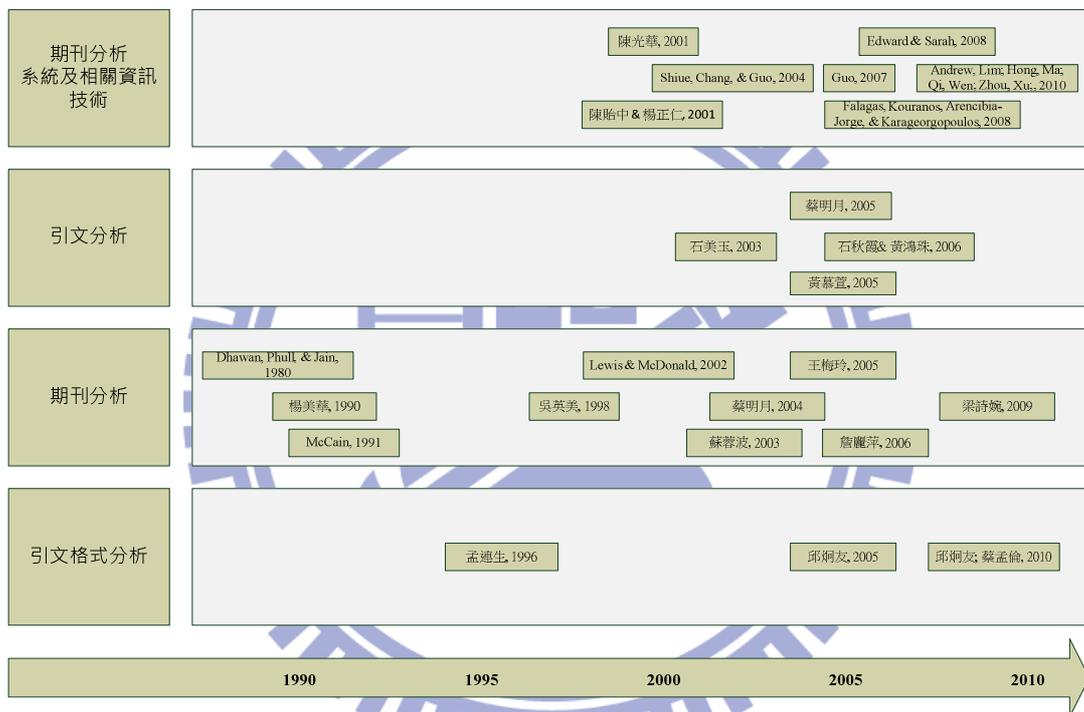


圖 2-1 研究文獻關係

### 2.1 期刊使用分析

期刊使用分析的目的不外乎是為了瞭解期刊的品質、使用狀況、是否符合使用者需求等。本節針對期刊使用分析進行回顧目前國內外常使用的分析方式，以及常用來輔助期刊分析的資訊系統，並介紹資訊系統所提供的分析指標。

## 2.1.1 期刊分析方式

期刊為定期而繼續出版的刊物，根據英美編目第二版（Anglo-American Cataloguing Rules, 2nd Edition, AACRII）的定義：「是指有編號或年月編次標示，意欲無限期發行者，包括期刊、報紙、年刊（報告、年鑑等）、學會會報、紀要、議事錄等、以及有編號的出版品」。（胡述兆、王梅玲，民 92）。期刊是研究時事、科技新知、最發展與未來趨勢的最重要資源，凡在專書上網法找到或尚未論及的新理論、新發明、新概念與新趨勢，常仰賴期刊論文提供資料。而期刊的發行量雖然大，但根據研究，有 35% 的期刊論文從來沒有被引用過，而 49% 的期刊論文只被引用過一次，只有 16% 的期刊論文被多次引用（石美玉，民 92）。可見慎選切合使用者需要以及有用的期刊是極為重要。

國內外對於期刊分析評選的方式並無特定的規範，在許多文獻也指出，沒有一種單獨的標準可以做為期刊的價值或重要性的指標，因此需因應環境以及條件不同來選擇適當的期刊分析評選方式。分析評選方式大致可分為主觀及客觀二種，綜合目前國內外常用的主觀期刊分析評選方式如下：

### 一、意見調查法

意見調查法的優點在調查問題可以不只侷限於目前既有的館藏，並可得知讀者的直接回饋。

### 二、專家諮詢法

此法為根據專家意見來選擇核心期刊。具體方式為條列出某一學科的專業期刊或相關期刊清單，提供教師及研究人員評估，請他們圈選出在教學及研究活動中，利用率最高的幾種期刊，再綜合它們的意見。此法之優點為納入最了解在其學術領域中那些期刊為必需之專家學者的意見，可切合核心使用者群之真正需要。

客觀的期刊分析方式則有下列方式：

### 一、引文分析法

引文分析法是以期刊文獻所列之參考書目為研究對象，以統計各種期刊上發表的一次文獻被其他文獻引用的多寡。

## 二、載文法

可分為載文量法與載文率法兩種。

(一) 載文量法利用布萊德福定律，將某一主題下之所有期刊，依其所含相關論文篇數的多寡可以找出某一學科的核心期刊，或是決定要獲得某一比例的文章需要多少種的期刊，並且可以成為期刊淘汰的依據。但此一方法對載文量多，或刊期較頻繁的大型期刊，將佔有較大的優勢。相對的，對載文量少，但內容符合需要之小型期刊，便極有可能被排除在核心期刊之外。為補強這個缺點，必需同時使用所謂載文率法配合分析。

(二) 載文率法係計算期刊刊載某一學科或專業的相關文獻數，佔該期刊全部載文量的百分比，國外研究文獻曾提出若相關主題的論文篇數在 50% 以上者，即顯示此份期刊對此一主題的研究貢獻大。

## 三、索引摘要法

有別於載文量法及載文率法需要大量的期刊作為分析樣本群，索引摘要法僅取一種或數種在某一學科或專業領域中權威性較強，收錄較齊全的索引摘要刊物作為分析統計之依據。

## 四、使用分析法

直接分析現有館藏的使用狀況，是各種期刊評鑑方法中相當實用及重要的一種，可以直接反應使用者的使用狀況，但會因使用者的配合程度以及人為因素而造成分析的誤差。一般的期刊使用分析採行之方法大致上有：

### (一) 出納記錄

直接計算期刊被使用次數，但對於期刊不外借之單位，此法恐受到限制。

### (二) 調查單

於期刊上黏貼調查單，請讀者於使用資料後，於調查單上作記號。

### (三) 讀者問卷調查

在讀者離開圖書館之際，請讀者勾選曾經使用過的期刊。

### (四) 重新歸架計算

要求讀者於取出期刊閱覽完畢後，將期刊置於閱覽桌上或指定之地方，館方於一定時間內，將讀者留下之期刊加以記錄分析。

### (五) 複印記錄

由讀者申請複印之記錄單，統計使用次數。

### (六) 館際合作申請記錄

利用館際合作申請記錄，可了解讀者需要的，但目前館內沒有的期刊的狀況，但對館藏期刊之使用情形則無法提供資訊。

### (七) 電子期刊使用統計

近年電子期刊之發展，多數之電子期刊系統已可提供電子期刊之使用統計，直接採用亦是相當重要的參考。

## 五、使用費率法

本法是考慮期刊被使用的經濟效益，藉由統計各期刊的使用次數，將期刊價格除以使用次數，可以得知平均使用一次期刊所需付出的成本用。

## 六、標準期刊清單法

以權威性之團體所推薦之期刊為標準，例如以 Index Medicus 評估生物醫學期刊，以科學引文索引 (Science Citation Index, SCI) 評估科學技術各類期刊，或如其他著名圖書館之館藏期刊清單，其他組織或圖書館已完成評鑑之核心期刊清單等均可作為參考。中外有許多圖書館直接採用期刊引用報告 (Journal Citation Reports, JCR) 為衡量工具，JCR 將 SCI 及社會科學引文索引 (Social Sciences Citation Index, SSCI) 兩大引文索引資料庫所收錄超過三年以上之期刊，依其被引用次數 (Total Cites)、影響係數 (Impact Factor)、

立即索引（Immediacy Index）等分別列出名次，為期刊的選擇提供了重要的參考依據。

#### 七、多要素分析法

為避免使用單一評鑑方法造成偏頗之結果，而將多種評鑑方法綜合使用。有逐法篩選者，亦有數法同時使用，並賦與不同權重以統計排序者。具代表性者如 Dhawan, Phull, & Jain（1980）出之綜合模式，該模式在確認學科領域期刊之質與量之際，也同時結合實際利用情況與讀者意見來確認核心期刊的正確性。

綜觀國內外相關研究採用的期刊評鑑模式，石美玉（民 92）針對近年來的國內外較著名或有特色的研究實例進行分析，在國內的六篇相關研究針對其評鑑方法統計，多要素分析法為所有個案都採用，顯見絕大多數國內個案皆認同單一期刊分析評鑑方法容易造成結果之偏頗，引文分析法則僅有一案採用，標準期刊清單則有 80% 的案例採用。針對國外的六篇相關研究法頗為分散，僅有引文分析是共同採用之方法。就目前所蒐集到的文獻可以看出只要是在小範圍或單一主題、單一圖書館之核心期刊研究，引文分析法便經常被使用，在資料量不大之狀況下，同時搭配其他期刊分析評鑑方法以補足引文分析法可能之缺失，如此所得之結果是可以較不引起爭議的。

### 2.1.2 期刊分析系統

儘管國內外對於期刊分析的方式有許多，所涵蓋的深度以及廣度也不盡相同，但對於沒有充裕時間或人力的研究人員及圖書館館員而言，較廣較深的期刊分析範圍及方式就產生較大的問題。當可以用於協助期刊分析評鑑的系統出現之後，理當成為最直接便捷的參考來源，甚至於直接當成評鑑期刊的主要依據，以下就幾個可用於期刊分析評鑑的系統進行簡單介紹：

#### 一、期刊引用報告（Journal Citation Reports）

由 Institute for Scientific Information (ISI) 發行，其引用資料來自 Web of Science (WOS) 資料庫中 8,000 多種期刊，包含 200 多個學科全球 3,300 多家出版商，利用自然科學、社會科學領域的所有專業的期刊積累和列出引用數和文章計數，對期刊進行評價和比較。它是提供期刊數據以評估期刊的工具，做為期刊的品質評估及作者生產力與影響力的參考。針對 JCR 提供的期刊評鑑指數概述如下：

### (一) 影響係數 (Impact Factor, IF)

Impact Factor 為目前最多使用者參考的期刊評鑑指數，主要計算方式為該期刊前二年出版的文章在此今年平均被引用的頻率。例如：2010 年的影響係數，即為該刊 2008 年與 2009 年出版的文獻在 2010 年被引用的總數，除以該刊 2008 年及 2009 年出版的文章總數。其計算公式如下。

$$IF = \frac{\text{該期刊近二年出版之文章在當年的被引用次數}}{\text{該期刊近二年出版之文章數}}$$

### (二) 五年影響係數 (Five-Year Impact Factor)

其概念跟影響係數相同，只是將用於計算的引用時間由二年拉長至五年，其原因為有些領域需要較長的時間才能達到引用的高峰。

### (三) Eigenfactor Metrics

Eigenfactor Metrics 為 2009 年新版 JCR 所提供的新功能，包括 Eigenfactor Score 和 Article Influence Score。由華盛頓大學 (University of Washington) 發展出來的指標，其計算原理類似 Google Page Rank 的概念，Page Rank 是計算連結的網頁，而 Eigenfactor 則是計算引文，但不只計算引文的數量，更重要的是計算引文的來源。其特色為於計算的引用時間範圍為五年，忽略期刊的自我引用，將社會科學以及自然科學期刊一起計算，也加入了博士論文、新聞、以及一般期刊來更客觀的評估期刊的重要性，透過期刊引用網路社群反映期刊的權威性和引文的影響

力，打破原先影響指數引用次數權重均等的概念。

#### (四) 即時引用指數 (Journal Immediacy Index)

用於計算某期刊在該年度多快被引用的方法。也就是說 2010 年的即時引用指數，就是以 2010 年該期刊的文獻被引用數除以該刊 2010 年出版的論文總數。此指數與影響係數最大的差別，在於計算的時間分界，此指數直接顯示期刊當下的影響力，較適合觀察更新快速的學科。

## 二、Scopus

由全球超過 500 位使用者以及圖書管理人員所共同設計發展而成，包含來自全球 5,000 個出版公司，涵蓋超過 18,000 種期刊之摘要及參考書目，以確保其廣泛的跨領域性。它提供美國地區以外的資料，涵蓋範圍廣，大量收錄從 1996 年以後的期刊與論文。Scopus 提供的期刊評鑑指數概述如下：

#### (一) h-index

h-index 是在 2005 年由加州大學聖地牙哥分校 (University of California in San Diego) 的凝態物理學教授 Jorge Hirsch 所提出 SJR 指標，主要目的是藉由 h-index 來描述科學家的研究產出之影響與質量。其計算方式為一個科學家所發表的文章中，有 h 篇文章被引用次數大於等於 h，而其他文章被引次數皆小於等於 h，則其 h-index 指數為 h。在 Scopus 也提供計算一群文獻或期刊之 h-index。

#### (二) SCImago Journal Rank (SJR)

SJR 為西班牙的組織機構評鑑公司 SCImago 所提出，用於衡量學術來源的科學聲望，即每篇論文的加權引用價值。SJR 在 2008 年得到了 Nature 的報導被認為是對 ISI 的期刊影響係數強有力的挑戰，其演算法也是類似 Google Page Rank 的概念，在賦予分值的時候考慮被引期刊的級別，某一期刊在被另一期刊引用時將賦予該期刊其自身的聲望或地位，也就是說引用具較高 SJR 的來源比引用具較低 SJR 的來源更有價值。

### (三) Source Normalized Impact per Paper (SNIP)

SNIP 為萊登大學 (Leiden University) 技術研究中心的 Henk F. Moed 教授所提出。其特色為衡量期刊的上下文引用影響，它可依據作者引用其他論文的次數、引文影響成熟度，用於評估的資料庫對學科領域所有文獻的覆蓋程度。三年引文期中某一來源出版物中的每篇論文，平均被引用次數與該學科領域的引文潛力 (Citation potential) 之間的比值。引文潛力係指一篇文章估計在指定的學科領域中，達到的平均被引用次數。

### 三、Journal-ranking

Journal-ranking.com 是一個免費註冊使用線上的期刊評鑑服務，目前收錄超過 7,000 種期刊。由香港科技大學的林良材教授所建立，Center for Journal Ranking (CJR) 運作該網路服務。它也提供了期刊的評比指標，期刊影響指數 (Journal Influence Index) 來供使用者參考，跟 JCR 的 Eigenfactor 以及 Scopus SJR 一樣，其概念都來自於 Google Page Rank，引文網路結構決定該引用的品質權重，使文獻的品質不單只由引用的數量決定。

上述的三個期刊分析評鑑系統，都是以引文文獻分析為基礎，較新的分析評鑑指標則再配合上，引文網路結構決定該引用的品質權重。引用文獻分析即使有許多的參考價值但還是存在著有諸多爭議，在眾多的批評中，以自我引用、引用錯誤、引用均等以及二次引用為多。引文分析雖然存在著不少干擾變數，卻仍在資訊計量學甚至資訊科學學域中，昂首佇立，自成一家，各方研究欣欣向榮。究其原因，不外是：沒有一種既簡單、直接又實際的技術可以掌握所有文獻的使用情形；價值的判斷早在文獻被引用時即已完成；可以容易且精確的選擇引用的數據；引文分析是唯一對作者使用文獻最容易觀察到的指標 (蔡明月，民 94)。

總之，引用文獻資料庫提供豐富的各項的量化數據以及評比指標，但該數據應視為參考之用，而不應視為評估期刊品質的唯一指標，應用時搭配其它的分析評鑑方式，才能分析評鑑出真正有高品質及有用的期刊。

## 2.2 學術論文引文格式

在學術論文的寫作當中，通常會引用前人的研究成果或觀念來提高論文的信服力，藉由引用文獻表明對前人研究成果的承認、借鑒、繼承、修正、反駁、批判或是向讀者提供更進一步研究的參考線索等，於是著文與引文之間建立起一種引證關係（孟連生，民 85）。而讀者則可以藉由引用文獻來延伸對該文章閱讀，由文章之間建立起來的引證關係來探索相關的知識網路，讓研究範圍可以延伸的更深更廣。因此引文格式的正确使用是很重要的一環，其重要性不亞於文章本身，故在引用資料的過程中，如未按照一定的格式規範撰寫，或者是沒有記載出處，日後讀者想要利用引文來延伸閱讀就會發生困難，甚至產生錯誤引用的狀況，對作者而言甚至會被冠上抄襲的惡名。

而在每一學門領域中，對於論文的撰寫引用資料或者是編製參考文獻時，都有其學術規範存在，研究人員在從事學術論文寫人作或投稿至學術期刊時也希望從所屬學門的規範。從各學門的期刊投稿格式規定中，雖然皆有其專屬文獻引用格式，但其實絕大部份也都是以國際知名的寫作規範所變化而來。在 1991 年被著名的 booklist 雜誌選為三大格式手冊分別為，Chicago Style、American Psychological Association Style（APA）以及 Modern Language Association Style（MLA），國內外對於引文規範的討論多以此三種格式為對象，以下就這三種格式進行說明：

### 一、Chicago/Turabian Style

Chicago Style 包含 Kate L. Turabian 主編的 A Manual for Writers of Term Papers, These, and Dissertations，而不是單指 The Chicago Manual of Style，二者都是由芝加哥大學出版，Kate L. Turabian 以 The Chicago Manual of Style 為架構編制，提供大學生及研究生的正式論文、報告的寫作格式手冊 A Manual for Writers of Term Papers, These, and Dissertations，至今已經過六十多年，該手冊已成為知名的論文寫作手冊之一。但雖然 A Manual for Writers

of Term Papers, These, and Dissertations 在學術界的地位頗為重要，但在探討學術格式規範時，其重要性還是不及 The Chicago Manual of Style。The Chicago Manual of Style 最新版為 2003 年出版的第 15 版，而目前 A Manual for Writers of Term Papers, These, and Dissertations 最新版為 2007 年以 The Chicago Manual of Style 第 15 版為藍本的第 7 版，但此版本與上一版相隔 11 年之久。其使用範圍極廣，包含了文史、哲學及科學等。

## 二、American Psychological Association Style

APA 格式是指美國心理協會 (American Psychological Association) 所出版的出版手冊中，有關投稿該協會旗下所屬二十九種期刊時必須遵守的規定。該手冊詳細規定文稿的架構、文字、圖表、數字、符號等的格式通稱為 APA 格式 (APA style)，相關領域的期刊、大學報告、學位論文也常參考其格式，做為要求投稿者及研究生之依據。APA 出版手冊自 1944 年出版以來，經過 1974 年的二版，1983 年的三版，1994 年的第四版，於 2001 年 7 月發行第五版，在 2007 年六月為了符合網路新科技以電子參考資源出版 APA Style Guide to Electronic References，以補足現今的需求。APA 格式為心理、教育及社會科學領域學術期刊經常採用的書目格式。

## 三、Modern Language Association Style

MLA 格式為由美國現代語言學會所制定，最初出現在 1951 年語言學會期刊 PMLA 上的 MLA Style Sheet，此規範立即受到學術出版界及學會會員的歡迎，接著經過多年的增補修定，1977 年正式發行成為 MLA Handbook for Writers of Research Papers。接著在 1984 年發行第二版，在 1988、1995 及 1999 發行三、四、五版，目前為 2003 年出版的第六版。MAL 格式根據使用者的不同，分為給高中以及大學生使用的 MLA Handbook for Writers of Research Papers；給研究生與專家學者使用的則是 MAL Style Manual and Guide to Scholarly Publishing 二種，MAL Style Manual and Guide to Scholarly

Publishing 在 1998 年出版第二版來因應電子文獻的使用需求。MLA 格式被人文學領域廣泛採用，特別是文學及語言等。

## 2.3 相關資訊技術

該節將針對系統所使用到的資訊技術，包含用於系統在期刊刊名比對用到的近似字串比對、跟 WOS 溝通的 Web Service、以及系統的開發平台 ASP.Net 進行介紹。

### 2.3.1 近似字串比對

近似字串比對 (Approximate String Matching) 為字串比對 (String Matching) 所衍生出來的問題。近似字串比對主要是研究字串 P (Pattern) 的近似子字串 (substring)，是否在較長的字串 T (Text) 出現的方法。近似字串比對的問題一直以來就不斷的被研究，用以改進其準確率以及效率，也不斷的有新的演算法被提出來，在目前常見的近似字串比對演算法可分為單一樣式演算法 (Single Pattern Algorithm)，以及多樣式演算法 (Multi Pattern Algorithm) 也就是一次比對多個 P 的演算法。目前常見的 Boyer-Moore Algorithm 以及 Knuth-Morris-Pratt Algorithm 就是單一樣式演算法，Boyer-Moore Algorithm 被認為在實際應用中，最有效的單一樣式演算法。而 Wu-Manber Algorithm 以及 Aho-Corasick Algorithm 則是常見的多樣式字串比對演算法。

### 2.3.2 Web 服務

隨著資訊技術的發展，依照需求而發展出不同的資訊系統，系統之間的差異性也越來越大，為了讓系統之間能互相交流，必須要有共同的協定，Microsoft、IBM、以及 Sun 等資訊大廠為了解決這個問題，陸續向全球資訊網協會 (World

Wide Web Consortium, W3C) 提出以 XML 為基礎的標準協定，做為網際服務的溝通語言及工具，Web 服務 (Web Services) 就是因應這樣的需求而產生。Web 服務是一種服務導向架構 (Service-oriented architecture) 技術，主要是透過 Web 協議給不同平台的應用程式提供服務，讓平台之間能動過服務相互溝通。Web 服務是以 Web 的開放標準為基礎它包含：

#### 一、eXtensible Markup Language (XML)

用來描述資料的一種標記語言，主要是說明資料是什麼，以及攜帶資料資訊。XML 被廣泛用來作為跨平台之間互動資料的形式，主要針對資料的內容，通過不同的格式化描述方式 XSLT, CSS 等，可以完成最終的展現格式，例如產生對應的 HTML, PDF 或者其他的檔案格式。

#### 二、Simple Object Access Protocol (SOAP)

是一個以 XML 為基礎的編碼技術製作出來的通訊協定，SOAP 的出現是為了簡化網頁伺服器在從 XML 資料庫中提取資料時，無需花時間去格式化頁面，並能夠讓不同應用程式之間透過 HTTP 通訊協定，以 XML 格式互相交換彼此的資料，所以應用程式之間不需要知道彼此的作業平台是那一種，或是各自如何實作等細節資訊，就可以透過 SOAP 來相互溝通。其作業方式為將溝通用的內容封裝成 SOAP 訊息，當 Client 端發出 SOPA 請求給 Server 端的 Web 服務時，Server 在收到請求後會依據請求將處理完的資料回應一份 SOAP 給 Client 端。

#### 三、Web Services Description Language (WSDL)

WSDL 是由 Microsoft 及 IBM 共同訂定的標準。主要是描述 Web 服務的細節，也是使用 XML 格式之語言，讓 Web 服務應用程式能以一種標準方法來描述自己擁有的介面及能力，以便讓互動更容易進行。

#### 四、Universal Description, Discovery and Integration (UDDI)

UDDI 是由 Microsoft、IBM 以及 Ariba 所共同制定的註冊及儲存標準。主要

是 Web 服務的目錄註冊服務，其架構也是以 XML 為基礎，其目的為提供 Web 服務的提供者透過 UDDI 來公布所提供的 Web 服務，讓用戶端可以快速的悉知哪些可利用的 Web 服務。

目前 W3C 針對 Web 服務的定義涵蓋了許多相異且無法介分的系統，但通常在實作還是會根據 SOAP 協定進行傳輸 XML 格式的資料。不管是定義還是實作，在 Web 服務的伺服器端還是會提供一個機器可選的描述 WSDL 來讓用戶端來瞭解服務的細節，雖然 WSDL 並非是 SOAP 伺服器端的必要條件，但目前 JAVA 的 Web 服務開發框架就把 WSDL 設定為必要的條件。甚至有一些工業標準化的組織，將 Web 服務的定義，強制的把 SOAP 及 WSDL 包含進去。

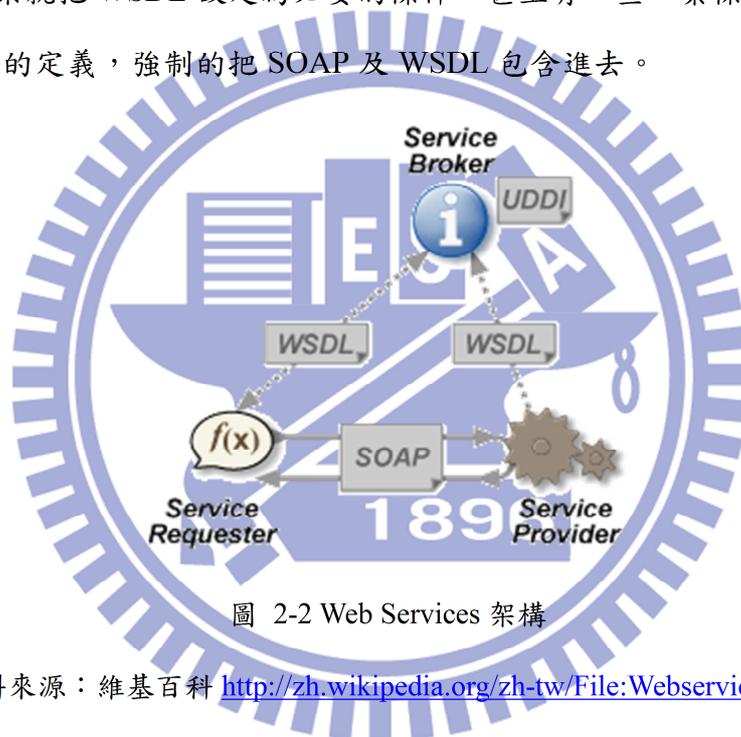


圖 2-2 Web Services 架構

資料來源：維基百科 <http://zh.wikipedia.org/zh-tw/File:Webservices.png>

### 2.3.3 ASP.NET

ASP.NET 是由微軟 (Microsoft) 在 .NET Framework 提供的一種開發平台 (Development Platform)，主要是用來做為 Web Application 的開發，早期的 Web Application 都是由 Common Gateway Interface (CGI) 所開發，常見的 CGI 程式例如 Perl 或 C，大都是處理表單 (Form) 資料為主。但是 Web 平台的迅速發展，CGI 的開發已經逐漸不符需求，開始出現例如 Active Server Pages (ASP)、

Hypertext Preprocessor (PHP) 或 JavaServer Pages (JSP) 這樣的簡化程式設計流程與提供強大資料處理能力的模組，應用在 Web 伺服器的外掛模組 (plug-in module)。在 ASP 平台紅極一時的時候，微軟也因為 ASP 的諸多限制，例如程式語言只支援 VBScript、Jscript，只能存取支援 COM 的元件，直譯式語言的效能問題等，因而著手重新設計 ASP 的架構，導入基礎類別庫 (Base Class Library) 的支援，強化了程式模型以及平台的支援，在 2002 年推出包含於 Microsoft .NET Framework 1.0 的 ASP.NET 1.0 版，宣告新一代的 ASP.NET 時代的開端，目前最新的版本為 ASP.NET 4.0，內含於 Microsoft .NET Framework 4。

ASP.NET 的許多控制來簡化繁瑣的開發工作，讓程式設計人員可以有效的提昇開發效率，大幅的增加生產力，另外也內建 ASP.NET Asynchronous JavaScript and XML (AJAX) 的技術、Master Page、Profile 使用者設定檔、Web Part Framework、網站管理工具、以及帳號權限管理系統等等的解決方案或技術，讓程式設計人員在程式設計上直接應用，減少重複造輪子的工作，這也是本系統選擇以 ASP.NET 平台來開發的主要原因。

# 第三章 系統分析與設計

本章將以參考文獻計量為基礎，朝向設計出期刊使用統計系統（SCAJ）目標進行，3.1 依據對圖書館員進行的需求訪談，將所提出的需求進行分析說明，3.2 說明系統的是以何種概念做為設計的方向，3.3 則進入系統的設計部份，將規劃出系統的架構以及模組功能的設計。

## 3.1 系統需求分析

系統需求是依據圖書館所面臨的期刊使用統計問題，以設計期刊使用統計系統（SCAJ）目標進行，針對大學圖書館館員進行需求訪談。依據訪談結果整理及分析出系統需求，將需求分為資料需求、功能需求、介面需求、以及報表需求四個面向，如圖 3-1 系統需求面向所示。系統主要是以資料需求的輸入（Input）以及報表需求輸出（Output）為主，以下針對這四個面向的需求進行詳細說明。

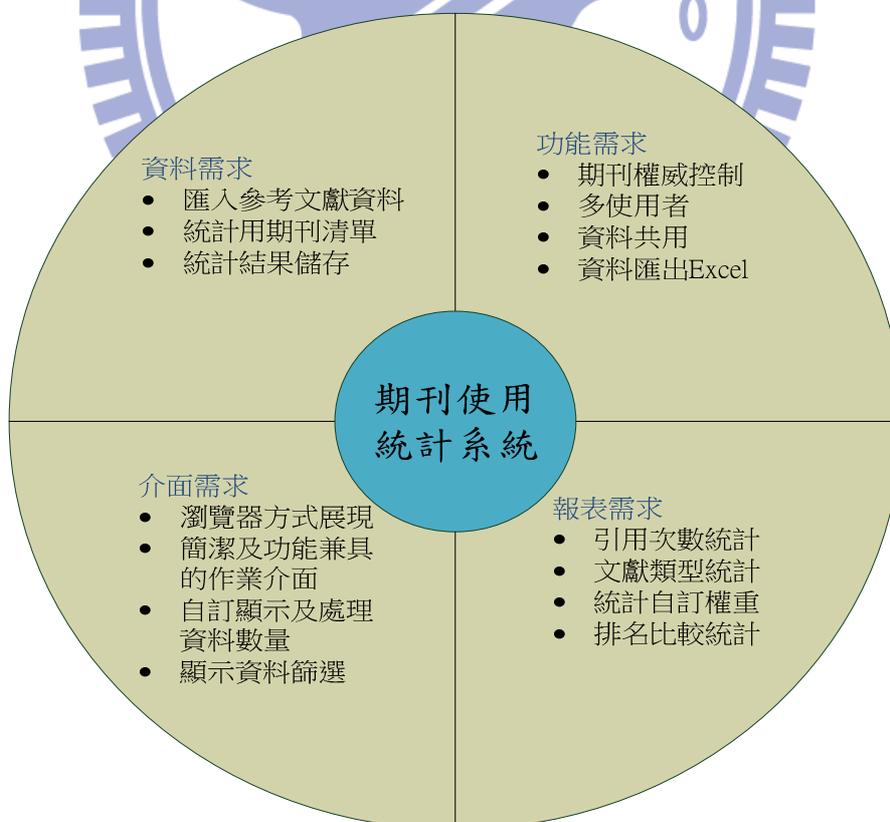


圖 3-1 系統需求面向

## 一、 資料需求：

### (一)匯入參考文獻資料

使用者上傳所要分析的引文資料文字檔，系統將上傳文字檔分析，分析出引文類型及期刊名稱，並依分析的結果儲存於資料庫內。引文類型分類如下：

1. 西文期刊：以英文為主的期刊論文。
2. 會議論文：不分語言的會議論文。
3. 網路資源：來自於網路上的文獻。
4. 圖書：不分語言的圖書文獻。
5. 中文期刊：以中文為主的期刊論文。
6. 博碩士論文：不分語言的博碩士論文。
7. 研究報告：不分語言研究報告文獻。
8. 其它：無法分類於上列分類範圍的文獻。

### (二)統計用期刊清單

用於統計作業，除了引文資料外，額外的參考指標，讓使用者自行上傳統計用期刊清單包含所要分配的分數，經由系統分析後儲存於資料庫，於統計作業時給予自訂權重，產出計算後的分數，可配合引文資料一起計算或單獨計算。

### (三)統計結果儲存

統計作業時，使用者自訂給予不同的引文資料、統計用期刊清單，以及不同的權重，進行比較每次的結果變化，必需先做統計結果的儲存再取出比較。

## 二、 功能需求：

### (一)期刊權威控制

需設置期刊刊名資料庫做為系統的核心資料，參考文獻資料的分析及統計用期刊清單，皆需跟期刊刊名資料庫比對，來產生比對後結果，但可能產生同一期刊但期刊名稱不一樣的狀況，故需要設置刊名的權威控制來解決同一期刊但期刊名稱不一的問題。

#### (二)多使用者

需要讓多個使用者同時使用系統而互相不衝突，每個使用者有獨立的作業空間，不因其他使用者的變動而造成資料錯誤。

#### (三)資料共用

在資料共用主要希望讓使用者之間，相互參考文獻的資料，以及統計期刊清單的共用，讓使用者間也可以將對方的所建立的資料加入，當成一個參考指標來產出統計的結果。

#### (四)資料匯出 Excel

系統所產出的結果需建立匯出功能，提供匯出 Excel 格式讓使用者可以利用既有的工具程式，做後續的資料處理或建立圖表。

### 三、 介面需求：

#### (一)瀏覽器方式展現

不需安裝用戶端軟體就可以使用系統的所有功能。

#### (二)簡潔及功能兼具的作業介面

在作業畫面中，盡量將所需的功能加入，需要的訊息顯示，不需要的訊息隱藏，並可以依需求顯示隱藏訊息及功能，不要讓使用者在引文作業或其它的作業時，因為訊息太多而眼花瞭亂或資料量太大而傳輸太久造成作業緩慢。

#### (三)自訂顯示及處理資料數量

可由使用者自訂顯示的資料數量，可以一次處理更多的引文資料，來加快引文資料的處理速度。

#### (四)顯示資料篩選

對顯示的資料做篩選，來更快找到所需的資料，並進行相關作業。

#### 四、報表需求：

##### (一)引用次數統計

統計所分析資料的各期刊引用次數，包含權威控制及排除權威控制，以及同時顯示每筆資料的單獨引用次數跟合計次數。

##### (二)文獻類型統計

所分析的參考文獻資料類型統計。

##### (三)統計自訂權重

引文資料及期刊清單資料，可透過自訂權重的方式，產出統計報表，並讓使用者單獨或合併統計。

##### (四)排名比較統計

將自訂不同權重的統計報表儲存，再利用排名比較統計來產生報表，以便使用者觀查各儲存記錄的期刊排名變化。

## 3.2 設計概念

WEB2.0 網路運用新時代的當下，共建共享是其主要精神，系統將依據這個的精神去設計，並建置一個以使用者為導向的期刊使用統計系統 (SCAJ)，其設計概念如下：

### 一、簡易使用

建構在一個透過網際網路就能使用，來降低使用門檻，讓有需求的研究人員或圖書館館員都能使用的開放式系統，不必自行建立及開發，不必安裝用戶端軟體就能使用系統。

### 二、個人化作業環境

透過網路進入系統就能有獨立的作業環境，不會因操作而影響他人的作業或

他人的作業影響自己。

### 三、自訂權重報表

讓使用者自己上傳參考清單，來自由的訂定各項權重，做各項權重調整後的比較，可透過系統產出的數據，做為研究及評估的參考。

### 四、使用者回饋

透過更多使用者的使用，來回饋系統豐富資料庫內的資料，進而提高系統的準確度及其效率，並且能互相參考其它使用者所建立的資料或當成評估的條件。

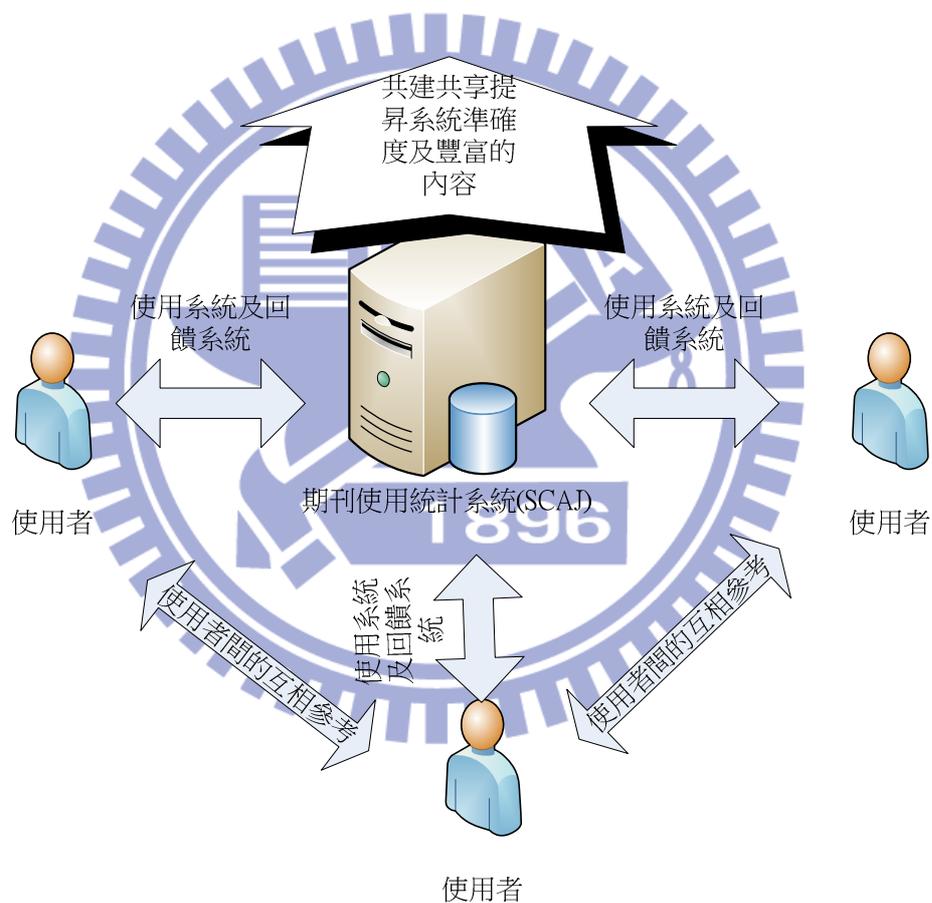


圖 3-2 系統設計概念

## 3.3 系統設計

本節主要是依據 3.1 節的系統需求分析結果來設計系統，完成如圖 3-3 系統架構所示的完整系統架構，並規劃設計系統的各模組及功能。3.3.1 為系統實體

架構的說明；3.3.2 針對主要的系統作業流程的設計進行說明；3.3.3 系統模組及功能是以圖 3-3 系統架構

圖 3-3 系統架構為參考，將系統以功能區分設計成六個模組，最後在系統中較為重要的參考文獻分析設計，於 3.3.4 參考文獻分析功能及 3.3.5 WOS Web 服務 (WOS Web Services) 二節中詳述其設計方法。

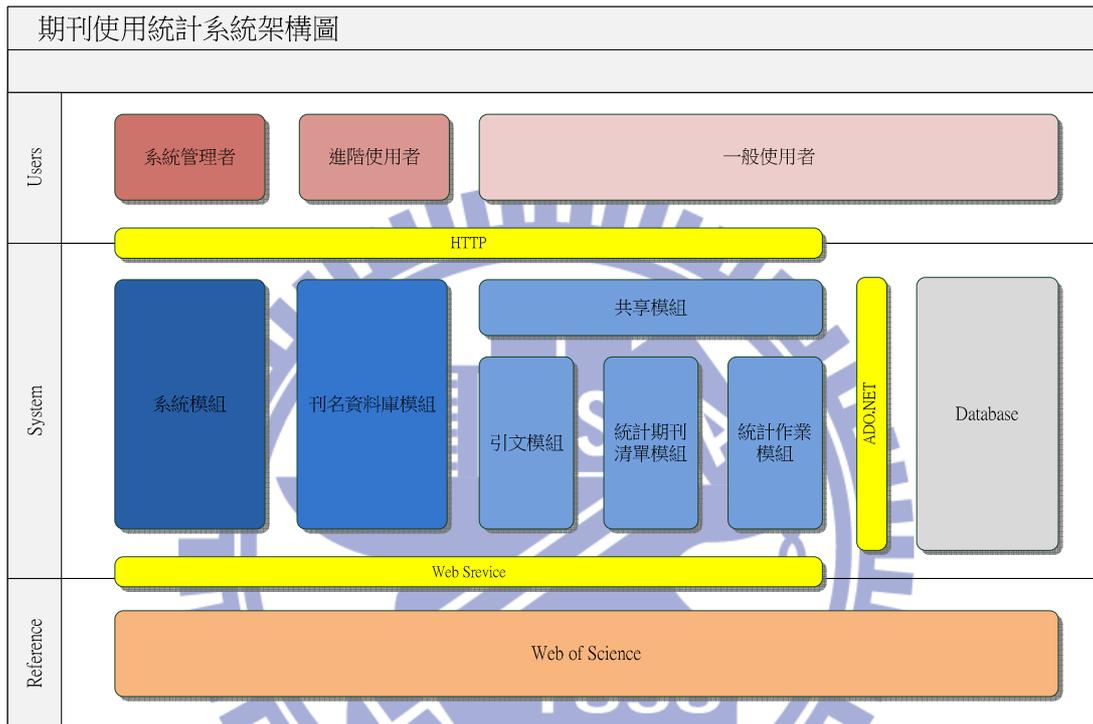


圖 3-3 系統架構

### 3.3.1 系統實體架構

系統採用三層式架構 (3-tier architecture) 來建置本系統，其架構如圖 3-4 系統實體架構所示，在實體的架構中，基於資安考量會將商業邏輯層建置於防火牆的非軍事區 (Demilitarized Zone, DMZ) 內，資料存取層 (Data Access Layer, DAL) 層置於防火牆內的 Intranet Zone。

#### 一、展示層 (Presentation Layer)

使用者端電腦連上網際網路 (Internet) 後，透過瀏覽 (Web browser) 以 HTTP

跟系統（Web Server）做為二者間的溝通協定。

## 二、商業邏輯層（Business Logic Layer, BLL）

負責在 Web Server 上執行處理邏輯運算，此層接受來自使用者端的請求（Request）服務，並且將程式處理結果傳回到使用者端，它的目的是用來作為使用者與資料庫之間的溝通橋樑，系統在該層是以 Internet Information Services（IIS）6.0 建置。

## 三、資料存取層（Data Access Layer, DAL）

DAL 主要提供資料給 BLL 來處理，系統建置 Microsoft SQL 2005 做為資料庫，Client 端無法直接對資料庫進行存取動作，必須經由 BLL 透過 ADO.NET 與 DAL 的連結對資料進行存取。

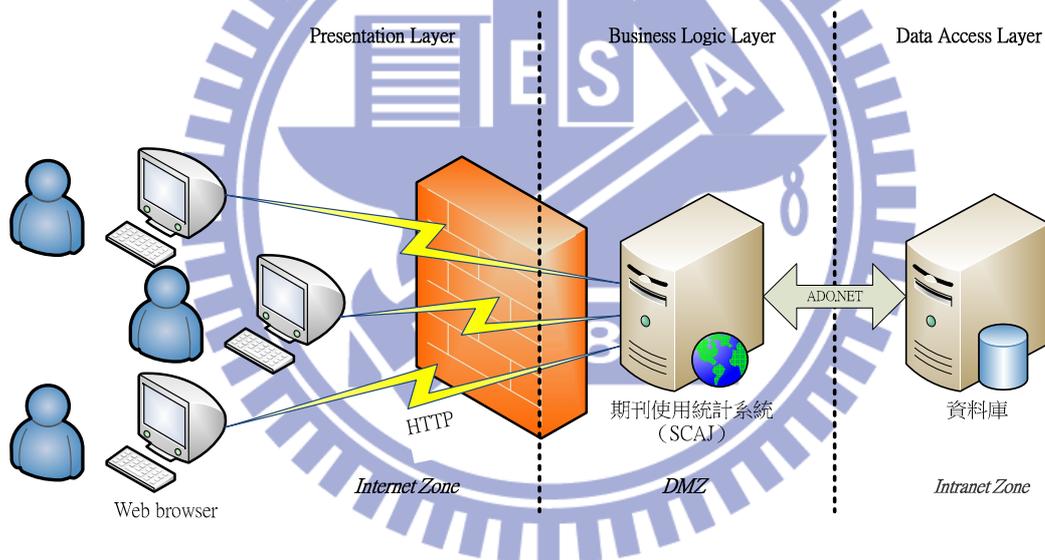
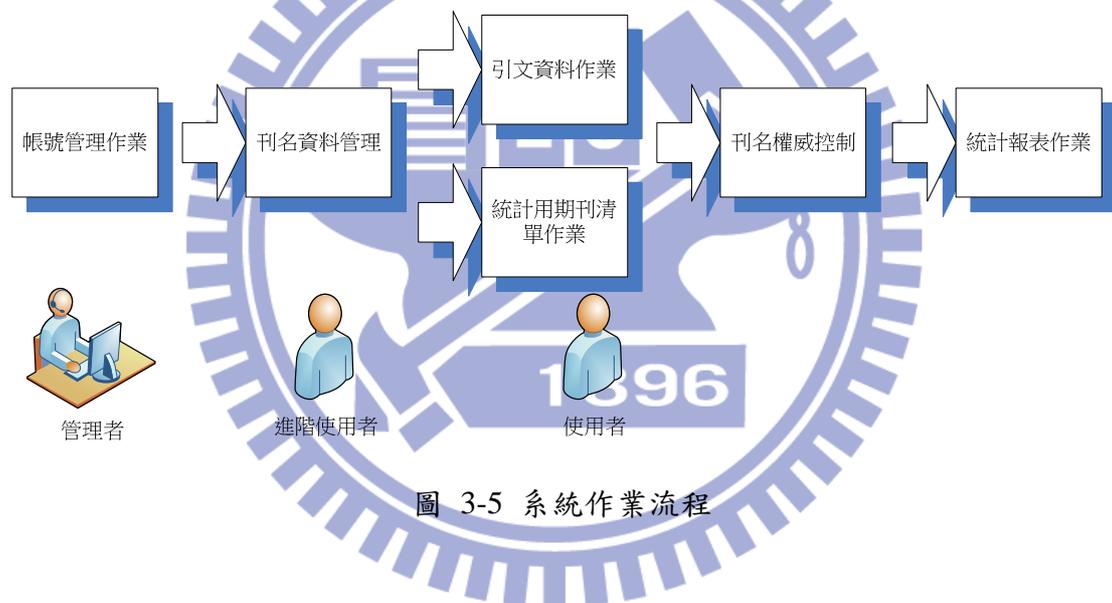


圖 3-4 系統實體架構

### 3.3.2 系統作業流程

依據系統需求分析，規劃設計主要的系統作業流程。在系統中，每項功能都有其相互關係，如圖 3-5 系統作業流程；先從管理者建立使用者帳號及權限開始，再由管理者及進階使用者匯入蒐集的期刊名，增加刊名資料庫內的資料量，如果刊名資料庫內的資料量不足時，會造成引文作業的系統分析結果準確率低，接著進行引文資料作業或統計用期刊清單作業。進行完前述二項作業後，可能產生相同期刊卻有不同刊名的狀況，在統計作業前必需完成刊名權威控制作業，將相同的期刊合併，以解決在統計時系統將其視為不同期刊的錯誤。最後再進行統計報表作業才能產生正確的結果，接下來依據作業流程部分做詳細的說明。



## 一、刊名資料庫管理

刊名資料庫為本系統之核心資料，存放所有的期刊名稱，目的為用來讓引文資料及統計用期刊清單做為期刊比對用，所以刊名資料庫內的資料愈多，相對的系統的準確程度也愈高。刊名資料庫內的資料蒐集如圖 3-6 刊名資料庫作業示意，可由管理者或進階使用者匯入蒐集的期刊清單，也可由使用者在引文資料作業及統計用期刊清單作業時增加未收錄的期刊名資料，其作業流程如圖 3-7 刊名資料庫作業流程；期刊名資料上傳作業及統計用期刊清單上傳作業，都是將上傳的資料一筆筆的分析，如果不存在內就新增至資料庫內，而引文作業則是由使用者在作業時所設定的期刊名來判斷是否新增。

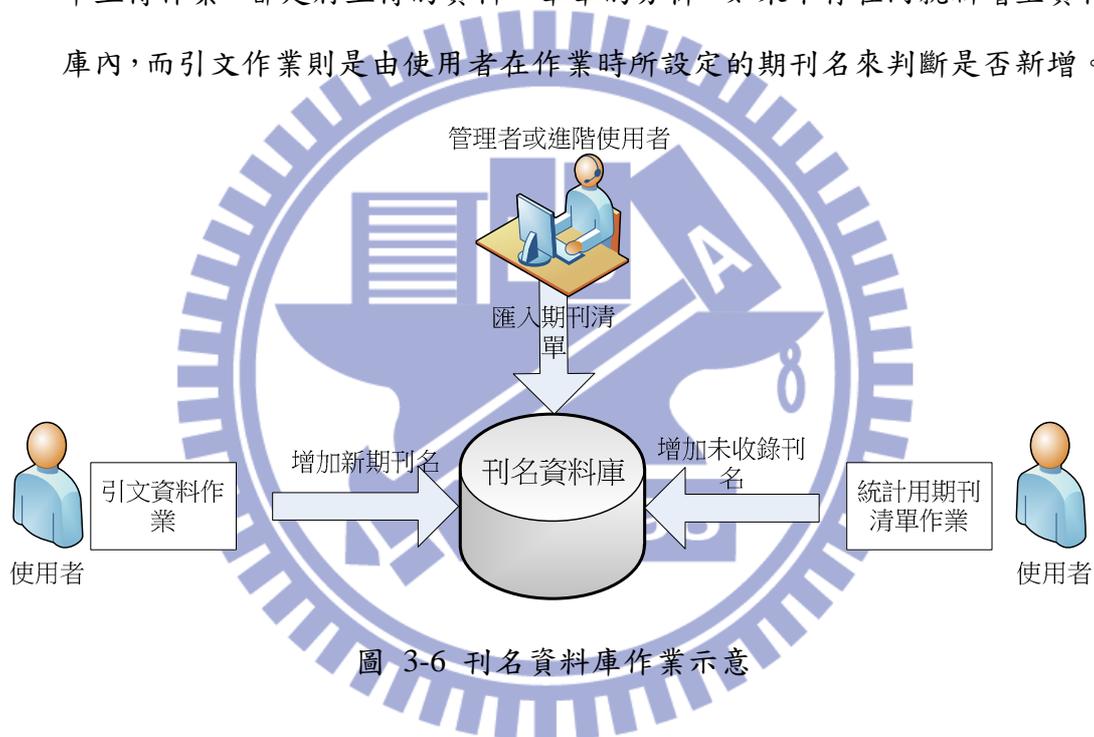


圖 3-6 刊名資料庫作業示意

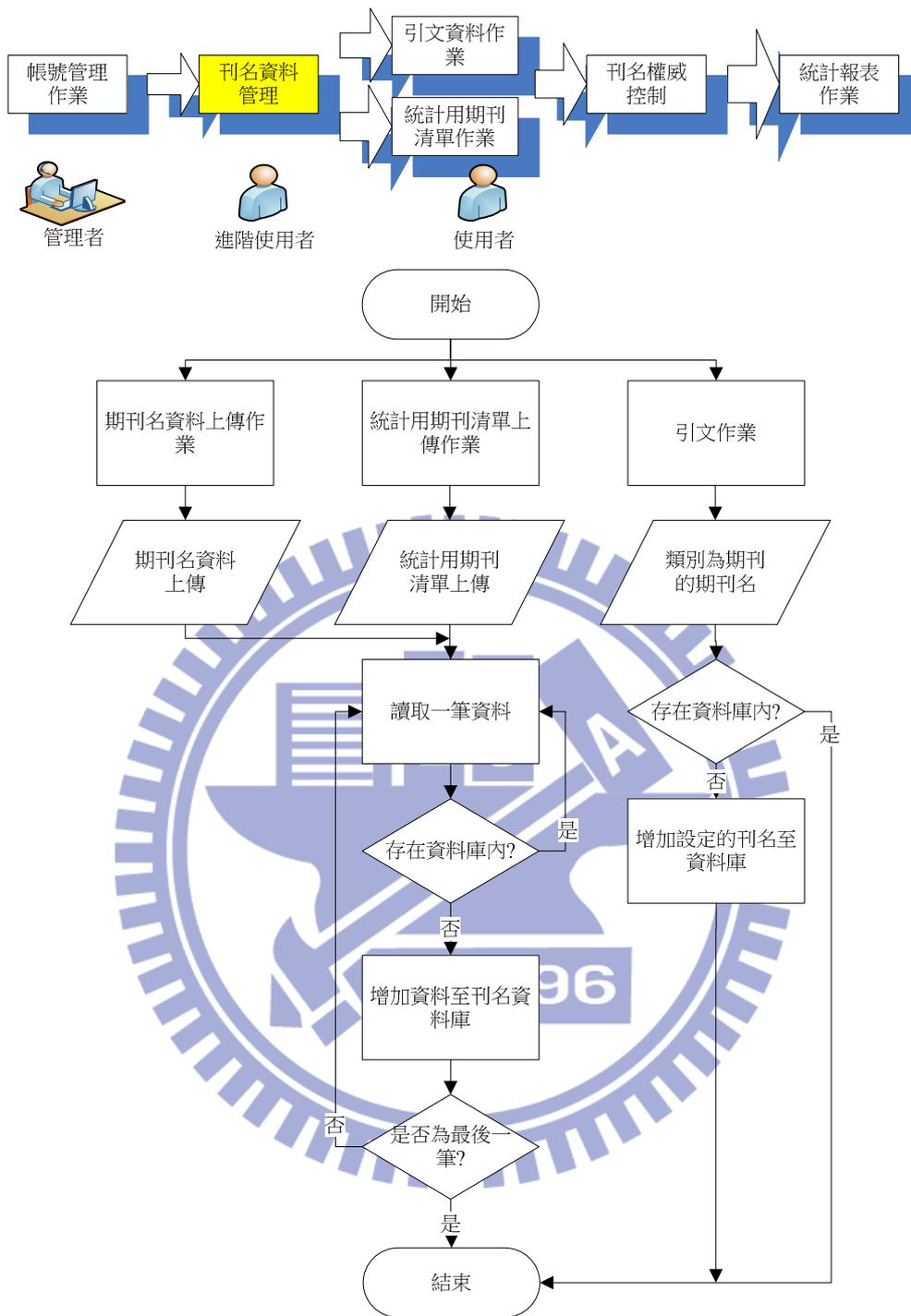


圖 3-7 刊名資料庫作業流程

## 二、引文資料作業

該作業為系統最主要的作業，其統程如圖 3-8 引文資料作業流程，讓使用者將要分析的引文資料檔案上傳，再經由系統做分析，將每一筆資料加以分類，並判別出西文期刊的刊名，完成分析作業之後，可由使用者修改成方便

識別的檔名，再選擇要分析的檔案，由使用者針對每一筆資料加以確認，確認類別的正確性以及類別為西文期刊的刊名是否正確。

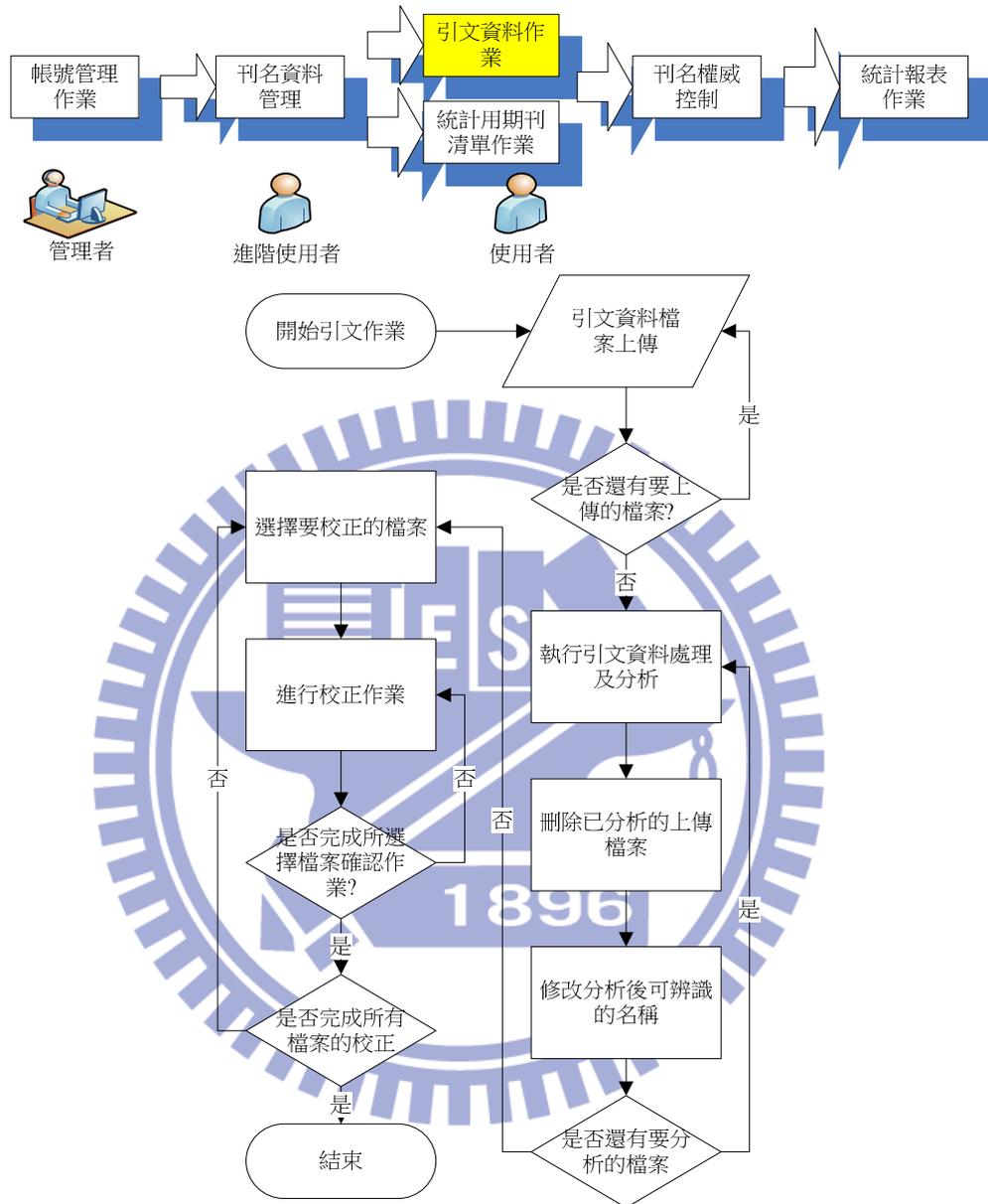


圖 3-8 引文資料作業流程

資料校正的詳細流程，如圖 3-9 引文資料校正作業流程，校正作業是以一次多筆的方式進行校正以加快作業速度；首先檢查資料是否完整，不完整的資料可以做刪除動作；再檢查分類是否正確，不正確可修改為正確的分類，再來檢查類別是否為西文期刊，如是是西文期刊就檢查是系統分析出的刊名

是否正確，不正確的話需自行修改成正確的期刊名稱，重覆比對完整個頁面的資料後再進行做儲存。

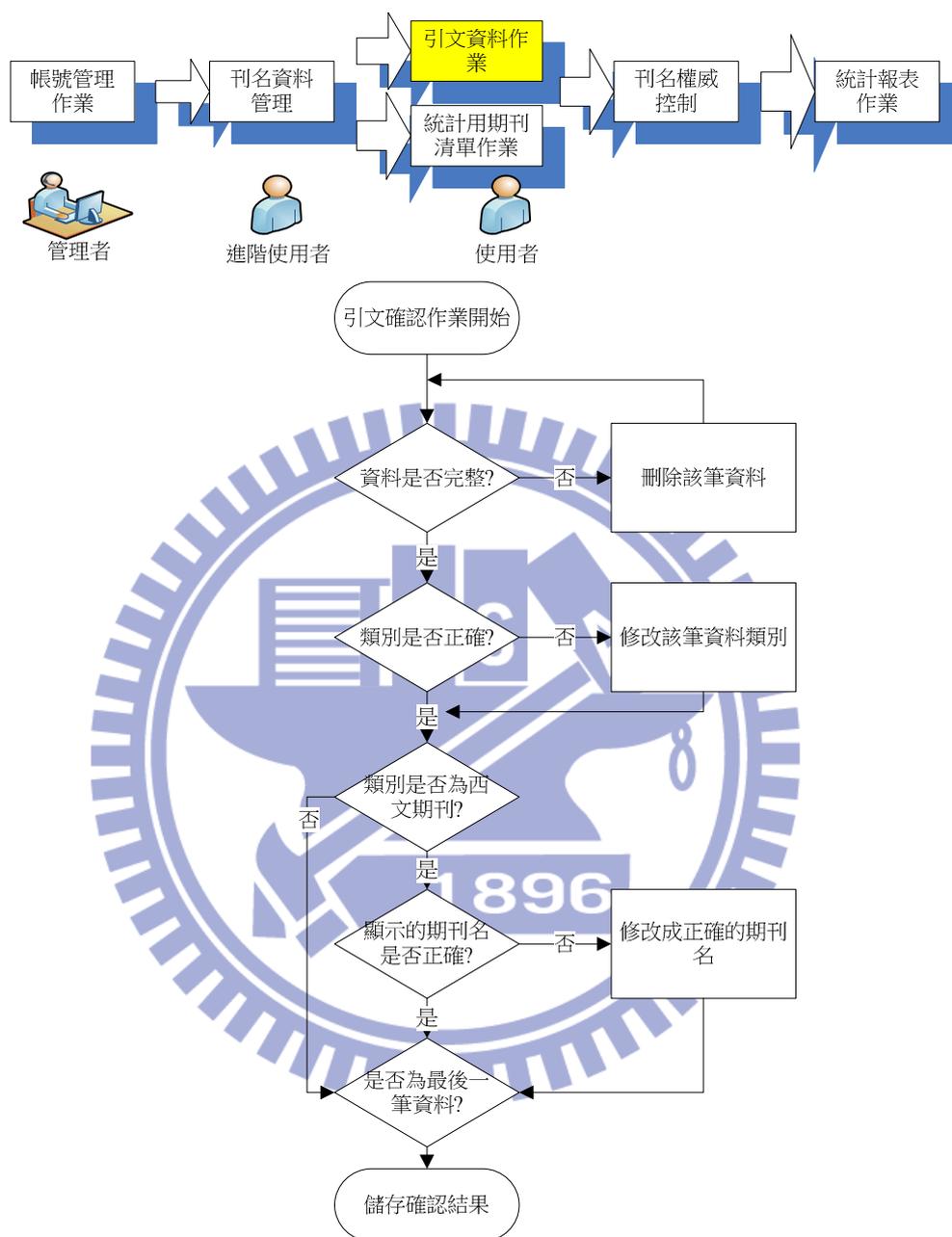


圖 3-9 引文資料校正作業流程

### 三、統計用期刊清單作業

讓使用者自行上傳期刊清單，目的是在統計作業時可以加入上傳的期刊清單並給予不同的權重，提供使用者引文外的另一種評估指標，例如可上傳 JCR 的清單或國科會核心期刊清單等。如圖 3-10 期刊清單作業流程說明，使用

者先將要上傳的期刊清單，自訂加入 1 至 100 的分數，統計後期刊所得到的分數為自訂分數\*權重%，整理後上傳儲存在資料庫內，最後系統會自動分析出是否有重覆的期刊名，由篩選出來的重覆期刊中刪除重覆的資料就完成該項作業，重覆的期刊名會造成統計時重覆計算導致結果錯誤。

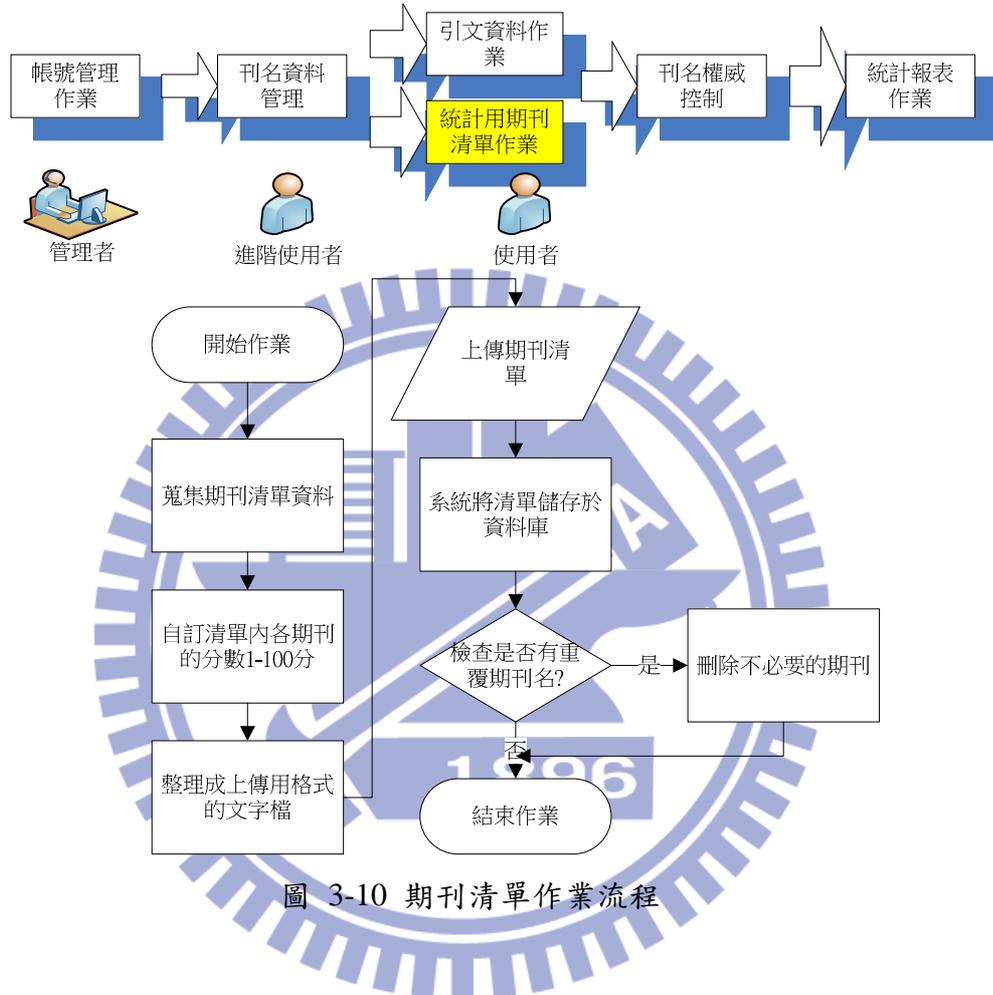


圖 3-10 期刊清單作業流程

#### 四、刊名權威控制

目的為控制刊名資料庫內的期刊名的正確性，如有同一期刊但期刊名稱不一樣的狀況，則可設置刊名的權威控制來連結這些刊名，以解決此問題。如果未設置好權威控制會造成引用次數的分散，統計出來的結果就會有錯誤。系統會自動合併相同的國際標準期刊代碼 (International Standard Serial Number, ISSN) 期刊刊名，也可利用系統的篩選功能找出一樣的期刊，再選權威控制功能來合併期刊，在引文資料的引用次數統計，以及統計自訂權重統計都必需考量刊名權威控制的正確性。

## 五、統計報表作業

- (一) 引文資料的引用次數部份，選擇一個或數個先前作業所完成的引文資料檔，送出統計後統計出各檔及合計的引用次數，並可另存為 Excel 格式檔，供使用者後續利用。
- (二) 引文資料分類統計，選擇一個或數個先前作業所完成的引文資料檔，送出統計後系統會合併所選的檔案，並且統計出各分類的數量，並可另存為 Excel 格式檔。
- (三) 合併統計報表，可讓使用者自訂統計權重，選擇一個或數個先前作業所完成的引文資料檔，給予權重百分比，或者是選擇一個或數個先前上傳的統計用期刊清單，再給予權重後送出執行統計，系統會依照所選擇的檔案及權重計算出各期刊所得到的分數。結果可儲存在系統內，供排名比較統計做排名比較，也可另存為 Excel 格式檔。系統所使用之正規化評分方式舉例如：若被引用次數最高之期刊為 IEEE Transactions on Pattern Analysis and Machine Intelligence，其被引用數為 200 次，而被引用數第二高期刊為 IEEE Transactions on Computers，其被引用數為 198 次，則將給予 IEEE Transactions on Pattern Analysis and Machine Intelligence 之評分為 100 分，而 IEEE Transactions on Computers 之評分則為  $(f - \min / \max - \min) \times 100$ ，即  $(198 - 0 / 200 - 0) \times 100 = 99$  分，以此類推。
- (四) 排名比較統計，選擇合併統計報表所儲存的二筆以上資料，送出執行統計後，系統將依據合併儲存記錄的排名，產生排名比較報表，可方便使用者觀查各儲存記錄的期刊排名變化，其作業流程如圖 3-11 排名比較統計作業流程，結果可另存為 Excel 格式檔，供使用者後續利用。

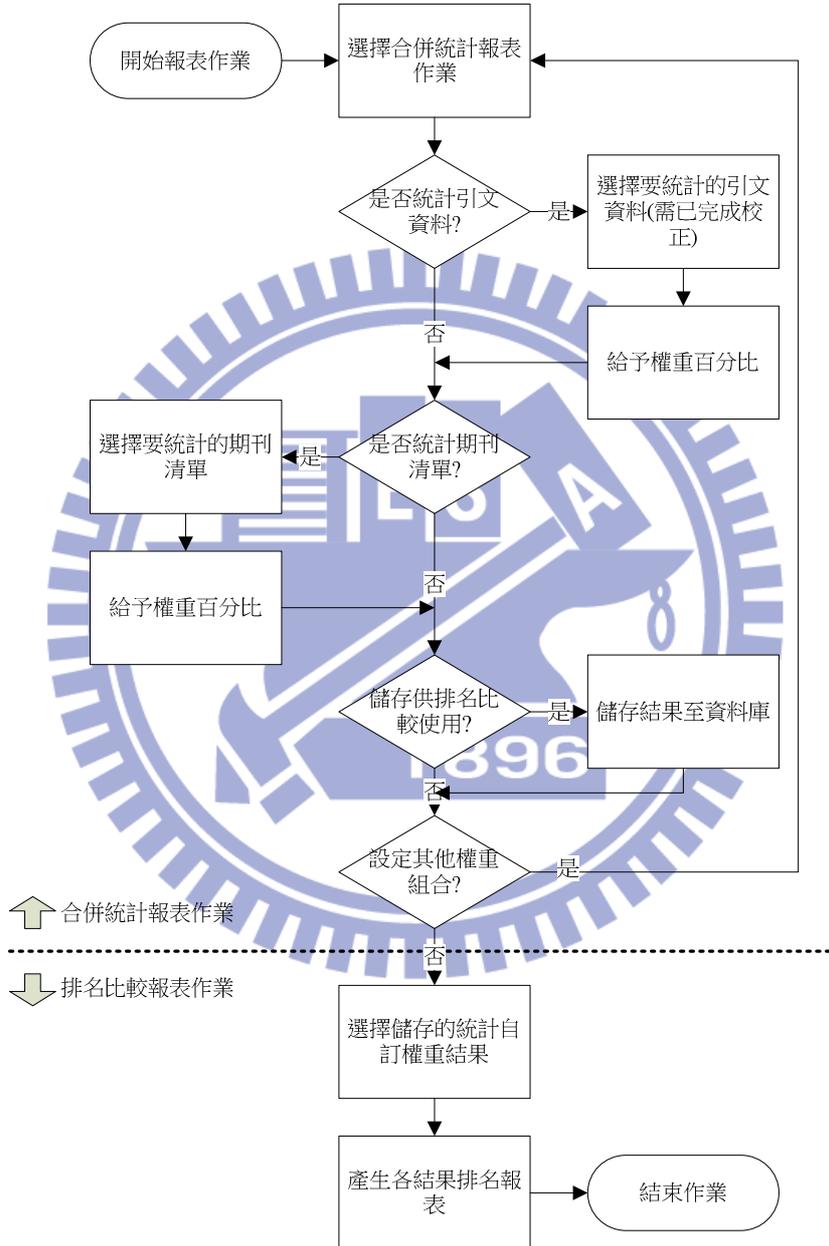


圖 3-11 排名比較統計作業流程

### 3.3.3 系統模組及功能

依據圖 3-3 系統架構設計，將系統以功能區分設計成六個模組，分別為系統模組、刊名資料庫模組、引文模組、統計期刊清單模組、共享模組、以及統計作業模組。每一個模組又分為幾個子功能，為做到簡潔的系統，盡可能將不同的子功能做在同一作業介面上；以資料庫為中心，各模組都跟資料庫做連結溝通，將所有的資料都存於資料庫內，以便系統運算分析利用或提供其它模組使用。如圖 3-12 系統模組所示，各模組功能簡介如下：

#### 一、系統模組

主要功能為管理系統帳號及使用權限，讓管理者透過管理模組進行帳號管理，以及設定適當的權限。

#### 二、刊名資料庫模組

主要功能為管理系統核心的期刊名資料，提供上傳整批期刊名資料，以及期刊名權威控制。

#### 三、引文模組

該模組為使用者在引文作業時最主要的工作環境，提供引文資料的整批上傳，上傳之後的分析、修改、校正及確認的作業。

#### 四、統計期刊清單模組

主要讓使用者上傳在統計時要用到的期刊清單，並包含所要分配的分數，以及重覆期刊的刪除作業，由系統自動分析是否有期刊重覆的狀況，重覆的期刊資料會於報表作業時產生不正確的輸出結果。

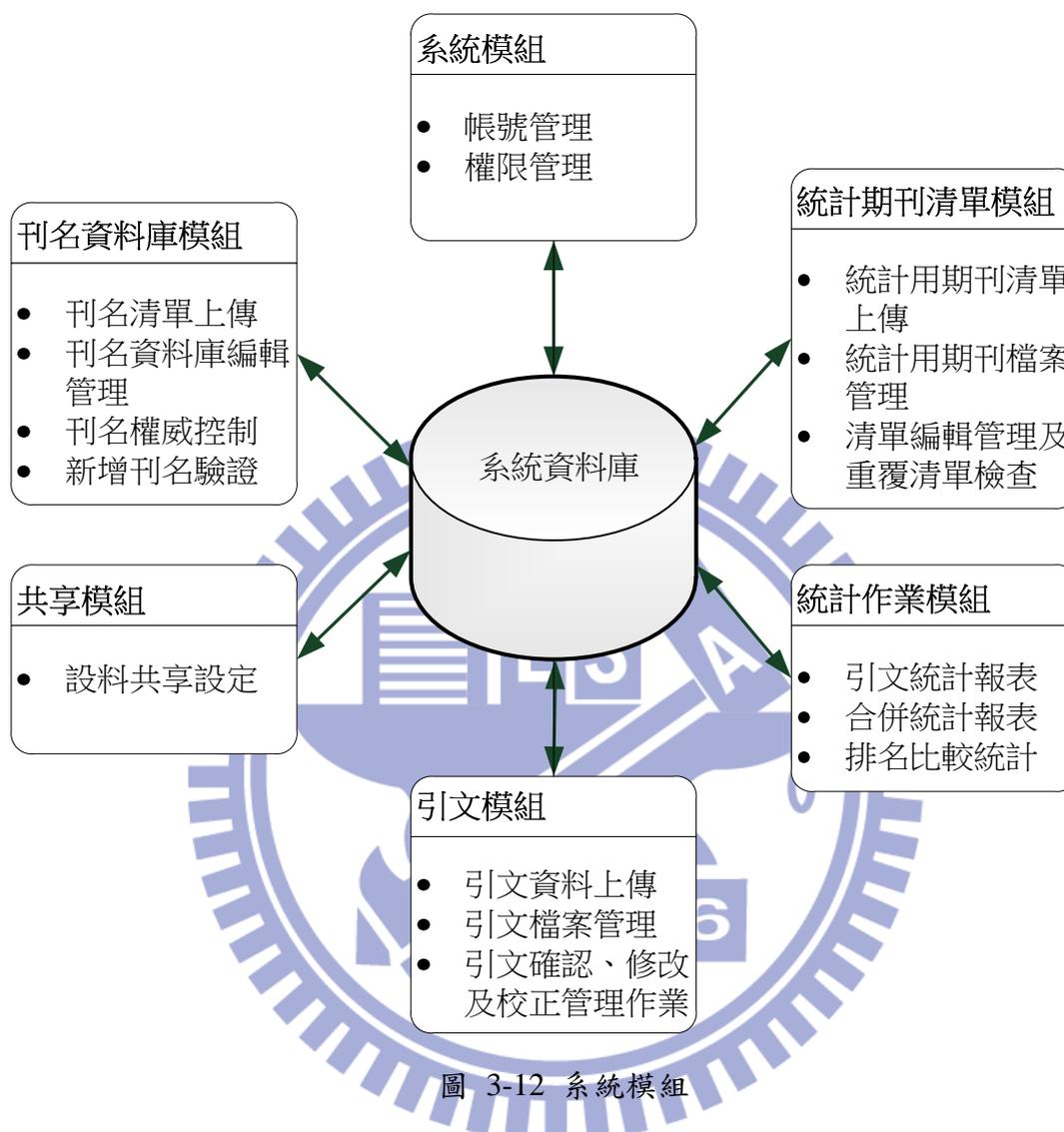
#### 五、共享模組

讓使用者分享所建立的資料，其他使用者可以利用別人分享出來的資料進行統計參考。

#### 六、統計作業模組

主要功能為各項統計報表的輸出，包含引文資料的引用次數，引文資料分類

統計，統計自訂權重，讓使用者單獨或合併統計，排名比較統計，讓使用者觀查各儲存記錄的期刊排名變化；並且各統計報表都提供 Excel 格式輸出。



依據上述六個模組規劃設計出如圖 3-13 系統功能。系統實做就以該架構為藍圖進行設計，首先完成刊名資料庫模組的設計，讓刊名資料庫可以進行上傳及整理功能，接著進行引文模組的設計，最後再完成其它的模組。

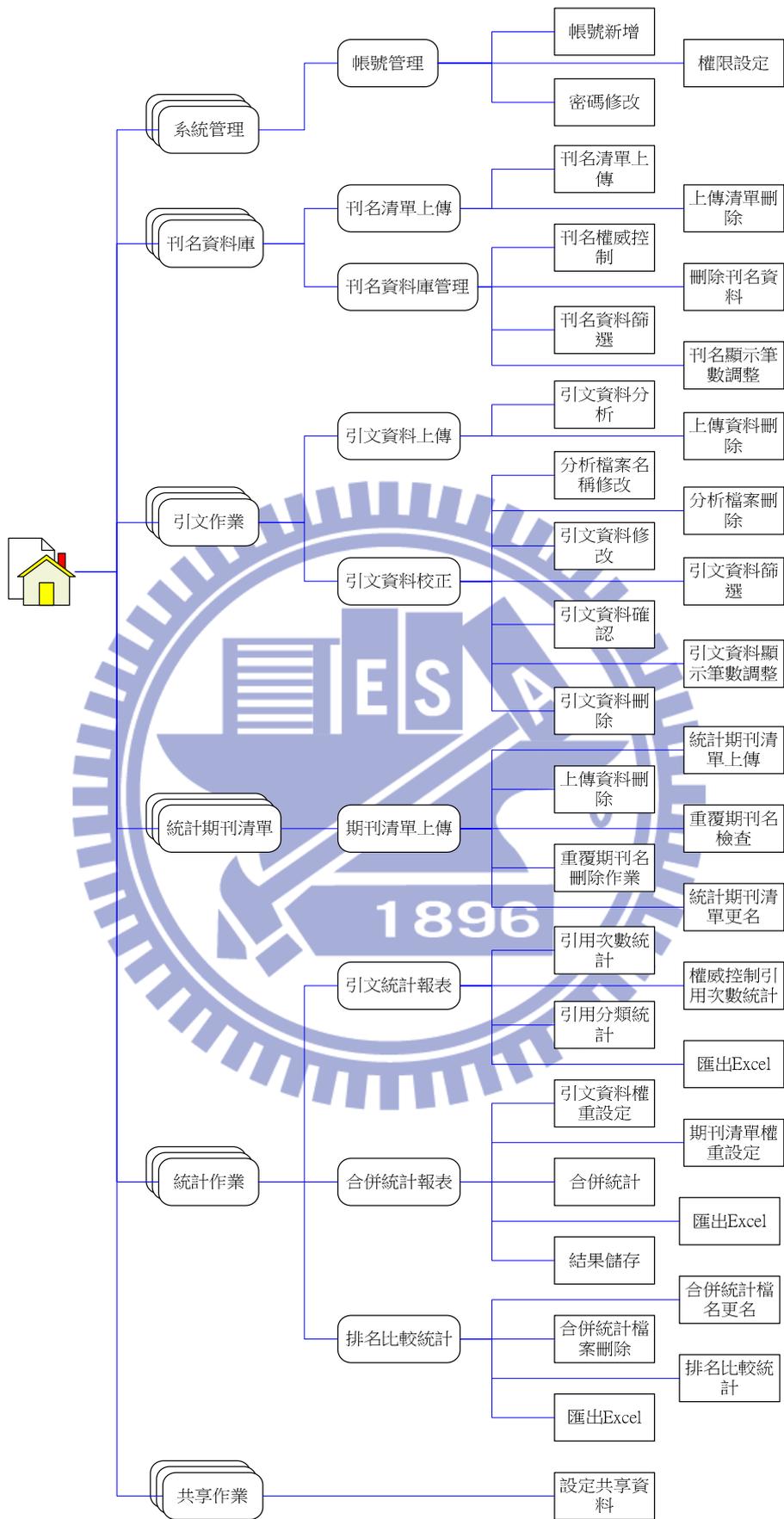


圖 3-13 系統功能

系統模組在使用者權限部份，因考慮到系統中有某些部份會影響到其他使用者所產出報表的結果，以及系統重要的刊名資料庫內容正確性，故將權限劃分為三種如下：

一、管理者

具有所有系統功能。

二、進階使用者

具有除系統模組外的所有功能，可管理刊名資料庫內的資料及設定刊名的權威控制，賦予系統核心刊名資料庫的管理任務。

三、一般使用者

具有引文模組、統計作業模組、以及統計期刊清單模組的系統一般性功能，停用刊名資料庫模組，是為了防止使用者更改刊名資料庫的內容，而造成其他使用者的統計結果產生誤差。

表 3-1 系統模組使用權限

模組名稱	權限劃分		
	管理者	進階使用者	一般使用者
系統模組	V		
刊名資料庫模組	V	V	
引文模組	V	V	V
統計作業模組	V	V	V
統計期刊清單模組	V	V	V

### 3.3.4 參考文獻分析功能

目前常見的參考文獻著錄方式有 American Psychological Association (APA) Style、Modern Language Association (MLA) Style、American Medical Association (AMA) Style、以及 Chicago Style。在這幾個常見的著錄方式中，如要利用程式透過各種書目格式的標準，以及近似字串比對演算法，來分析判斷出期刊名或是文獻類型是難以達到需求的，其原因說明如下：

- 一、無特定關鍵字、標示及特定的格式，足以讓程式能正確判斷文獻類型。
- 二、常用的書目格式在標示期刊名稱部份無明顯的區隔符號，程式難以分析出正確的期刊名稱。
- 三、文字檔無法展現書目格式的特定著錄方式，例如 APA、Chicago 及 AMA 會以斜體或 MAL 以底線的方式來著錄期刊名稱，轉換成文字檔再進入系統，這些特定的標示都會隨著消失。
- 四、著錄方式混亂，例如作者並未按照規定著錄。

系統在設計參考文獻分析處理時為因應上述狀況，故另外再設計參考文獻分析功能，其主要功能的為分析判定引文的類型，以及類型為西文期刊的文獻分析出其刊名。系統利用如圖 3-14 參考文獻分析處理流程所示，將參考文獻分析分為四個處理程序，來達到分析出文獻的類型及期刊名之目的，其程序之功能及處理方式依照順序如下，而關於 WOS Web 服務的設計將於下一節進行詳細說明：

#### 一、資料前置處理程序

將每筆引文資料做下列轉換，讓資料比對時的錯誤減少。

- (一) 轉換所有文字成為大寫。
- (二) 全形字元轉換成半形字元。
- (三) 去除無法轉換成半形的全形字元。
- (四) 去除前後及連續的空白字元。

例如：

轉換前 [66] R. Mohan, "Video Sequence Matching," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 6, pp. 3697–3700, May 1998.

轉換後 [66] R. MOHAN, "VIDEO SEQUENCE MATCHING," *IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING*, VOL. 6, PP. 3697-3700, MAY 1998.

## 二、西文期刊分析程序

該模組主要功能為分析判斷出是否為西文期刊及其刊名。首先分析取得可能的期刊名稱，再與刊名資料庫比對來確定是否為西文期刊及其刊名，如未比對到則進行可能期刊名比對 WOS 程序，如還是未比對到刊名，再進行篇名比對 WOS 程序。下列程序二、程序三及程序五只要有一程序產生比對結果，就不再往下進行其它程序，並將分析的引文資料判定類型為西文期刊，且判定分析出的期刊名為該引文資料的正確刊名，其詳細說明如下：

(一) 分析引文資料的期刊名子程序：利用下列三種方法來分析出可能的期刊名稱

1. 尋找引文資料內是否有 VOL.，如有則取 VOL. 前的雙引號跟逗號之間的字為可能期刊名。

例如：F. AURENHAMMER, "VORONOI DIAGRAMS - A SURVEY OF A FUNDAMENTAL GEOMETRIC DATA STRUCTURE," ACM COMPUTING SURVEYS, VOL. 23, NO. 3, PP. 345-405, 1991

2. 引文資料無雙引號，則找倒數第二個逗號到往前句號中間為可能期刊名。

例如：P. J. BURT AND E. H. ADELSON. A MULTIREOLUTION SPLINE WITH APPLICATION TO IMAGE MOSAICS. ACM TRANSACTIONS ON GRAPHICS, 2 (4) :217-236, 1983

3. 引文資料以逗號、句號及雙引號當成分割點，取第二長的字串當成可能的期刊名。

例如：*F. AURENHAMMER, "VORONOI DIAGRAMS - A SURVEY OF A FUNDAMENTAL GEOMETRIC DATA STRUCTURE," ACM COMPUTING SURVEYS, VOL. 23, NO. 3, PP. 345-405, 1991*

- (二) 刊名資料庫的比對子程序：將分析出來的可能期刊名跟刊名資料庫進行比對，比對結果符合刊名資料庫內的期刊名才算正確，並將引文資料類型設定為西文期刊。在比對之前需做字串整理，如去除標點符號、特殊字元及一些停用字 (Stop Word)，目的在增加比對的正確率。

例如：

字串整理前 *ACM TRANSACTIONS ON GRAPHICS,*  
字串整理後 *ACM TRANSACTIONS GRAPHICS*

- (三) WOS 刊名比對子程序：在上個程序未比對到確定的期刊名時執行，將上個程序分析出來的可能期刊名送至 WOS 查詢，查詢該期刊的文章一筆，來判定是否有該期刊存在。
- (四) 分析引文資料的期刊篇名子程序：西文期刊分析程序及 WOS 刊名比對程序都未比對出期刊名時，才執行本程序及下個程序 WOS 篇名比對程序，利用下列二種方法來分析出可能的期刊篇名

1. 引文資料有雙引號時，利用雙引號當成分割點，取最長的字串當成可能的篇名。

例如：*F. AURENHAMMER, "VORONOI DIAGRAMS - A SURVEY OF A FUNDAMENTAL GEOMETRIC DATA STRUCTURE," ACM COMPUTING SURVEYS, VOL. 23, NO. 3, PP. 345-405, 1991*

2. 引文資料無雙引號時，利用逗號及句號當成分割點，取最長的字串當成可能的篇名。

例如：P. J. BURT AND E. H. ADELSON. A MULTIREOLUTION  
SPLINE WITH APPLICATION TO IMAGE MOSAICS. ACM  
TRANSACTIONS ON GRAPHICS, 2 (4) :217-236, 1983

(五) WOS 篇名比對子程序：方法為利用上個程序分析出的可能期刊篇名送至 WOS 查詢，查詢是否有該篇名的文章，如有回傳資料則取得該文章的所屬期刊名稱。

### 三、引文類型分析程序

該模組主要功能在於分析判別引文類型，可判別的類型有網路資源、會議論文集以及可能為西文期刊或會議論文集，已判定為西文期刊的引文資料不執行該模組。

(一) 網路資源分析判別子程序：檢查引文資料是否有 HTTP: 或 HTTPS: 字串，如包含有該字串即判定類型為網路資源。

例如：*SONY ONLINE ENTERTAINMENT. "EVERQUEST". AVAILABLE FROM: HTTP://EVERQUEST.STATION.SONY.COM*

(二) 會議論文集分析判別子程序：檢查引文資料是否有字串 CONFERENCE、CONF、PROCEEDINGS、PROC、SYMPOSIUM 及 SYMP，如包含其中的字串即判定類型為會議論文。

例如：*M. PAN, C. CHU. FASTROUTE 2.0: A HIGH-QUALITY AND EFFICIENT GLOBAL ROUTER. IN ASIA AND SOUTH-PACIFIC DESIGN AUTOMATION CONFERENCE, PP.250-255, 2007*

(三) 可能為西文期刊或會議論文集分析判別子程序：檢查引文資料是否有字串 PP，如包含該字串即判定類型為西文期刊或會議論文集。

例如：*M. PAN, C. CHU. FASTROUTE 2.0: A HIGH-QUALITY AND EFFICIENT GLOBAL ROUTER. IN ASIA AND SOUTH-PACIFIC DESIGN AUTOMATION CONFERENCE, PP.250-255, 2007*

#### 四、刊名資料庫比對引文資料程序

該模組利用刊名資料庫來反向比對引文資料，如刊名資料庫內的期刊名字串包含在引文資料內，即判定可能為西文期刊，以便利使用者校正。

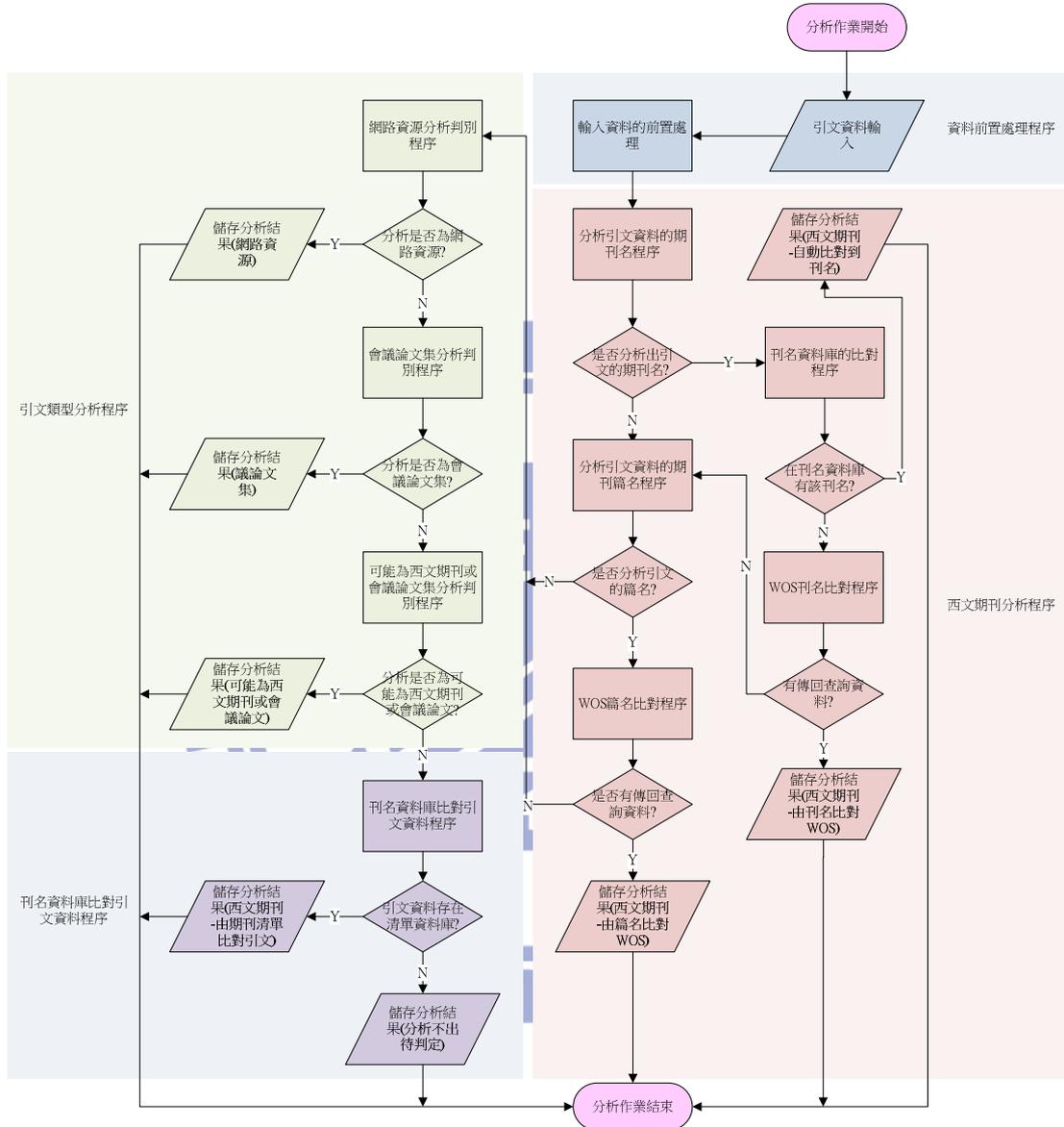


圖 3-14 參考文獻分析處理流程

### 3.3.5 WOS Web 服務

WOS 為美國 Thomson Reuters 於 1997 年間建置之網際網路版引用文獻索引資料庫系統，提供使用者理、工、醫、農、人文、及社會科學等各學科領域之文獻書目、作者摘要、及引用文獻等資料。該系統收錄期刊超過 10,000 種，每週更新其內容，每年提供超過 110 萬筆書目及 2300 萬筆引用文獻資料。WOS 共涵蓋下列三大資料庫：

#### 一、Science Citation Index Expanded (SCIE)

收錄科學技術類期刊計 7,000 餘種，涵蓋主題約 150 餘種。收錄年限於 2005 年，並已回溯自 1900 年，另外 1991 年起之資料包含作者摘要。

#### 二、Social Sciences Citation Index (SSCI)

收錄社會科學類期刊計 2,200 餘種，亦從近 3,300 種科技類期刊中挑選相關資料收錄，涵蓋主題約 50 餘種。收錄年限自 1956 年，另外 1992 年起之資料包含作者摘要。

#### 三、Arts and Humanities Citation Index (A&HCI)

收錄藝術人文類期刊計 1,200 餘種，亦從近 7,000 種科學和社會科學類期刊中挑選相關資料收錄，主題包括藝展評論、戲劇音樂及舞蹈表演、電視廣播等。收錄年限自 1975 年起，另外 1999 年起之資料包含作者摘要。

系統在設計參考文獻分析功能時，因應在刊名資料庫內未收錄的未知刊名，需要一個提供正確的書目資料以及期刊名稱的參考資料庫，於是收錄大量的引用文獻索引資料庫 WOS 就設計為系統參考的主要來源。本系統在執行分析階段，利用 WOS 所提供的 Web 服務進行資料查詢及校對，以輔系統的分析，故系統的架構會如圖 3-15 存取 WOS Web Service 網路架構所示。期刊評選系統因授權限制無法直接存取 WOS Web 服務，所以在分析階段會先連至交通大學的代理伺服器，經過帳號認證，通過之後再由代理伺服器的授權 Internet Protocol Address (IP Address)，透過 Simple Object Access Protocol (SOAP)，連線至 WOS Web Service

進行書目資料查詢，查詢結果再經由代理伺服器回傳至系統。

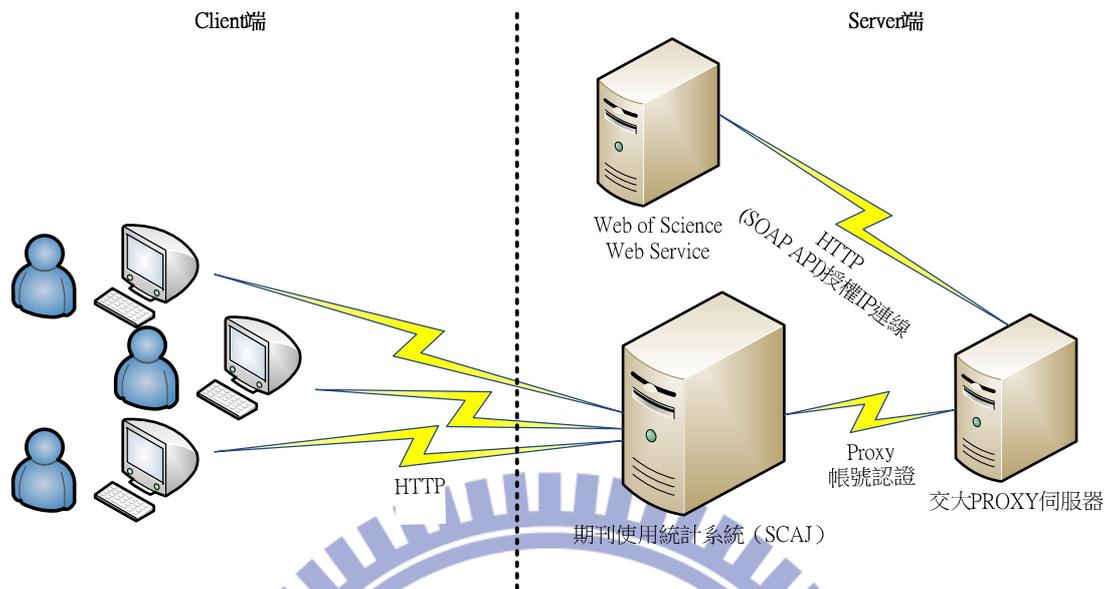


圖 3-15 存取 WOS Web Service 網路架構

WOS Web 服務提供 10 種作業 (Operations)，讓使用者自行撰寫程式利用 SOAP 簡單物件存取協定跟 WOS Web 服務進行資料查詢，本系統只使用 searchRetrieve 交易做為篇名及期刊名的查詢及比對，如表 3-2 WOS Web 服務 searchRetrieve 參數說明，使用 query 參數來查詢的範例如下：

- 一、查詢篇名，searchRetrieve ("WOS", TI= ("篇名"), "", "", "", 1, 1, "")，因查詢只要包含所查詢的篇名就會回傳結果，回傳的結果必需要再跟原本查詢的篇名再比對字串是否完全一樣，以防止回傳的結果錯誤。
- 二、查詢期刊名，searchRetrieve ("WOS", SO= ("期刊名"), "", "", "", 1, 1, "")，跟查詢篇名一樣是包含來運算，故回傳結果也需再做一次的字串比對。

表 3-2 WOS Web 服務 searchRetrieve 參數說明

Parameter	Description
<b>databaseID</b>	Identifies the WOK resource that this request will search. The "Codes and Descriptions" document for this resource provides more detailed descriptions of the other parameters in this document.
<b>query</b>	The search expression in WOK advanced search format. Please see the "Query Expressions" section below for more information.
<b>depth</b>	The time span that this search will cover such as 1 week, or 2000-2002.
<b>editions</b>	The editions that this search will cover.
<b>sort</b>	The sort order for record retrieval.
<b>firstRec</b>	The index of the first record this request will retrieve. The index of the initial record in the search result is 1.
<b>numRecs</b>	The number of records this request will retrieve. If the number of records in the search result is less than this value, then this request will return a smaller number of records. This value must not be greater than 100.
<b>fields</b>	The fields that will be retrieved for each record.

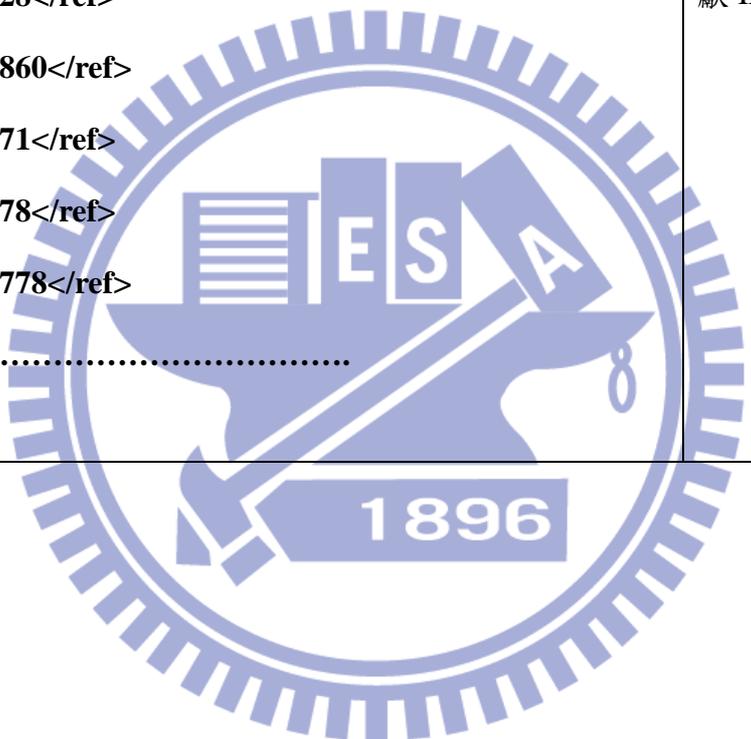
SOAP 簡單物件存取協定是基於 XML 的協定，故查詢結果回傳的 XML 資料更利於程式的後續處理，其回傳的 XML 資料範例重點說明如表 3-3 WOS Web 服務回傳 XML 內容。

表 3-3 WOS Web 服務回傳 XML 內容

XML Tag 範例	Tag 說明
<ut>000270345000023</ut>	文章 ID
<source_title>WORLD JOURNAL OF MICROBIOLOGY BIOTECHNOLOGY</source_title>	期刊名
<source_abbrev>WORLD J MICROB BIOT</source_abbrev>	期刊名縮寫
<item_title>Molecular cloning and characterization of multidomain xylanase from manure library</item_title>	篇名
<bib_id>25 (11) : 2071-2078 NOV 2009</bib_id>	卷期及日期
<article_no>DOI 10.1007/s11274-009-0111-6</article_no>	數位物件識別 號
<doctype code="@">Article</doctype>	文章類型
<languages count="1"> <primarylang code="EN">English</primarylang> </languages>	語言
<authors count="1"> <primaryauthor>Li, RP</primaryauthor> <fullauthorname> <AuRole>Author, Reprint Author</AuRole> <AuLastName>Li</AuLastName> <AuFirstName>Ruiping</AuFirstName> <address number="1">China Three Gorges Univ, Alan G MacDiarmid Res Inst Renewable Energy, Yichang 443002, Peoples R China</address>	作者資料

<p>&lt;AuCollectiveName&gt;Li, Ruiping&lt;/AuCollectiveName&gt;  &lt;/fullauthorname&gt;  .....  &lt;/authors&gt;</p>	
<p>&lt;keywords count="4"&gt;  &lt;keyword&gt;Xylan&lt;/keyword&gt;  &lt;keyword&gt;Xylanase&lt;/keyword&gt;  &lt;keyword&gt;Metagenomic DNA&lt;/keyword&gt;  &lt;keyword&gt;Carbohydrate binding module&lt;/keyword&gt;  &lt;/keywords&gt;</p>	<p>關鍵字</p>
<p>&lt;abstract avail="Y" count="1"&gt;  &lt;p&gt;The gene (manf-x10) encoding xylanase from an environmental genomic DNA library was cloned and expressed in Escherichia coli. The manf-x10 encoded a predicted protein of 467 amino acids residues with a molecular mass of 50.3 kD. Sequence analysis of manf-x10 gene revealed that the N-terminus had high homology to the catalytic domain of other bacterial xylanase enzymes. The optimal pH and temperature for xylanase activity were 7.0 and 40 degrees C, respectively. In the presence of 1 mM solution of Co<sup>2+</sup>, Fe<sup>2+</sup>, Mg<sup>2+</sup> and Zn<sup>2+</sup>, the relative xylanase activity was enhanced; however, it had almost no activity in the presence of 10 mM solution of Cu<sup>2+</sup>. The apparent K<sub>m</sub> and V<sub>max</sub> values obtained for the hydrolysis of rye arabinoxylan were 2.8 mg/ml and 49.5 mu</p>	<p>摘要</p>

<p>mol/min/mg, respectively. The C-terminus of the enzyme had high homology to a domain of unknown function found in several mannanase enzymes. Biochemical characterization of the C-terminus of the enzyme revealed a previously unrecognized carbohydrate binding module.&lt;/p&gt;</p> <p>&lt;/abstract&gt;</p>	
<p>&lt;refs count="33"&gt;</p> <p>&lt;ref&gt;69375728&lt;/ref&gt;</p> <p>&lt;ref&gt;135610860&lt;/ref&gt;</p> <p>&lt;ref&gt;73849571&lt;/ref&gt;</p> <p>&lt;ref&gt;90533778&lt;/ref&gt;</p> <p>&lt;ref&gt;148667778&lt;/ref&gt;</p> <p>.....</p> <p>&lt;/refs&gt;</p>	<p>文章的參考文 獻 ID</p>



## 第四章 系統實作與展示

將依據第三章系統分析與設計，進行期刊使用統計系統（SCAJ）的實作，並說明實作的成果以及系統展示。4.1 將說明系統平台的建置實作，4.2 展示實作的成果，包含系統各模組的功能以及系統的操作方式。

### 4.1 系統環境建置

本系統的開發環境是 Microsoft Visual Studio 2005，ASP.NET 2.0 C# 為開發程式語言，其原因為 ASP.NET 2.0 提供許多控制項（Control）可以簡化繁瑣的開發工作；並提供資料庫的控制項，可以快速方便的存取資料庫，與資料庫緊密結合。資料庫使用 Microsoft SQL Server 2005，跟 ASP.NET 2.0 搭配使用。WEB Server 則是採用安裝在作業系統（Operation System）Windows 2003 SERVER 下的 Internet Information Services（IIS）6.0。因近來虛擬化（Virtualization）技術成熟，在考量系統備份以及轉移方便，採取虛擬化的環境建置，如圖 4-1 系統環境建置架構所示，將系統建置在 hypervisor 的 VMware ESXi 伺服器虛擬化環境上，也就是伺服器上安裝 VMware ESXi，利用 VMware ESXi 模擬出虛擬的硬體環境，並在虛擬的硬體環境上安裝作業系統，其的好處是可以節省硬體設備來節能減碳，並提供系統快照（Snapshot），利用排程快速的備份整個系統，而整個作業系統是 VMware 通用格式檔，可快速遷移系統。

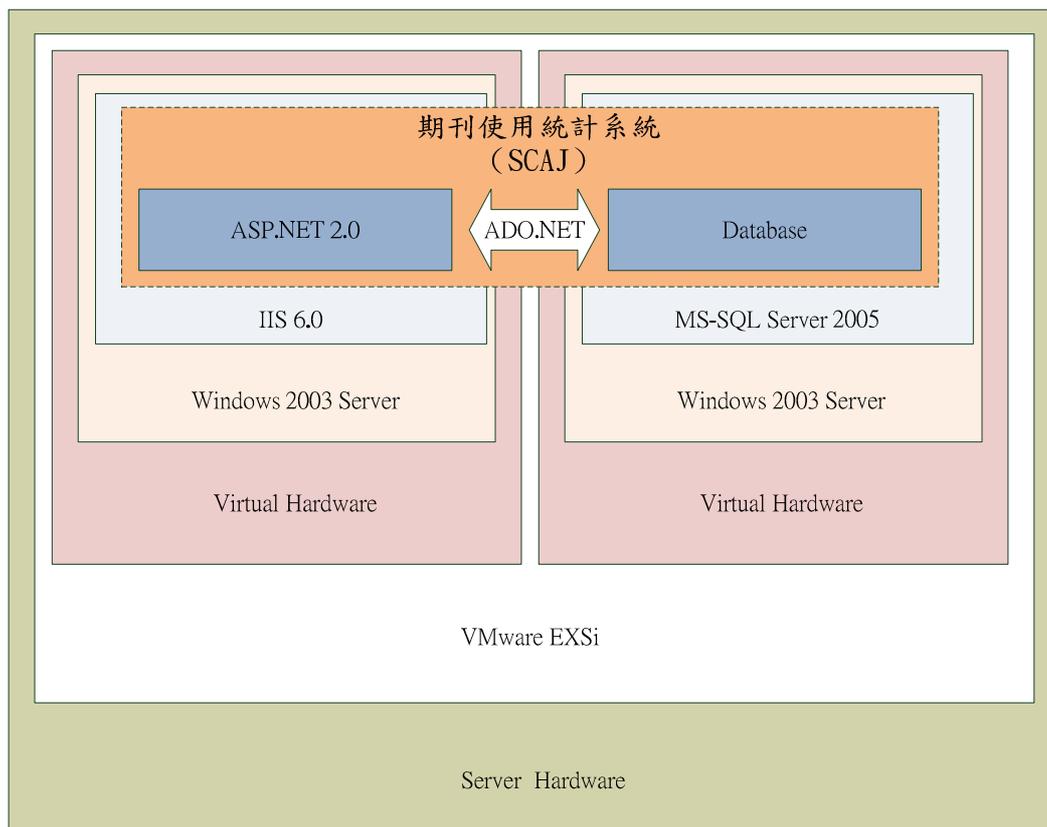


圖 4-1 系統環境建置架構



## 4.2 系統展示與功能介紹

本節將細分六個小節，以六大模組區分，就系統實作成果做詳細介紹。

### 4.2.1 系統模組

系統模主要功能為管理系統帳號及使用權限，讓管理者透過該模組來增刪使用者及設定權限。權限的設計依照系統分析的表 3-1 系統模組使用權限所規劃，權限設定以數字來表示，1 為管理者、2 進階使用者、以及 3 一般使用者。



圖 4-2 系統權限管理

## 4.2.2 刊名資料庫模組

刊名資料庫模組在系統設計上分為三個功能，以下就該模組的三個功能做詳細介紹：

### 一、刊名清單上傳功能

利用系統的刊名清單上傳功能，將蒐集到的期刊清單匯至刊名資料庫內，系統會利用 ISSN 以及期刊名稱為依據，來比對新增新的刊名或既有刊名的縮寫，並去除重覆部份，讓系統執行的比對準確率提高。該功能如圖 4-3 刊名清單上傳功能所示，選擇所要上傳的清單再進行上傳。匯入的資料範例如下需為 Unicode 編碼的純文字檔，並且提供二個欄位資料，1.期刊名、2.期刊的 ISSN，欄位之間以 Tab 字元分開。

<b>American Educational Research Journal</b>	<b>0002-8312</b>
<b>American Journal of Clinical Nutrition</b>	<b>0002-9165</b>
<b>American Journal of Epidemiology</b>	<b>0002-9262</b>



圖 4-3 刊名清單上傳功能

## 二、刊名資料庫管理功能

主要的功能用於管理刊名資料庫內的刊名資料，還有對刊名進行權威控制，其作業畫面如圖 4-4 刊名資料庫管理功能。權威控制是讓同一期刊，但可能有多個名稱或縮寫名稱，利用該功能將其合併，讓系統在統計時將其視為同一期刊，來產生正確的統計結果。

圖 4-4 展示了「刊名資料庫管理」系統的截圖。畫面顯示了期刊列表，包括期刊名稱、ISSN 和權威控制功能。圖中標註了「可選擇的刊名」和「已選擇要權威控的刊名」。

期刊名	ISSN	權威控制
Academy of Management Journal ACAD MANAGE J	0001-4273	加入
Academy of Management Perspectives (Formly: Academy of Management Executive)	1079-5545	加入
Academy of Management Review ACAD MANAGE REV	0363-7425	加入
ACCESS		加入
Accident Analysis and Prevention ACCIDENT ANAL PREV ACCID. ANAL. & PREV. ACCIDENT ANALYSIS & PREVENTION 33(2001)	0001-4575	加入
ACCIDENT PREVENTION		加入
ACCOUNT ORG SOC	0361-3682	加入
ACCOUNTING ORGANIZATIONS AND SOCIETY		加入
Accounting Review ACCOUNT REV	0001-4826	加入
Accounts of Chemical Research ACCOUNTS CHEM RES	0001-4842	加入
ACCREDIT QUAL ASSUR QUALITY ASSURANCE	0949-1775	加入

圖中顯示的權威控制操作表：

主要顯示	期刊名	變更主要顯示	移除權威控制
	ACCID. ANAL. & PREV.	選擇	移除
	ACCIDENT ANAL PREV	選擇	移除
	ACCIDENT ANALYSIS & PREVENTION 33(2001)	選擇	移除
	Accident Analysis and Prevention	選擇	移除

圖 4-4 刊名資料庫管理功能

### 三、新增刊名驗證功能

因系統為共用同一個刊名資料庫，為防止使用者在手動引文校正時，建立錯誤的刊名至刊名資料庫內，導致系統以錯誤的資料進行分析，故建立該新增刊名驗證功能。當權限等級為 3 的使用者，在引文作業時所產生的新刊名，為了防止錯誤資料的產生，需經權限等級為 1 或 2 的使用者在系統對該筆刊名確認，才會正式的儲存在刊名資料庫內；如否決該刊名，系統會將在該新增刊名時所做的異動都回復。其作業畫面如圖 4-5 新增刊名驗證功能所示。



圖 4-5 新增刊名驗證功能

### 4.2.3 引文模組

引文模組在系統設計上，將它分為引文資料上傳以及引文資料校正功能，以下就這二個功能做詳細介紹：

#### 一、引文資料上傳功能

引文資料的上傳分為二種格式，格式儲存的資料不同之處如表 4-1 上傳引文資料儲存比較說明。

表 4-1 上傳引文資料儲存比較

儲存資料	格式	Text 檔格式	XML 檔格式
引文資料		V	V
以年度自動分開儲存			V
研究生及論文相關資料			V
關鍵字資料			V
指導教授資料			V

(一) Text 檔格式：用於僅提供上傳引文的資料，不提供其它資訊，上傳的資料引文之間以換行分隔，每個檔案就包含一個群組的引文資料，其格式範例如下。

**1.Glover, S. W., Addison, John , Gleghorn, Colette , and Bramwell, John (2007) . Blackwell Publishing Ltd Journal Usage in NHS Hospitals: a Comparison Report of Total Usage at an Acute NHS Trust and a Specialist NHS Trust in the North West of England. Health Information and Libraries Journal, 24, 193-199.**

**2.Haycock, L. A. (2004) . Citation Analysis of Education Dissertations for Collection Development. Library Resources and Technical Service, 48 (2) , 102-106.**

(二) XML 檔格式：該功能提供博碩士學位論文格式的 XML 檔案上傳，為將 XML 博碩士論文引文資料上傳至系統，以及進行資料分析的作業，其內容的範例如下。

```
<book>
<Creator>梁詩婉</Creator>
<Creator.translated> Shih-wan Liang</Creator.translated>
<Title>研究型大學圖書館期刊評選模式之研究</Title>
<Title.translated> Citation Analysis Approach to Journal Collection of Academic Library</Title.translated>
<Contributor>黃明居</Contributor>
<Contributor.translated>Ming-jiu Hwang</Contributor.translated>
```

```

<Description.note.degree>碩士</Description.note.degree>
<Description.note.school>國立交通大學</Description.note.school>
<Description.note.department>數位圖書資訊</Description.note.department>
<Description.note.year>98</Description.note.year>
<Subject>引文分析</Subject>
<Subject>期刊評鑑</Subject>
<Subject>多因素分析</Subject>
<Subject.translated>Citation Analysis</Subject.translated>
<Subject.translated>Periodical Evaluation</Subject.translated>
<Subject.translated>Multiple-Factor Analysis</Subject.translated>
<Relation.reference>
1.Glover, S. W., Addison, John , Gleghorn, Colette , and Bramwell, John
(2007) . Blackwell Publishing Ltd Journal Usage in NHS Hospitals: a
Comparison Report of Total Usage at an Acute NHS Trust and a Specialist
NHS Trust in the North West of England. Health Information and Libraries
Journal, 24, 193-199.
2.Haycock, L. A. (2004) . Citation Analysis of Education Dissertations for
Collection Development. Library Resources and Technical Service, 48 (2) ,
102-106.
</Relation.reference>
</book>

```

資料上傳之後，系統會將 XML 資料抓出放至資料庫，參考文獻部份會以換行的識別字元做為每一筆資料的分隔。如上傳資料包含不同系所及不同年度，系統會自動識別的將各系所以年度分開儲存，如圖 4-6 引文資料上傳功能

所示，只上傳一個包含 94 至 96 三個年度的資料檔案，系統就自動識別分開儲存為三筆資料。另外會儲存研究生及論文相關、論文關鍵字中英文、論文指導教授中英文資料，提未來的相關研究使用。上傳完畢之後系統並不會立即執行分析功能，需自行點選”開始分析”來進行資料分析的作業，分析作業完畢之後就會消失在圖 4-6 引文資料上傳功能的分析檔案名稱清單內。

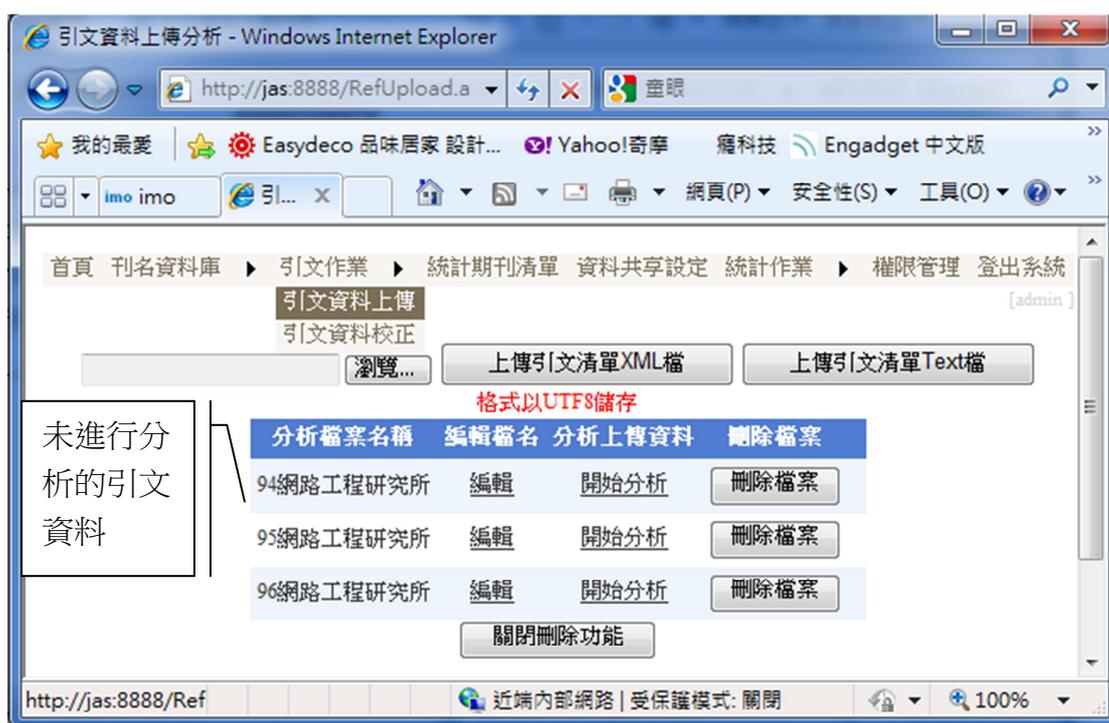


圖 4-6 引文資料上傳功能

## 二、引文資料校正

引文資料上傳功能的分析資料作業完畢之後，資料就會出現在引文資料校正功能的分析檔案名稱清單上，如圖 4-7 引文資料校正。選取要校正的分析檔案，就會出現下半部目前資料校正狀況清單，目的為讓使用者透過該功能進行人工校正，記錄已確認跟未確認的參考文獻跟類型，當已確認筆數等於全部筆數時，就代表全數確認完畢。

系統已分析完畢的檔案清單

分析檔案名稱	編輯檔名	檢視分析結果
92學年	<a href="#">編輯</a>	<a href="#">選取</a>
93學年	<a href="#">編輯</a>	<a href="#">選取</a>
94學年	<a href="#">編輯</a>	<a href="#">選取</a>
96學年	<a href="#">編輯</a>	<a href="#">選取</a>
95學年	<a href="#">編輯</a>	<a href="#">選取</a>

[開啟刪除功能](#)

校正狀況清單

分析狀態	全部筆數	已確認筆數	未確認	已確認
西文期刊(自動比對到刊名)	731	731	<a href="#">確認</a>	<a href="#">修改</a>
西文期刊(使用者建立刊名)	770	770	<a href="#">確認</a>	<a href="#">修改</a>
西文期刊(WOS)	231	231	<a href="#">確認</a>	<a href="#">修改</a>
會議論文集(自動)	920	920	<a href="#">確認</a>	<a href="#">修改</a>
會議論文集(使用者)	460	460	<a href="#">確認</a>	<a href="#">修改</a>
網路資源(自動)	598	598	<a href="#">確認</a>	<a href="#">修改</a>
網路資源(使用者)	13	13	<a href="#">確認</a>	<a href="#">修改</a>
圖書	786	786	<a href="#">確認</a>	<a href="#">修改</a>
中文期刊	6	6	<a href="#">確認</a>	<a href="#">修改</a>
其它(使用者)	30	30	<a href="#">確認</a>	<a href="#">修改</a>

圖 4-7 引文資料校正

圖 4-8 引文資料校正作業，為校正作業的主要頁面，類別的欄位預設值為系統判定的類別，如果錯誤時可以利用下拉式選單修改。如類別為期刊時，需檢視藍色字的刊名是否正確，如錯誤可按”修改”鈕進行刊名修正，類別為非期刊時則不會出現藍色字的刊名欄位，最後再按儲存與確認，系統就會將檢視過的資料記錄成已確認資料。



圖 4-8 引文資料校正作業

## 4.2.4 統計期刊清單模組

統計期刊清單模組的主要功能為管理統計時可以用到的期刊清單，清單可以加入所要統計分配的分數，如圖 4-9 統計期刊清單功能，匯入的資料跟刊名資料匯入一樣需為 Unicode 編碼的純文字檔，可提供三個欄位資料，1.期刊名、2.要給予的分數，其分數範圍由 1 到 100 分、3.ISSN，第三個欄位可以不提供，每個欄位間以 Tab 字元分開，範例如下。

<b>American Educational Research Journal</b>	<b>100</b>	<b>0002-8312</b>
<b>American Journal of Clinical Nutrition</b>	<b>90</b>	<b>0002-9165</b>
<b>American Journal of Epidemiology</b>	<b>80</b>	<b>0002-9262</b>

選取已上傳的清單，會由系統自動分析是否有期刊重覆的狀況，重覆的期刊資料會於報表作業時產生不正確的輸出結果，如有重覆的期刊會顯示在下方讓使用者修改。





圖 4-9 統計期刊清單功能

## 4.2.5 共享模組

該共享模組功能主要為提供使用者將自己的資料共享給其它使用者使用。如圖 4-10 共享模組功能，左邊為引文資料，右邊為統計期刊清單的分享，勾選分享並儲存設定之後，其他使用者就可以在統計作業模組中的合併統計報表功能中看到，並且利用分享的資料做為統計參考。



圖 4-10 共享模組功能

## 4.2.6 統計作業模組

統計作業模組主要功能為資料的統計，以及統計之後的比較，系統設計上將該模組分為三個功能，分別為引文統計報表、合併統計報表、以及排名比較功能。以下就這三個功能做說明：

## 一、引文統計報表

該統計報表可執行引用次數統計分為權威控制以及無權威控制，還有引用分類統計共三項統計，其選擇畫面如圖 4-11 引文統計報表選擇所示，選擇要統計的資料之後，再由下拉式選單選擇要統計的功能。進行引文統計之後產出如圖 4-12 引用次數統計結果的報表，系統是以引用次數的合計來排序，並顯示將每個年度引用次數，讓使用者可以瞭解分佈狀況。

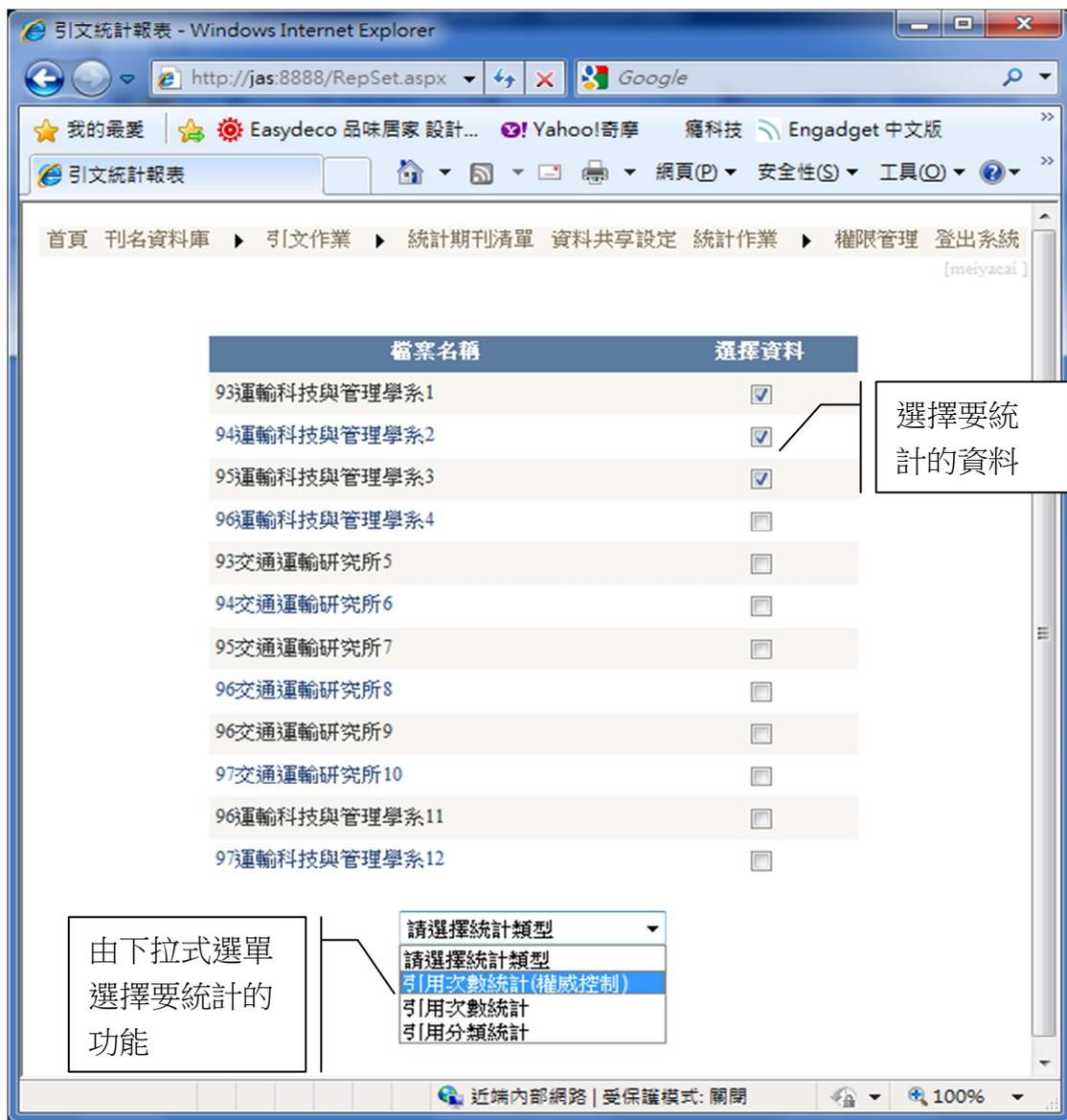


圖 4-11 引文統計報表選擇

引文統計報表 - Windows Internet Explorer

http://jas.8888/RepSetRe: Google

我的最愛 引文統計報表

首頁 刊名資料庫 ▶ 引文作業 ▶ 統計期刊清單 資料共享設定 統計作業 ▶ 權限管理 登出系統 [dudu]

10 匯出至Excel

JournalTitle	masterID	92學 年	93學 年	94學 年	96學 年	95學 年	sum
LECTURE NOTES IN COMPUTER SCIENCE (LNCS)	15368	67	54	64	142	45	372
IEEE Transactions on Pattern Analysis and Machine intelligence	689	65	91	69	17	71	313
IEEE Transactions on Computers	656	49	73	38	35	21	216
COMMUNICATIONS OF THE ACM	9060	30	39	44	30	28	171
IEEE Journal on Selected Areas in Communications	625	28	24	45	20	26	143
IEEE Transactions on Information theory	674	21	40	20	27	27	135
IEEE Transactions on Circuits & Systems for Video Technology	650	32	43	24	12	20	131
Pattern Recognition	1423	28	31	26	6	26	117
IEEE Transactions on Image Processing	670	29	26	16	7	29	107
IEEE Communications Magazine	608	14	23	35	17	17	106

1 2 3 4 5 6 7 8 9 10 ...

網際網路 | 受保護模式: 關閉 100%

圖 4-12 引用次數統計結果

進行分類統計之後產出的結果如圖 4-13 分類統計結果所示，其產出的報表可以直接在系統上瀏覽，或者是轉存成 Excel 檔，讓使用者使後續的參考及利用。



Type	ReferenceCounts
會議論文集(自動)	5579
西文期刊(使用者建立刊名)	5202
圖書	4745
會議論文集(使用者)	4231
西文期刊(自動比對到刊名)	3272
網路資源(自動)	3220
西文期刊(WOS)	456
網路資源(使用者)	142
其它(使用者)	117
中文期刊	54

圖 4-13 分類統計結果

## 二、合併統計報表

合併統計報表主要功能為利用已上傳的統計用期刊清單、校正過的引文資料、以及其他人分享出來的資料，給予不同的權重，由系統計算出合併統計的排名及分數，來做為期刊評鑑的參考之用。如圖 4-14 合併統計選擇所示，可選擇自己所上傳的統計用期刊清單及已校正過的引文資料，並分配給予不同的權重，但總權重合計必須為 100%。

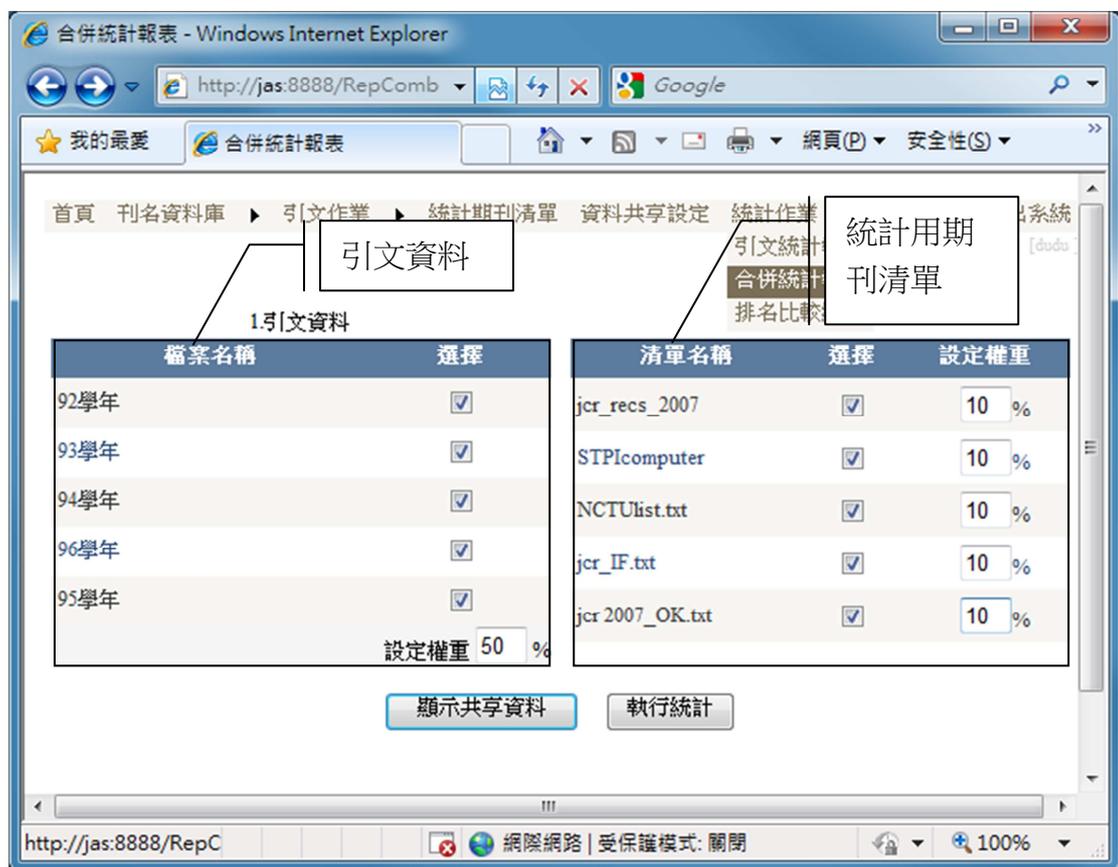


圖 4-14 合併統計選擇

如要將其他使用者分享出來的資料也納入參考，可點選顯示共享資料，則出現如圖 4-15 合併統計共享資料選擇，一樣是分配給予不同的權重，但總權重合計必須為 100%。

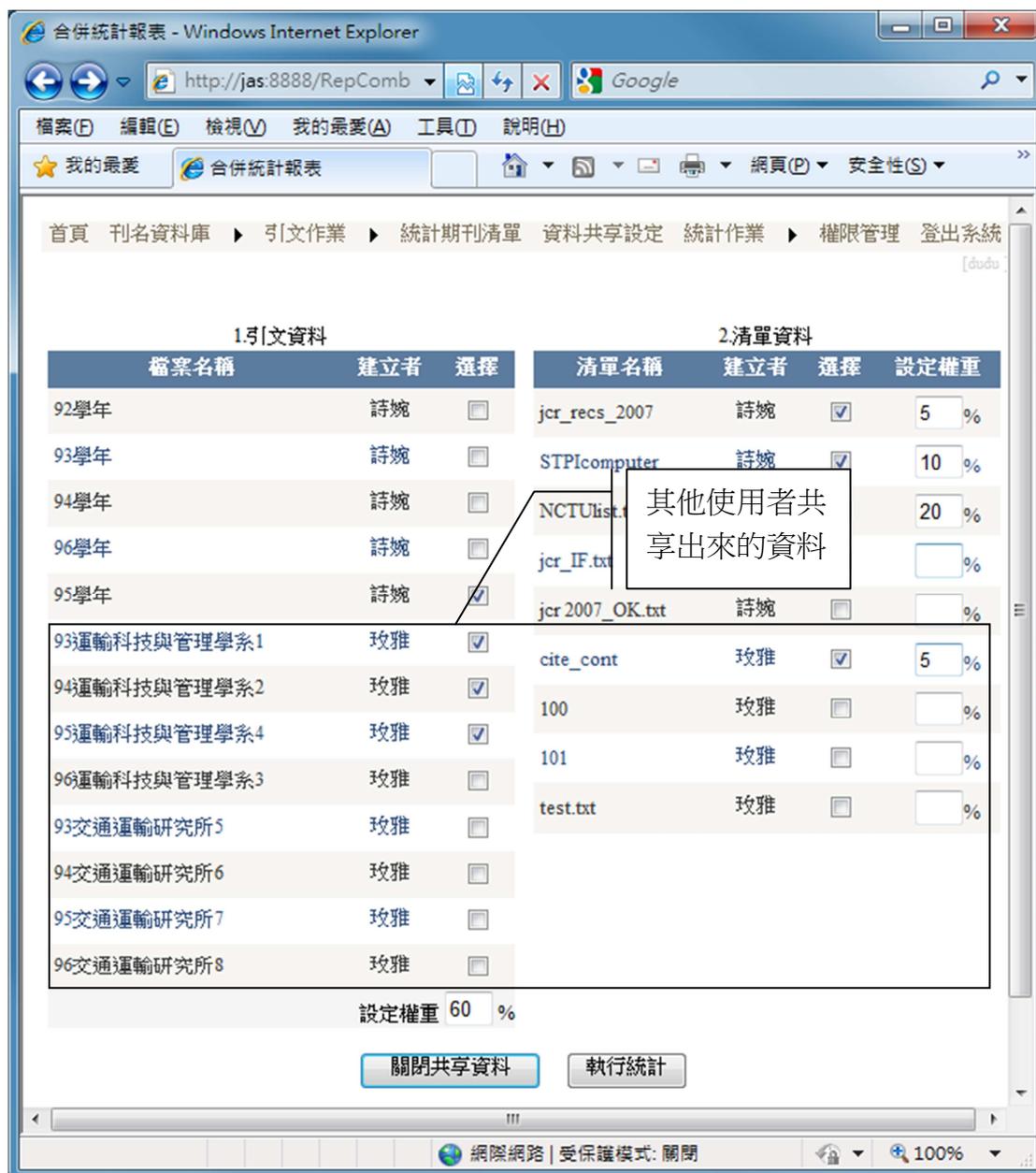


圖 4-15 合併統計共享資料選擇

執行統計之後產出的結果如圖 4-16 合併統計報表結果，系統會依據計算分數進行排名，並可以將計算結果匯出 Excel 檔、轉存結果至排名比較、以及轉存結果至統計用期刊清單。



圖 4-16 合併統計報表結果

### 三、排名比較功能

該功能主要是讓使用者用來比較多個已儲存的合併統計報表，可用來觀察及參考給予不同權重之後的排名變化。其操作將圖 4-16 合併統計報表結果所產生的報表轉存結果至排名比較，再進入圖 4-17 排名比較功能，選擇要排多比較的合併統計報表之後，系統會自動將期刊合併，最後產出如圖 4-18 排名比較結果報表。



圖 4-17 排名比較功能

排名比較統計 - Windows Internet Explorer

http://jas:8888/RepCompResu.aspx?fileID=4

我的最愛 排名比較統計

首頁 刊名資料庫 引文作業 統計期刊清單 資料共享設定 統計作業 權限管理 登出系統

15 匯出至Excel

title	journalID	100%引[文	92-94引[文	93-95引[文	94-96引[文	citations and JCR
Academy of Management Journal	6	672	489	515	0	611
Academy of Management Review	8	672	0	0	525	611
ACM Transactions on Information Systems	14	107	99	171	141	47
Acta Arithmetica	15	672	0	515	525	611
Advances in Computers	31	0	0	0	0	410
American Economic Review, the	42	367	489	515	260	611
American Educational Research Journal	43	367	271	275	0	611
American Journal of Epidemiology	45	367	271	275	260	611
American Journal of Sociology	50	152	271	171	116	431
American Mathematical Monthly, the	52	473	346	515	525	611
American Sociological Review	57	367	489	515	353	611
Annals of Statistics, the	77	672	489	515	525	611
Annual Review of Biochemistry	82	672	489	0	0	611
Annual Review of Fluid Mechanics	85	672	489	515	525	611
Applied Mathematics & Computation	95	123	113	120	116	410

1 2 3 4 5 6 7 8 9 10 ...

網際網路 | 受保護模式: 關閉 100%

圖 4-18 排名比較結果



# 第五章 系統評估

在本章節中將針對系統進行各項評估，以系統產出的數據來評估系統的準確率，並分析出系統的缺點，做為往後系統的改進參考。5.1 節以系統需求分析為依據來評估系統實際完成的狀況，5.2 節將說明資料分析評估的方式，5.2.1 小節將說明評估的資料為何，5.2.3 小節說明系統在西文期刊評估部份的正確程度，5.2.3 小節將討論系統加入 WOS 的分析參考之後對系統的改進程度。

## 5.1 需求完成度評估

完成系統的設計與實作之後，回顧 3.1 節，由圖書館館員提出的系統需求，所完成的需求分析，以該需求來進行完成度評估。系統依據分析的需求去設計及實作，如表 5-1 需求完成度評估所示，在需求完成狀況欄位說明系統於實作時以何種方式達成館員的系統需求。

表 5-1 需求完成度評估

	系統需求	需求完成狀況
資料需求	匯入參考文獻資料	提供純引文資料之 Text 檔格式，或者 XML 格式儲存的博碩士學位論文引文資料，供匯入系統分析
	統計用期刊清單	提供匯入或轉存統計用期刊清單，並且可以給予期刊分數，在統計時使用
	統計結果儲存	系統可將統計結果儲存，或是轉存成統計用期刊清單
功能需求	期刊權威控制	提供期刊權威控制介面，並自動將相同 ISSN 的期刊合併
	多使用者	利用登入帳號區分使用者，並可賦予不同權限
	資料共用	提供共享模組，可讓使用者的資料透過該模組，分享給其他使用者使用
	資料匯出 Excel	所有的統計報表皆提供匯出 Excel，以利使用者後續利用

介面需求	瀏覽器方式展現	瀏覽器上即可使用系統所有功能
	簡潔及功能兼具的作業介面	減少圖片以及不必要的訊息
	自訂顯示及處理資料數量	系統設計讓使用者自訂顯示的資料數量，以符合使用者的需求
	顯示資料篩選	會大量顯示資料的功能，都設計資料篩選功能
報表需求	引用次數統計	系統提供權威控制及排除權威控制引用次數的統計報表
	文獻類型統計	系統提供統計分析的參考文獻類型統計報表
	統計自訂權重	系統可讓使用者單獨或合併統計，引文資料及期刊清單資料，並透過自訂權重的方式，產出統計報表
	排名比較統計	系統提供排名比較統計報表，用來讓使用者觀查各儲存記錄的期刊排名變化

## 5.2 資料分析評估

本節將說明以資料分析的過程及結果，資料分析評估主要是以博碩士的參考文獻資料，匯入進行評估作業。

### 5.2.1 系統評估資料

系統的評估將採用以實際系統作業方式，用系統正規的做法，將資料匯入之後，再利用系統進行分析，最後再以人工來校正資料，校正後的準確資料跟系統分析出的初始狀態比較，以客觀的評估系統。系統在進行評估前，已先行匯入至刊名資料庫內 15,323 筆的刊名，其資料來源如下：

- 一、圖書館事業發展三年計畫建立之西文核心期刊清單—國家圖書館 92 年委由行政院國家科學委員會科學技術資料中心執行之圖書館事業發展三年計畫 92-02 研究成果報告所建立之西文核心期刊清單。
- 二、2007 年 SCI 及 SSCI 期刊清單。
- 三、2008 年交通大學、中央大學、清華大學以及陽明大學四校的西文期刊清單。

#### 四、交大資工系 92-96 學年度博碩士論文的引用文獻。

系統的評估來源資料是採用交通大學運輸科技與管理學系九十三至九十六學年，以及交通運輸研究所的九十三至九十六學年，共二系所八個學年的博碩士學位論文資料，合計 266 篇學位論文 12,851 筆參考文獻，如表 5-2 系統評估資料所示。依照順序先匯入運輸科技與管理學系九十三年，再進行系統的資料分析，分析完畢再進行人工校正，校正完畢之後再進行下個年度的匯入、分析及校正工作依此類推，之所以將每年度分析校正完再匯入下年度，其目的是用於評估資料庫在加入新的資料之後，每個年度的系統分析改進狀況。

表 5-2 系統評估資料

資料分析順序	畢業年	系所	學位論文數量	參考文獻數量
1	93	運輸科技與管理學系	40	1874
2	94	運輸科技與管理學系	33	1803
3	95	運輸科技與管理學系	28	1134
4	96	運輸科技與管理學系	42	2168
5	93	交通運輸研究所	30	1472
6	94	交通運輸研究所	28	1199
7	95	交通運輸研究所	34	1728
8	96	交通運輸研究所	31	1473

### 5.2.2 西文期刊分析評估

系統最主要是以西文期刊使用統計為主，西文期刊分析的正確程度為攸關本系統可用度的重要因素之一。本節將依據上一節所述，系統評估資料依序完成評估資料的作業所獲得的數據，針對西文期刊分析進行評估說明。表 5-3 西文期刊分析正確率，為在沒有人工介入的狀況下，由系統進行自動分析文獻類型為西

文期刊，並且比對到期刊刊名的正確率。評估結果顯示，系統分析西文期刊的平均正確率為 98%，也就是說如果系統判斷為西文期刊，並由刊名資料庫比對到該刊名，只會有 2% 的錯誤。表 5-3 西文期刊分析正確率，欄位說明如下：

- A. 參考文獻：該年度的參考文獻總筆數。
- B. 系統分析為西文期刊：在沒有人工介入的狀況下，由系統進行自動分析，判定文獻類型為西文期刊，並且比對到期刊刊名。
- C. 系統判定正確：由欄位 B.系統分析為西文期刊 的參考文獻，再進行人工校正後為確定是系統判定無誤的筆數。
- D. 正確率：系統針對西文期刊的分析正確率，其計算公式如下。

$$C. \text{分析後數量} / B. \text{系統分析為西文期刊} = D. \text{正確率}\%$$

表 5-3 西文期刊分析正確率

順序	畢業年 系所	A.參考文獻 (筆)	B.系統分析為 西文期刊(筆)	C. 系統判 定正確(筆)	D.正 確率
1	93 運輸科技與管理	1874	396	393	99%
2	94 運輸科技與管理	1803	640	629	98%
3	95 運輸科技與管理	1134	379	369	97%
4	96 運輸科技與管理	2168	682	656	96%
5	93 交通運輸研究所	1472	353	349	99%
6	94 交通運輸研究所	1199	315	311	99%
7	95 交通運輸研究所	1728	625	614	98%
8	96 交通運輸研究所	1473	593	589	99%
合計		12851	3983	3910	98%

表 5-4 西文期刊分析比率，為在沒有人工介入的狀況下，由系統自動的正確辨識出文獻類型為西文期刊，並且比對到期刊刊名，所佔全部的西文期刊文獻比率。評估結果顯示，系統的西文期刊分析平均比率為 75%，也就是在系統在匯入資料之後，經由系統進行自動的比對分析作業，在完全沒有人工校正的狀況之下，可以正確的分析辨識出 75% 的西文期刊及正確刊名。表 5-4 西文期刊分析比率，欄位說明如下：

- A. 參考文獻：該年度的參考文獻總筆數。
- B. 西文期刊：參考文獻類型為西文期刊筆數。
- C. 系統判定正確：由系統判定自動判定為西文期刊的文獻筆數。
- D. WOS 分析：系統利用 WOS 比對到的西文期刊文獻筆數。
- E. 分析比率：系統自動分析出的西文期刊，佔實際西文期刊的比率，其計算公式如下。

$$(C. \text{系統判定正確} + D. \text{WOS 分析}) / B. \text{西文期刊} = E. \text{分析比率}\%$$



表 5-4 西文期刊分析比率

順序	畢業年及系所	A.參考文獻 (筆)	B.西文期刊 (筆)	C.系統判定正確 (筆)	D.WOS 分析 (筆)	E.分析比率
1	93 運輸科技與管理	1874	739	393	106	68%
2	94 運輸科技與管理	1803	927	629	90	78%
3	95 運輸科技與管理	1134	533	369	56	80%
4	96 運輸科技與管理	2168	1100	656	90	68%
5	93 交通運輸研究所	1472	601	349	56	67%
6	94 交通運輸研究所	1199	443	311	24	76%
7	95 交通運輸研究所	1728	797	614	51	83%
8	96 交通運輸研究所	1473	831	589	57	78%
合計		12851	5971	3910	530	75%

依據資料分析順序來檢視各年的期刊分析比率，可由圖 5-1 西文期刊分析比率看出，運輸科技與管理學系的 93~95 年度各為 68%、78% 及 80% 逐年度上昇，但在 96 年度減少到 68%，在交通運輸研究所部份，93~95 年度各為 67%、76% 及 83% 也是逐年上昇期刊的分析比率，但跟運輸科技與管理學系同樣，96 年度就又減少至 78%，因參考文獻的內容及研究的趨勢，並不在本研究的範圍之列，故以目前的數據顯示，系統在前三年的期刊分析比率有逐年進步的狀況，最後一年的分析狀況不如預期有微幅下降的狀況。

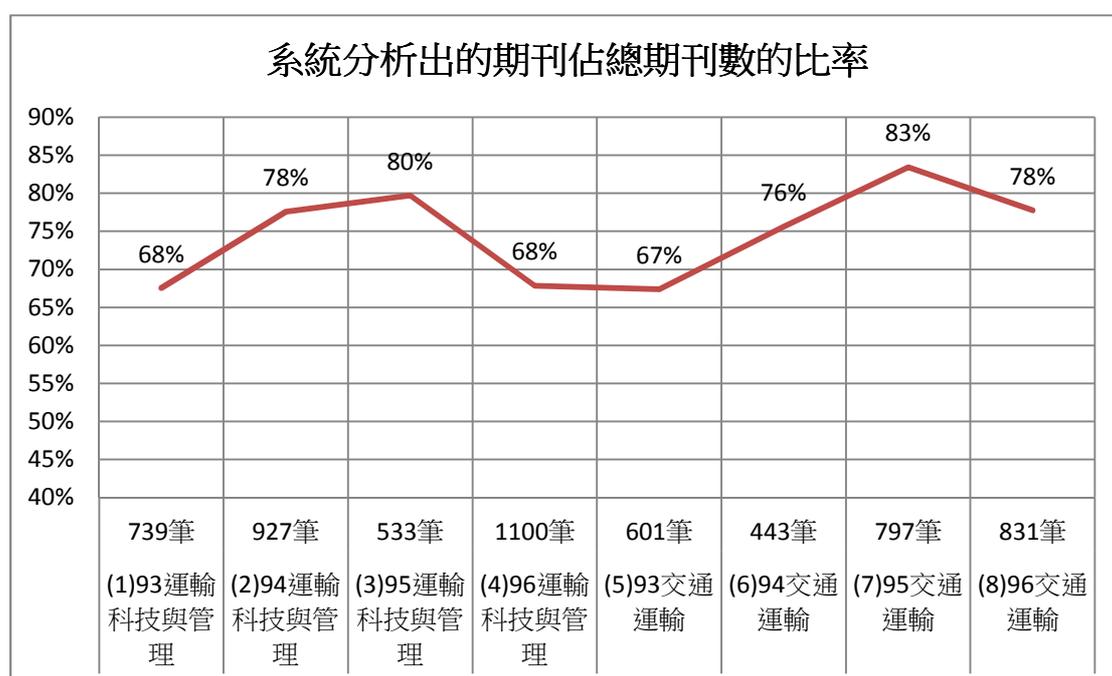


圖 5-1 西文期刊分析比率

系統分析出參考文獻為來源期刊佔實際參考文獻來源為期刊的比率為 75%，而在未分析出的 25% 之中則是由系統判定為如表 5-5 西文期刊未成功分析的分佈所揭示的各種類型，各類型欄位說明如下：

- A. 無法辨識：無法由系統分析判定該比參考文獻為何種類型。
- B. 未比到刊名：系統判定該參考文獻來源期刊，但無法判斷出該文獻的刊名。
- C. 清單比對：由系統中的刊名資料庫內的期刊名，反向比對該筆參考文獻是否

在期名資料庫之內。

- D. 刊名比對 WOS：利用 WOS 的刊名比對程序來分析，分析結果為期刊文獻，但比對到的刊名是錯誤。
- E. 篇名比對 WOS：利用 WOS 的篇名比對程序來分析，分析結果為期刊文獻，但比對到的刊名是錯誤。
- F. 會議論文或期刊：由系統判定為會議論文或期刊。

表 5-5 西文期刊未成功分析的分佈

順序	畢業年及系所	A.無法辨識	B.未比到刊名	C.清單比對	D. 刊名比對 WOS	E. 篇名比對 WOS	F.會議論文或期刊
1	93 運輸科技與管理	103	58	32	0	3	44
2	94 運輸科技與管理	46	54	25	1	9	68
3	95 運輸科技與管理	29	32	23	0	0	19
4	96 運輸科技與管理	64	59	177	0	0	46
5	93 交通運輸研究所	71	36	17	1	3	66
6	94 交通運輸研究所	20	33	29	0	5	18
7	95 交通運輸研究所	14	49	37	0	1	23
8	96 交通運輸研究所	49	38	88	0	2	7
合計		396	359	428	2	23	291

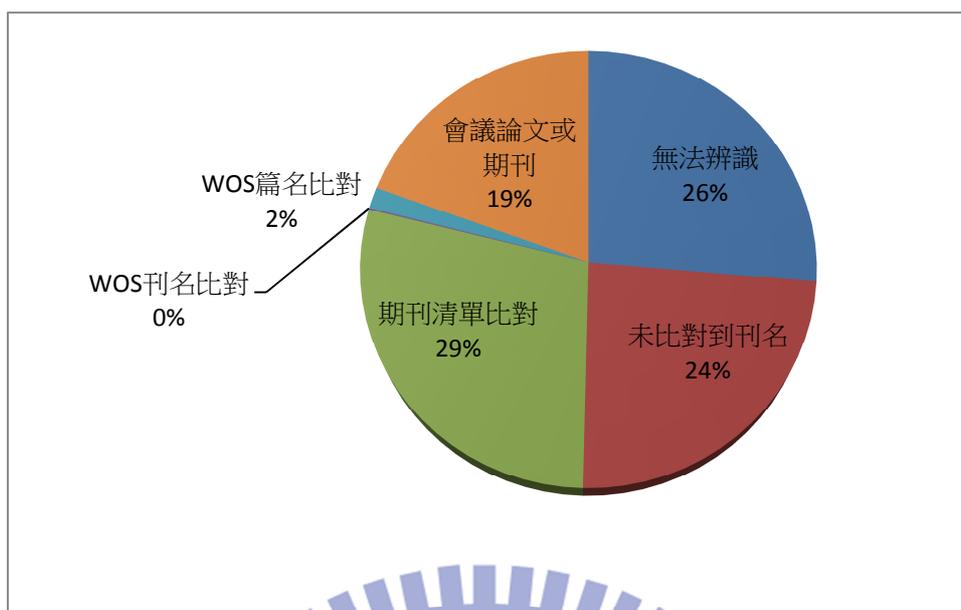


圖 5-2 西文期刊未成功分析的類型比率

### 5.2.3 WOS 分析評估

系統在西文期刊分析時，所利用到 WOS 比對程序進行評估，該程序是針對系統在參考文獻資料，抓取到可能是西文期刊刊名或者是期刊篇名字串，在系統的刊名資料庫比對不到該字串時就會利用該程序。表 5-6 系統利用 WOS 分析期刊的評估也是在人工校正後所完成的評估表，該表分為利用刊名比對跟篇名比對。

檢視評估表，刊名比對部份的準確率平均達到 89%，比篇名比對平均 83% 來的要高，但平均準確率仍達到 86%。不過比對利用 WOS 到的準確率雖高，但比對到的數量不算多，究其原因為刊名資料庫內的刊名資料在評估資料分析之前建立了許多的刊名資料在刊名資料庫內，導致在未進入 WOS 比對程序前就已判定出文獻的刊名。表 5-6 系統利用 WOS 分析期刊的評估，欄位說明如下：

- A. 刊名比對：利用 WOS 刊名比對程序所比對到的筆數。
- B. 刊名正確：利用 WOS 刊名比對程序比對到結果為正確的筆數。
- C. 刊名正確率：WOS 刊名比對程序的正確率，其計算方式為：

$$B. \text{刊名正確} / A. \text{刊名比對} = C. \text{刊名正確率}\%$$

D. 篇名比對：利用 WOS 篇名比對程序所比對到的筆數。

E. 篇名正確：利用 WOS 篇名比對程序比對到結果為正確的筆數。

F. 篇名正確率：WOS 篇名比對程序的正確率，其計算方式為：

$$E. \text{篇名正確} / D. \text{篇名比對} = F. \text{篇名正確率}\%$$

G. 新增刊名：利用 WOS 刊名以及篇名比對程序，新增在系統刊名資料庫內未收錄的期刊刊名。

表 5-6 系統利用 WOS 分析期刊的評估

順序	畢業年及系所	A.刊名比對	B.刊名正確	C.刊名正確率	D.篇名比對	E.篇名正確	F.篇名正確率	G.新增刊名
1	93 運輸科技與管理	29	28	97%	96	78	81%	24
2	94 運輸科技與管理	21	17	81%	95	73	77%	11
3	95 運輸科技與管理	5	3	60%	72	53	74%	3
4	96 運輸科技與管理	13	12	92%	83	78	94%	13
5	93 交通運輸研究所	11	9	82%	57	47	82%	8
6	94 交通運輸研究所	5	5	100%	25	19	76%	5
7	95 交通運輸研究所	13	13	100%	43	38	88%	8
8	96 交通運輸研究所	8	8	100%	55	49	89%	9
合計		105	95	89%	526	435	83%	81

檢示圖 5-3 WOS 刊名程序比對結果,以及圖 5-4 WOS 篇名程序比對結果顯示, WOS 比對程序所比對到的新增期刊刊名數量都持續下降,表示系統的分析的資料量愈多以及使用率愈高,系統內刊名資料庫內的資料就愈齊全,分析的準確性就會愈高。這次的評估資料利用 WOS 的比對程序,新增加了 94 筆刊名資料,也就是原本刊名資料庫內沒有的刊名,證實 WOS 的比對程序是有其存在的必要性。

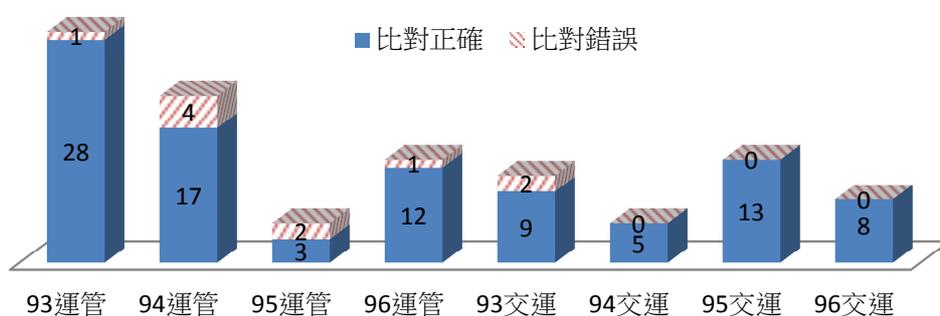


圖 5-3 WOS 刊名程序比對結果

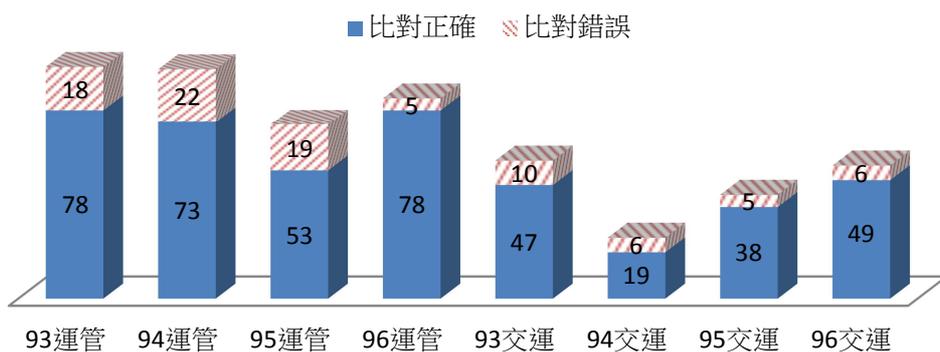


圖 5-4 WOS 篇名程序比對結果

## 第六章 結論與建議

本章針對研究提出總結，並提出建議以及未來可能的研究方向。6.1 節主要提出本研究的結論以及貢獻，6.2 節主要針對研究的結果、以及本研究未盡部份提出建議。

### 6.1 研究結論

探討圖書館期刊資源被使用的情況，一直以來都是圖書資訊學領域的一項重要研究。在期刊費用高漲以及採購經費縮減的當下，圖書館也利用各種方式想瞭解期刊資源的使用情況，期以在期刊採購時能以最低的代價，為最多的讀者，提供最佳的讀物。本研究針對上述課題進行探討，提出圖書館期刊使用統計之系統設計與實作，並且建置完成期刊使用統計系統 (SCAJ)，系統可廣泛應用於參考文獻資料的統計工作，協助解決參考文獻著錄凌亂問題，並減少人工作業時間。用於圖書館的期刊資源使用狀況時，可利用使用者的論文產出，統計論文的期刊引用，回推圖書館期刊資源的使用狀況，協助研究圖書館期刊使用率的問題。

總結本研究最主要的貢獻如下：

- 一、解決圖書館期刊使用率的統計問題。
- 二、協助分析凌亂的參考文獻著錄。
- 三、減少統計作業時間進行更大範圍的研究。
- 四、期刊評估輔助，加入期刊清單並給予不同的權重，產生期刊排名及分數。

## 6.2 研究建議

最後的研究建議將分為以下二個小節對研究的結果，以及研究未盡可加強之處提出建議。

### 6.2.1 研究結果建議

本研究的目的為，以參考文獻計量為基礎，設計出一套可協助研究人員或圖書館，進行論文參考文獻研究時較易於分析的系統。故以參考文獻計量為基礎，在研究進行的過程中，發現作者應該需對論文參考文獻著錄更為嚴謹。當所匯入系統的參考文獻資料著錄的方式正確時，系統針對西文期刊的分析正確率應該為相當的高，甚至不需要人工進行校對，由系統做自動分析就能取得可用的統計資料。所以本研究對典藏系統設計，以及論文作者提出以下建議：

- 一、典藏系統設計：在系統設計時可能考量使用者輸入的方便性，在參考文獻的部份都會以一個欄位放入全部的資料，最多以分筆的方式輸入參考文獻的資料，如能像常見的書目管理軟體一樣，規範及定義詳細的參考文獻輸入，雖然將造成輸入人員的負擔，但系統上的正確資料將可提高系統的參考文獻相關後續研究利用。
- 二、論文作者：作者在投稿時除了必須耗費心力在論文的內容之外，還必需配合出版單位所規定的引文格式，在百家爭鳴的引文格式當中，整理引文格式也是一個費時的工作。目前引文格式的整理可藉由書目管理系統的軟體來解決，例如 EndNote、Refernce Manger、RefWorks 及 ProCite 等，利用書目管理系統的功能，作者輕易的將引文格式轉變成所需的格式，利用這類的工具軟體來轉換引文格式可減少錯誤著錄的產生。目前許多著名的學術資料庫也都會與書目管理系統結合，將所需的書目資料快速的轉匯至書目管理系統，可節省許多作者在目書管理上的時間。

## 6.2.2 未來研究建議

進行完本研究之後，仍有許多在本研究未盡的部份，以下提出幾點可供未來進行相關研究的建議：

- 一、關鍵字、指導教授進行分析統計：匯入系統的 XML 檔案格式的博碩士學位論文引文資料，其內容除了包含參考獻資料之外，還包含許多學位論文的相關資訊。雖然目前系統已針對關鍵字、以及指導教授的資料存入系統中，但未對其資料進行分析統計，在未來系統可以針對這些資料做延伸的研究。
- 二、自由調整統計報表評分公式：目前系統的統計設計是以固定公式，對期刊進行評分，而使用者能調整的部份只有權重。未來如能讓使用者可以自由的設置評分公式，如此能讓使用者更有彈性靈活的應用系統。
- 三、持續增加系統的精準度：目前以系統評估來看期刊分析比率有 72%，未分析到的 28% 部份，其中絕大部份包含作者參考文獻著錄的錯誤，但著錄錯誤只能待作者的論文寫作素養改善才可能獲得改進。未來如能加入一些近似字串比對演算法在分析的程序中，應能改善一些因錯誤的著錄而未分析到部份。
- 四、增加外部參考：目前系統的外部參考只利用到 WOS，未來如能再加入其它收錄大量期刊的資料庫，例如 SCOPUS 索摘資料庫，如此在未知的期刊分析上應能更加的準確。

## 參考書目

- Andrew, Lim; Hong, Ma; Qi, Wen; Zhou, Xu;. (n.d.) . *Journal-Ranking.com: An Online Interactive Journal Ranking System*. Retrieved 2009, from the World Wide Web:  
<http://www.journal-ranking.com/ranking/web/docs/whitepapers/iaai01lima.pdf>
- APA Style. (n.d.) . Retrieved 2010, from the World Wide Web:  
<http://www.apastyle.org/>
- Dhawan, S. M., Phull, S. K., & Jain, S. P. (1980) . Selection of Scientific Journals: A Model. *Journal of Documentation*, 36 ( 1 ) ,24-32.
- Edward, G., & Sarah, H. (2008) . Comparison of Journal Citation Reports and Scopus Impact Factors for Ecology and Environmental Sciences Journals. *Issues in Science and Technology Librarianship*, 54, 5.
- Eigenfactor Methods. (n.d.) . Retrieved 2009, from the World Wide Web:  
<http://www.eigenfactor.org/methods.htm>
- Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008) . Comparison of SCImago journal rank indicator with journal impact factor. *The FASEB Journal*, 22 ( 8 ) , 2623-2628.
- Guo, G. M. (2007) . A computer-aided bibliometric system to generate core article ranked lists in interdisciplinary subjects. *Information Sciences*, 177 ( 17 ) , 3539-3556.
- Journal Citation Reports. (2005) . Retrieved 2010, from the World Wide Web:  
[http://science.thomsonreuters.com/m/pdfs/mgr/jcr4\\_sem\\_0305.pdf](http://science.thomsonreuters.com/m/pdfs/mgr/jcr4_sem_0305.pdf)
- Lewis, J. S., & McDonald, J. D. (2002) . Defining an undergraduate core journal collection. *Serials Librarian*, 43 ( 1 ) , 45-59.
- McCain, K. W. (1991) . Core journal networks and cocitation maps: New bibliometric tools for serials research and management. *Library Quarterly*, 61 ( 3 ) , 311-336.
- MLA Style. (n.d.) . Retrieved 2009, from the World Wide Web:  
<http://www.mla.org/style>
- Scopus. (n.d.) . Retrieved 2010, from the World Wide Web:  
<http://www.info.sciverse.com/scopus/>
- SCImago Journal Rank. (2010) . Retrieved 2010, from the World Wide Web:  
<http://info.scopus.com/journalmetrics/sjr.html>
- Shiue, Y. C., Chang, R. I., & Guo, G. M. (2004) . A Computer-Aided Bibliometrics System for Journal Citation Analysis and Departmental Core Journal Ranking List Generation. *Journal of Educational Media & Library Sciences*, 42 ( 2 ) ,

199-219.

*Source Normalized Impact per Paper*. (2010). Retrieved 2010, from the World Wide Web: <http://info.scopus.com/journalmetrics/snip.html>

*Style Guides*. (n.d.). Retrieved 2010, from the World Wide Web: <http://www.aresearchguide.com/styleguides.html>

*The Chicago Manual of Style*. (n.d.). Retrieved 2010, from the World Wide Web: <http://www.chicagomanualofstyle.org/home.html>

*Web Services Tutorial*. (2009). Retrieved 2009, from the World Wide Web: <http://www.w3schools.com/webservices/default.asp>

*WOS Webservices*. (n.d.). Retrieved 2009, from the World Wide Web: <http://scientific.thomson.com/ts/media/faq/webservices.zip>

王梅玲 (民 94)。大學研究評鑑與引文分析應用。圖書館學與資訊科學, 31 (1), 5-13。

石秋霞、黃鴻珠 (民 95)。學術論文被引用次數之分析研究：以 Science Citation Index Expanded 及 Scopus 為例。教育資料與圖書館學, 44 (1), 15-34。

石美玉 (民 92)。建立西文核心期刊清單暨紙本及電子期刊館藏合作發展計畫研究成果報告。圖書館事業發展三年計畫 92-02 合作發展圖書館館藏計畫。臺北市：國家圖書館。

吳英美 (民 87)。我國三所師範大學圖書館教育類西文期刊館藏之研究 (上)。國立中央圖書館臺灣分館館刊, 5 (1), 58-73。

吳英美 (民 87)。我國三所師範大學圖書館教育類西文期刊館藏之研究 (下)。國立中央圖書館臺灣分館館刊, 5 (2), 50-67。

孟連生 (民 85)。試問引文索引法的性質與功能。資訊傳播與圖書館學, 3 (1), 15-19。

邱炯友 (民 94)。學術論文引文格式之發展與使用調查。中華民國圖書館學會會報 (75), 172-183。

邱炯友、蔡孟倫 (民 99)。學術論文引文格式對於學術資料庫電子文獻規範之探討：以 APA、MLA、Chicago/Turabian 為例。圖書資訊學研究, 4 (2), 41-64。

胡述兆、王梅玲 (民 92)。圖書資訊學導論。台北：漢美。

梁詩婉 (民 98)。研究型大學圖書館期刊評選模式之研究 (碩士論文)。國立交通大學：新竹市。

陳光華 (民 90)。學術期刊引用文獻資料庫之現況與建置。大學圖書館, 5 (1), 33-48。

陳貽中、楊正仁 (民 90 年 10 月)。以估測相似度為基礎的近似字串比對。NCS 全國計算機會議。台北中國文化大學。

黃慕萱 (民 94)。國內七所研究型大學論文發表概況分析。圖書與資訊學刊, 55, 0-23。

楊美華 (民 77)。大學圖書館西文期刊之評量。圖書館學與資訊科學, 14 (2), 208-231。

楊美華 (民 79)。核心期刊館藏之建立。教育資料與圖書館學, 28 (2), 207-236。

詹麗萍 (民 95 年 1 月)。引文分析與電子期刊使用評鑑。數位時代圖書館館藏發展研討會論文集。台中市中興大學圖書資訊學研究所。

蔡孟倫 (民 97)。學術論文引文格式對於學術資料庫電子文獻規範之探討 (碩士論文)。淡江大學: 台北縣。

蔡明月 (民 93 年 5 月)。引用分析與學術評鑑。引文分析與學術評鑑研討會。台北市國立政治大學。

蔡明月 (民 94)。引文索引與引文分析之探討。圖書館學與資訊科學, 31 (1), 45-536。

蘇蓉波 (民 92)。從博士學位論文引用文獻探討成功大學圖書館館藏支援程度。國立成功大學圖書館館刊 (12), 53-69。

