

國立交通大學

資訊學院 數位圖書資訊學程

碩士論文

中國古籍文件校勘系統設計與開發之研究

The Design and Development of Web-based Chinese Ancient
Book Automatic Version Comparison System

研究生：陳堡皇

指導教授：黃明居 教授

中華民國 101 年 6 月

中國古籍文件校勘系統設計與開發之研究

The Design and Development of Web-based Chinese Ancient Book Automatic Version Comparison System

研究生：陳堡皇

Student : Pao-Huang Chen

指導教授：黃明居

Advisor : Dr. Ming-Jiu Hwang

國立交通大學
資訊學院 數位圖書資訊學程
碩士論文



A Thesis
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 2012
Hsinchu, Taiwan, Republic of China

中華民國 101 年 6 月

中文摘要

中國古籍文件校勘系統設計與開發之研究

研究生：陳堡皇

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



中國古籍在歷經多次朝代與時間的傳抄、轉印與轉刻狀況下，往往會出現一種古籍擁有多個版本，後代古籍之校勘研究者需花費許多人力逐字比對與核對，一般使用者亦無法了解各版本之差異為何。本研究是利用電腦快速、記憶、比對的技術，以及網際網路無遠弗屆的特性，建置一個讓研究者與一般使用者容易使用的校勘系統平台，除了協助古籍研究較容易實施註釋及校勘等工作，並且可容易發現大量人工校勘很難查察的錯誤，節省時間與提高校勘的質與量之外，亦讓對於古籍有興趣的使用者可以同時閱讀不同版本的古籍，並且閱讀其中之差異。

關鍵字：古籍、註釋、校勘

英文摘要

The Design and Development of Web-based Chinese Ancient Book Automatic Version Comparison System

Student : Pao-Huang Chen

Advisor: Dr. Ming-Jiu Hwang

Degree Program of Computer Science

National Chiao Tung University



The Chinese nation has a long history of five thousand years of civilization which leaving many precious cultural relics and books. Especially the ancient books which made copies and transferred through many dynasties have many versions in the same book. The results bring confusion and trouble to users and researchers. Therefore, the present study utilizes the computers technology and Internet borderless features, including the speed, memory and comparison to design a systematic platform that easy to use for researchers and general users. After inserting data to the system, it will automatically show different words or paragraph in the different ancient book versions and help researchers make annotation and textualism. In addition, the users can figure out many errors which difficult to find manually. It also improves the quality and quantity of textualism and reaches the goal of automatic version comparison.

Keywords : Ancient book, Annotation, Automatic version comparison

誌 謝

感謝求學過程中指導教授黃明居老師在各方面的悉心指導與協助，給我很多研究論文的方向，更讓我看到他對學術領域的堅持與期望，讓我可以為此研究領域更堅定自己的方向；另外，還要感謝擔任口試委員的柯皓仁老師與林信成老師，對論文未周延部份給予許多建議與指導，使本研究論文更臻完善。

此外，感謝相互鼓勵的揚喬、薇雅和易賢同學，在學術和技術上給予建議，以及當初不斷鼓勵我報考研究所的文華學長，除了在工作上給予支援之外，也適時關心課業及論文進度，使得我得以順利的完成論文。

最後，還要感謝家人長久以來的支持與鼓勵，在我工作與課業上徬徨無助或沮喪時，適時給予溫暖與解決方法，讓我無後顧之憂，專心致力於研究，並得以順利完成學業。

目錄

中文摘要.....	i
英文摘要.....	ii
誌謝.....	iii
目錄.....	iv
表目錄.....	vi
圖目錄.....	vii
第一章 緒論.....	1
1.1 研究動機與目的.....	1
1.2 研究內容.....	2
1.3 研究範圍與限制.....	3
1.4 研究流程.....	3
1.5 論文架構.....	5
第二章 文獻回顧與系統功能需求.....	6
2.1 文獻回顧.....	6
2.1.1 古籍定義.....	7
2.1.2 古籍內容數位化發展.....	8
2.1.3 校勘、校讎學.....	16
2.1.4 校勘的方法.....	18
2.1.5 古籍自動校勘.....	21
2.2 系統功能說明.....	23
2.3 使用的程式工具.....	24
2.3.1 AppServ.....	24
2.3.2 PHP.....	28

2.3.3	MySQL.....	32
第三章	系統分析與設計.....	35
3.1	系統設計概念.....	35
3.2	設計架構.....	36
3.3	系統模組設計.....	37
3.3.1	古籍文件管理模組.....	39
3.3.2	古籍文件校勘模組.....	40
3.3.3	古籍內容註釋模組.....	43
3.3.4	輔助工具模組.....	45
3.3.5	個人收藏模組.....	46
3.4	資料表設計與欄位定義.....	47
第四章	系統功能展示.....	54
4.1	系統實作環境.....	54
4.2	系統實作結果.....	55
4.2.1	古籍文件管理功能.....	57
4.2.2	古籍文件校勘功能.....	62
4.2.3	古籍內容註釋功能.....	68
4.2.4	輔助工具功能.....	70
4.2.5	個人收藏功能.....	71
4.3	系統功能完成度與效益.....	72
第五章	結論與建議.....	73
5.1	研究結論與貢獻.....	73
5.2	建議.....	74
參考文獻	76

表目錄

表 2-1：古籍的定義	8
表 2-2：PHP 資料庫支援一覽表.....	31
表 3-1：校勘系統使用者資訊欄位一覽表	48
表 3-2：古籍類別名稱資訊欄位一覽表	48
表 3-3：古籍列表名稱資訊欄位一覽表	48
表 3-4：古籍篇名名稱資訊欄位一覽表	49
表 3-5：古籍內容資訊欄位一覽表	49
表 3-6：古籍註釋資訊欄位一覽表	50
表 3-7：古籍校勘結果資訊欄位一覽表	50
表 3-8：字詞庫名稱欄位一覽表.....	51
表 3-9：儲存字詞庫內容資訊欄位一覽表	51
表 3-10：個人收藏資訊欄位一覽表.....	52
表 4-1：系統功能完成概況表	72

圖目錄

圖 1-1：研究流程圖	4
圖 2-1：本研究文獻資料整理	6
圖 2-3：中華電子佛典學會「漢文電子大藏經」	13
圖 2-4：文淵閣《四庫全書》全文檢索系統	15
圖 2-5：自動校勘核心計算方法流程圖	22
圖 2-6：古籍自動校勘系統模組圖	23
圖 2-7：Apache Service、PHP、MySQL、User 使用關係圖	25
圖 2-8：PHP 程式碼片段	26
圖 2-9：MySQL 命令模式介面	27
圖 2-10：phpMyAdmin 管理介面	28
圖 2-11：PHP 使用趨勢狀況	29
圖 2-12：CGI 關係圖	30
圖 2-13：資料庫階層式架構圖	33
圖 3-1：系統設計概念圖	35
圖 3-2：MVC 系統設計架構圖	37
圖 3-3：系統模組設計架構圖	38
圖 3-4：古籍文件管理模組流程圖	40
圖 3-5：古籍文件校勘模組（對校法、本校法、他校法）流程圖	41
圖 3-6：對校法、本校法、他校法流程圖	42
圖 3-7：古籍內容註釋模組流程圖	44
圖 3-8：文字定位流程圖	45
圖 3-9：輔助工具模組流程圖	46
圖 3-10：個人收藏模組流程圖	47

圖 3-11：資料表關聯圖.....	53
圖 4-1：系統實作環境圖.....	55
圖 4-2：系統功能架構圖.....	56
圖 4-3：古籍文件自動校勘系統頁面.....	57
圖 4-4：古籍類別、列表、篇名「新增」功能.....	58
圖 4-5：古籍類別、列表、篇名建立完成顯示.....	58
圖 4-6：古籍文件上傳功能.....	59
圖 4-7：古籍列表建立.....	59
圖 4-8：產生建立的古籍列表資料夾.....	60
圖 4-9：古籍篇名文件內容上傳.....	60
圖 4-10：顯示已上傳的文件內容.....	60
圖 4-11：古籍「類別」、「列表」與「篇名」項目中編輯與刪除功能..	61
圖 4-12：校勘功能選擇頁面.....	62
圖 4-13：對校法概念圖.....	63
圖 4-14：對校法版本選擇.....	63
圖 4-15：對校法之顯示結果.....	64
圖 4-16：本校法概念圖.....	65
圖 4-17：選擇本校法的底本與待校勘的相似段落.....	65
圖 4-18：本校法之顯示結果.....	66
圖 4-19：他校法概念圖.....	67
圖 4-20：其他相關古籍選擇.....	67
圖 4-21：他校法之顯示結果.....	68
圖 4-22：古籍註釋工具「新增註釋」.....	69
圖 4-23：古籍文件中具有註釋的範圍.....	69
圖 4-24：古籍註釋工具「本篇註釋」.....	70

圖 4-25：輔助工具表建立畫面	70
圖 4-26：輔助工具建立與管理畫面	71
圖 4-27：開啟輔助工具功能與關鍵字詞的解釋	71
圖 4-28：個人工具「註釋收藏」顯示畫面.....	72



第一章 緒論

本章描述本研究的動機與目的，以及希望解決的問題。1.1 節說明研究動機與本研究希望達成的目的；1.2 節為研究內容，以過去古籍進行校勘的方式，利用電腦技術協助校勘；1.3 節為研究的範圍與限制；1.4 節為研究流程；1.5 節則說明本論文各章節的內容架構。

1.1 研究動機與目的

中華民族是具有五千多年文明史的悠久民族，其先人所留下的珍貴文物與典籍資產浩瀚無窮，尤其是這些古籍在歷經多次朝代與時間的傳抄、轉印與轉刻，往往會出現各種謬誤，例如：一種古籍擁有多個版本、各個版本之間的內容差別很大的現象，而已經電子化、數位化的古籍中，往往因為版本不明、錯（漏）字以及使用簡體字，造成研究者或使用者的困擾。

因此，如何利用資訊數位化及網際網路的便利環境，建置一個容易運用和閱讀的系統平台，並利用電腦自動發現且標識出不同版本之古籍間的文字差異，以及提供校勘工作所需的各種輔助工具，可以找出大量人工校勘不易發現的錯誤，節省人工校驗時間；且也可以透過此系統，讓古籍文件與文件之間做資料的連結，加速使用者掌握涵義與尋找更多的相關資料。

另一方面，一般的使用者不容易同時閱讀不同版本的古籍，且看出差異之處，因此，如何設計一套系統，可匯入多種不同版本的古籍文件資料，提供可運用的勘誤工具，讓使用者達到快速閱讀、標記與註釋，為本研究主要的動機所在。

本研究的目的包括：

- 1.設計與建置一個古籍文件自動校勘平台，探討相關技術與功能，並利用此平台縮短研究者校勘時間，提高校勘的質與量。
- 2.提供一份較為可靠的電子本古籍文獻，使瀏覽於網路上的校勘研究者或對古籍有興趣的人校勘古籍與了解。
- 3.藉此系統使對古籍有興趣的人學習和培養閱讀與校勘古籍文獻的知識與能力。



1.2 研究內容

為達成本研究之目的，研究內容包括：

- 1.回顧過去如何進行古籍校勘、校勘方法的說明及演進。
- 2.探討利用電腦比對技術，系統如何顯示並加以記錄，幫助古籍校勘工作。
- 3.透過電腦校勘系統比對出差異的文字後，可予以註釋。
- 4.建置古籍文件自動校勘系統，提供一個增加提高校勘效率的平台。
- 5.測試與評估系統是否可正確顯示校勘所需的結果。

1.3 研究範圍與限制

由於研究者可以取得資源有限，為避免研究問題過於複雜，且保持研究之單純性，本研究有以下限制：

1. 以純文字檔(.txt)上傳至系統，並轉成網頁格式後處理文字間差異性的比對，尚未擴及至多種線上文件格式。
2. 系統開發環境以網際網路作業環境為主。
3. 系統設計主要是以「對校」、「他校」及「本校」等三種校勘方法為主，至於陳垣第四種校法之「理校」，屬讀書心法，電腦尚無法透過經年研讀所得之領會，因而未能取代人工。
4. 有關古今字形與原始檔案缺漏字問題，不在本研究範圍。

1.4 研究流程

研究者觀察到目前校勘者進行校勘比對時，校勘資料的來源大致上有電子資料庫(如愛如生漢籍資料庫、四庫叢刊)、前人研究成果(如校勘記、博碩士論文、研究專書)，獲得資料後多數仍採用人工方式實施校勘，在瞭解現行校勘方法與特性後，擬定系統需求功能並加以設計，接著將古籍文件資料匯入系統測試，並評估顯示內容是否正確，最後提出結論與未來可精進方向；以上過程區分「界定研究主題」、「系統設計」、「系統評估」及「結論與建議」等四個階段，研究流程如圖 1-1：

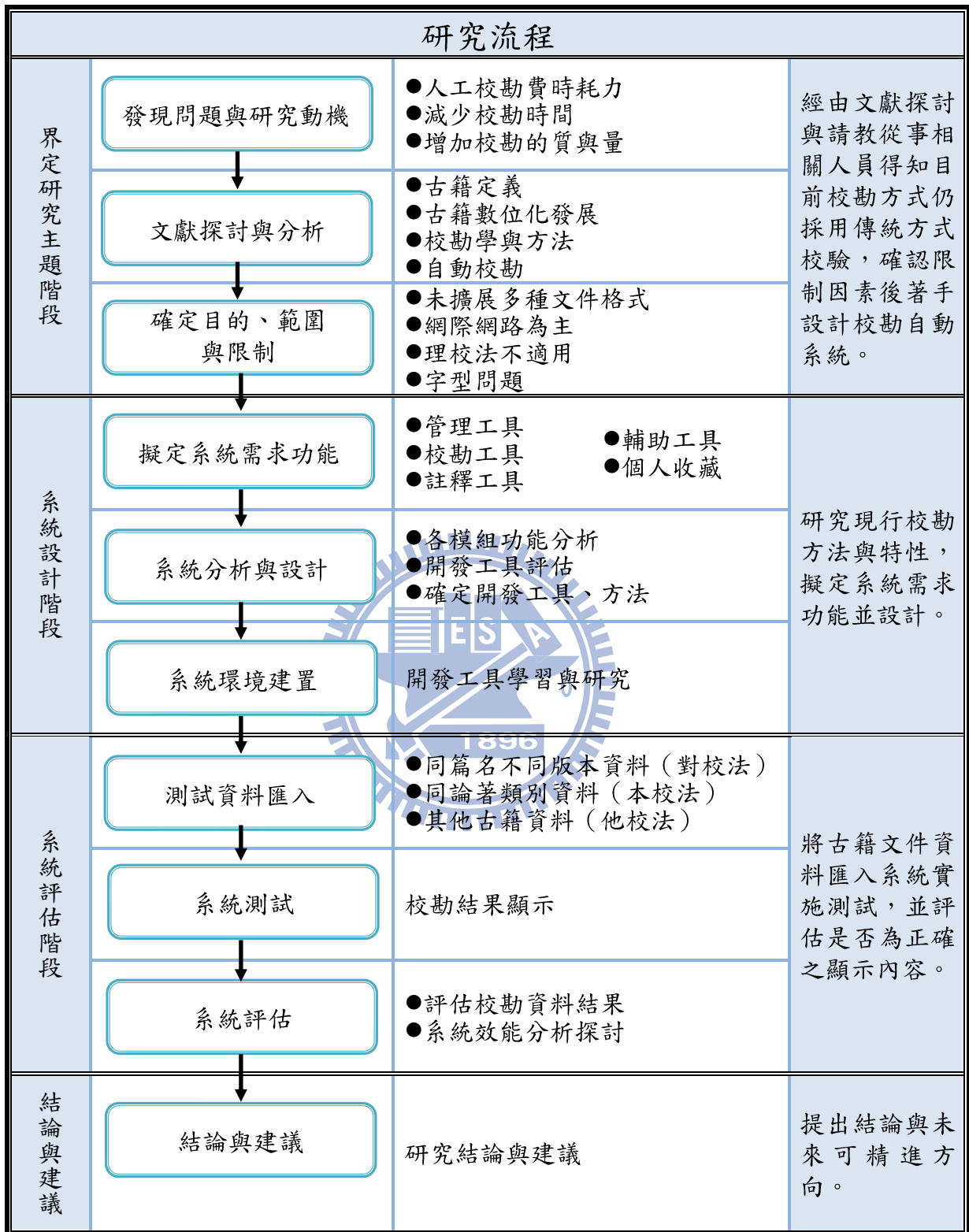


圖 1-1：研究流程圖

1.5 論文架構

本研究共分為五個章節，整理論文架構如下：

- 1.第一章為緒論，透過研究動機與目的，說明校勘研究者所遇到的問題並提供可行的解決方案，且訂定研究限制及研究流程。
- 2.第二章為文獻回顧與系統功能需求描述，介紹古籍及相關整理方法、校勘方法、自動校勘系統研究，並依據文獻予以改進設計，描述系統功能需求與所使用的工具程式。
- 3.第三章為系統分析與設計，描述系統架構、資料處理方法，並解釋各模組所處理的資料。
- 4.第四章為系統功能展示，說明系統各項功能是否符合需求，並實際將古籍文件資料輸入，評估顯示結果是否可解決校勘研究者的問題。
- 5.第五章為結論與建議，為本研究之總結，說明本研究之貢獻，並提出未來可繼續研究之議題。

第二章 文獻回顧與系統功能需求

本章內容分為三節，2.1 節就研究所需，說明相關文獻與運用，2.2 節為描述本系統擬定的功能需求，2.3 節為建置本系統所使用的程式工具。

2.1 文獻回顧

圖 2-1 為本研究文獻資料整理分佈圖，區分為古籍定義、古籍數位化發展、校勘與校讎學、校勘的方法、自動校勘相關研究文獻等五大部份。

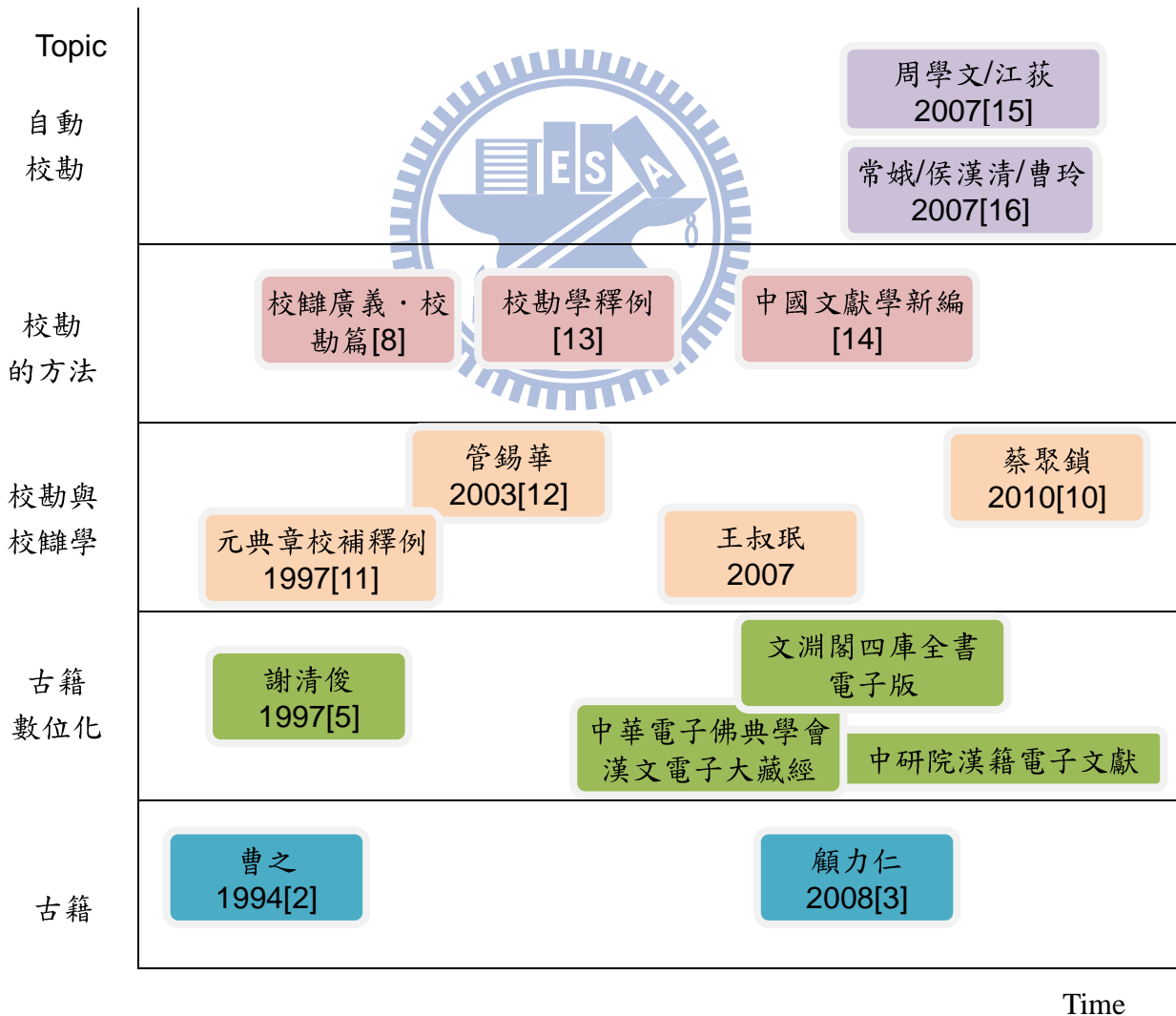


圖 2-1：本研究文獻資料整理

2.1.1 古籍定義

何謂古籍？即中國古代書籍的簡稱，以紙來抄寫或印刷的中國古代圖書，主要是指1911年以前出版的書籍（包含抄本、刻本），並採用中國古典裝幀形式（如包背裝、線裝、蝴蝶裝等），其內容主要是以研究中國古代傳統的思想、學術與文化的文獻¹；而曹之在《中國古籍版本學》一書中對「古籍」的範圍歸納出四種定義：凡是線裝書都叫古籍、凡是古人所著之書都叫古籍、凡是用古代漢語寫成的著作都叫古籍及1911年以前抄寫、印刷的圖書都叫古籍²，其中曹之採用狹義但得到大多數人認可的第四種說法。

而《中國文獻編目規則》中整合曹之歸納古籍定義的四種標準「裝訂形式」、「著作者時代」、「語言表達形式」及「成書年代」，將古籍下一個較為周延的定義為「係中國古代書籍的簡稱，主要指寫於1912年以前，並具有中國古典裝訂形式的漢字古籍³」。

另外李致忠認為，凡產生在1911年以前，內容是研究中國古代傳統文化、方法是中國古代傳統著作方式、裝幀具有中國古代圖書傳統裝幀形式的典籍，就是中國古籍⁴；茲將前述學者對古籍的定義整理如下表2-1。

¹ 互動百科，2010，古籍，「線上資料」，來源：<http://www.hudong.com/wiki/古籍#13> 2012年6月14日

² 曹之，1994，中國古籍版本學，頁3-5，台北，洪業文化事業有限公司。

³ 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁10，國立臺灣大學圖書資訊學研究所，台北。

⁴ 李致忠，2001，古籍版本知識500問，頁1，北京，北京圖書館出版社。

表 2-1：古籍的定義

年代	學者	古籍的定義
1994	曹之	1911 年以前抄寫、印刷的圖書都叫古籍。
2000	顧力仁	以《中國文獻編目規則》中對古籍的定義為「係中國古代書籍的簡稱，主要指寫於 1912 年以前，並具有中國古典裝訂形式的漢字古籍。
2001	李致忠	凡產生在 1911 年以前，內容是研究中國古代傳統文化、方法是中國古代傳統著作方式、裝幀具有中國古代圖書傳統裝幀形式的典籍。
2010	互動百科	中國古代書籍的簡稱，以紙來抄寫或印刷的中國古代圖書，主要是指 1911 年以前出版的書籍（包含抄本、刻本），並採用中國古典裝幀形式（如包背裝、線裝、蝴蝶裝等），其內容主要是以研究中國古代傳統的思想、學術與文化的文獻。

另外，錢存訓先生亦曾經在「歐美地區中國古籍存藏概況」一文中提到：「古籍比其他古物更為重要，因為歷史上的盛衰與興亡，思想上的起伏與變化，制度上的變革和延續，以至我們今天所沿襲的禮俗和傳統，名詞和習慣，大都是從古書的記載中得以流傳至今，因而影響到我們的日常生活⁵」。由此可知，古籍不僅具有歷史的意義與價值之外，亦匯聚了民族文化的精華，累積聖人先賢的智慧及經驗，後人讀書治學的來源，更是中華民族文化傳統禮俗的延續。

2.1.2 古籍內容數位化發展

為什麼要發展「古籍內容數位化」？在數位時代來臨時，為了便利儲存及方便閱讀與研究，遂有許多將古籍數位化之發展，然而，過去只是單純的將原稿掃描至資料庫中，提供影像化的閱讀，但是卻有以下的限制⁶：

⁵ 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁 iii，國立臺灣大學圖書資訊學研究所，台北。

⁶ 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁 58，國立臺灣大學圖書資訊學研究所，台北。

- 1.無法進行檢索、儲存及編輯。
- 2.若原稿的文字太小或是損壞嚴重，僅單純透過影像無法判讀細小的文字。

謝清俊教授曾在「中央研究院古籍全文資料庫的發展概要」一文中提到，因為古籍內容數位化在網路環境中有下列的優越性，不但不會在科技的洪流中微沒頂，反而是讓古籍重新展現最佳的風貌。分別是⁷：

- 1.古籍的電子版本可無限複製，而且幾乎無複製的成本。
- 2.電子古籍是取之不盡、用之不竭的資源，可供全體人民共享。
- 3.藉由四通八達的網路，電子古籍可以瞬息千里。
- 4.沒有運輸和分配的問題。
- 5.資料匯集後會產生新的訊息，並且經過相互參照，能發現前人所未見，對研究工作非常重要。
- 6.無損耗，不需保養，不怕遺失，儲存方便，體積小。
- 7.可以用計算機做快速檢索、應用與處理。

另外，大陸學者張普亦認為利用電腦從事「漢語語料資料庫」的研究，可以得到多重的效益，包括⁸：

- 1.進行各個歷史時期漢語研究、古漢語教學、自動翻譯、古文文本的自動分詞、自動校對、古籍版本的輔助校勘等研究領域。

⁷ 謝清俊、林晰，1997，中央研究院古籍全文資料庫的發展概要，「線上資料」，來源：http://cdp.sinica.edu.tw/paper/1997/19970301_2.html 2011年11月20日

⁸ 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁2，國立臺灣大學圖書資訊學研究所，台北。

2.古籍版本電子化與光碟的出版，有助於在網際網路上傳播，並且可以自動索引、標注和檢索。

所以，借重現代化電腦技術將古籍內容轉成為電子檔案的數位形式，並透過網際網路、硬碟等媒介來傳播與保存，使中國豐富的古籍資源不因天災或存管不當造成資料遺失，除了可以散布給有興趣的民眾使用與了解外，更能夠提供使用者快速查詢資料，以及增加學術研究的便利性。

而近年有許多學術單位、出版社及政府支持的機構依其特性及需求，透過數位化發展文獻典籍資料庫，以下針對中央研究院的「漢籍電子文獻一翰典全文檢索系統」、中華電子佛典協會的「漢文電子大藏經」、香港迪志文化出版有限公司的「文淵閣四庫全書電子版」等三個「全文資料庫」，分別介紹其概要及特色：

1. 台灣中央研究院「漢籍電子文獻一翰典全文檢索系統」⁹：

該院有鑑於古籍數位化的必要，於1984年至1990年期間，由歷史語言所及計算機中心合作開發「二十五史全文資料庫」，首開中文全文資料庫之先河，而後在此基礎上增加了《阮刻十三經注疏》、超過兩千萬字的台灣史料、一千萬字的大正藏以及其他典籍，合計字數一億三千四百萬字等項目，擴展成為「漢籍電子文獻一翰典全文檢索系統」（舊稱：漢籍全文資料庫）。更於1989年時提出了「數位典藏計畫」，將善本古籍、金

⁹ 中央研究院，1997，漢籍電子文獻介紹，「線上資料」，來源：<http://hanji.sinica.edu.tw/index.html> 2011年10月21日

石拓片、古籍附圖、台灣地方文獻及期刊報紙等，均納入古籍數位化之工作，將歷年以寫本、雕版、活字版、套版和鉛字印刷等方式所呈現之古代文籍，轉為電子媒體之形式，由於利用電子計算機之處理與檢索之功能，與網路無遠弗界之傳輸力量，使得古籍得以重新以新的面貌展現給廣大民眾，其特色包括¹⁰：

- (1) 品質最佳：文獻失誤率低於千分之一。
- (2) 規模最大：開發及完成的合計超過三億字，每年平均以三千萬字建檔，校對完成者約一千萬字。
- (3) 資料經過一定規模的組織。
- (4) 軟體設計最佳。



¹⁰ 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁 60，國立臺灣大學圖書資訊學研究所，台北。



圖2-2：漢籍電子文獻—翰典全文檢索系統

資料來源：<http://hanchi.ihp.sinica.edu.tw/>

2. 中華電子佛典協會「漢文電子大藏經」¹¹：

中華電子佛典協會於1998成立，其宗旨為建立出符合學術標準的數位化中文佛典。所使用的方法除了收集並校對目前各界已輸入的數位化佛典，並利用光學辨識來進行數位化。該協會選擇日本大正十三年至昭和九年(1924~1934)東京大正一切經刊行會所編輯出版「大正新修大藏經」來進行數位化。所以該工作並非將藏經佛典逐字鍵入，而是將紙本的佛經利用光學辨識方法轉換為數位的佛典，其特色包括¹²：

- (1) 提供多種版本，便利需求：版本包括普及版，即一般文字檔；App版，即「行末句點」格式之文字檔；HTMLhelp版，具有目錄、索引、全文檢索多功能；HTML版，可以直接使用網路瀏覽器閱讀。

¹¹ 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁60，國立臺灣大學圖書資訊學研究所，台北。

¹² 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁61，國立臺灣大學圖書資訊學研究所，台北。

- (2) 跨行檢索功能：使用跨行檢索及忽略行首資訊的中文化搜尋程式功能，以避免古籍文獻依紙本格式排列檢索，而碰到換行時，則無法完全檢索到一專門詞彙。
- (3) 版本比對校對方式，發揮校勘功能及提昇經文品質：在比對不同版本的藏經中，若發現「大正藏」的錯誤，則以紅色顯示經文。
- (4) 結合電子辭典閱讀藏經：該系統在 HTMLhelp 的瀏覽器上，將丁福保「佛學大辭典」電子檔字辭典與經文相連接。

海內外電子佛典的工作進行多年，目前以中華電子佛典協會所做的「大正藏」數位化工作最具代表性，也累積了許多寶貴的實務經驗。



圖 2-3：中華電子佛典學會「漢文電子大藏經」

資料來源：<http://www.cbeta.org/result/search.htm>

3. 香港迪志文化出版有限公司的「文淵閣四庫全書電子版」：

香港地區自1989年起，由香港中文大學中國文化研究所發展有關中國傳世典籍資料庫的研究工作，後來正式成立「古文獻資料庫中心」

(CHANT Center)，並先後推出《先秦兩漢古籍逐字索引叢刊》及《漢達古籍資料庫檢索系統》¹³。

大陸地區則自1980年代起開始了古籍數位化的工作，但是研究和成果比較少，直到1990年代中期，國家圖書館制訂了一個龐大的古籍特藏文獻數位化計畫，如碑碣精華、西夏碎金、敦煌遺珍、方誌以及甲骨文、永樂大典等，其中北京大學的《中國基本古籍資料庫》，計畫收錄古籍一萬種¹⁴，書同文數位化技術有限公司推出了《四部叢刊》全文檢索系統等相關系列的产品；迪志文化出版有限公司、上海人民出版社推出了文淵閣《四庫全書》全文檢索系統，共收圖書三千四百六十九多種，約七億字，三萬六千餘冊，該公司以掃瞄方式將原書圖像數位化，再用光學文字辨識技術，將圖像轉為近七億的電腦編碼文字，並研究出多種檢索功能。其特色包括¹⁵：

- (1) 圖文並列：「電子版」分為「原文及標題檢索版」及「原文及全文檢索版」，前者包含原書頁的圖像及標題檢索，後者另增加全文檢索功能。
- (2) 使用Unicode字集：於Unicode的基礎上，建立近三萬中文字集。
- (3) 跨平台：適用於臺灣(Big5)及大陸(GBK)中文視窗，以及其他如英文及日文等語文平台。

¹³ 郝淑東、張亮、馮睿，2007，古籍數字化的發展概述，頁 114-116，《情報探索》，2007 年第 7 期。

¹⁴ 郝淑東、張亮、馮睿，2007，古籍數字化的發展概述，頁 114-116，《情報探索》，2007 年第 7 期。

¹⁵ 顧力仁，2000，「資訊時代圖書館中國古籍組織與利用之探討」，博士論文，頁 62，國立臺灣大學圖書資訊學研究所，台北。



圖 2-4：文淵閣《四庫全書》全文檢索系統

資料來源：

http://www.content.sp.npu.edu.tw/teacher/pcwu/DocLib10/w3_full.ppt

綜合古籍內容數位化的應用與發展得知，不論是佛經、醫書、農書或經史子集等領域之古籍，皆為中華民族重要的知識資源，應該透過現代化科技的優勢，將深層的文化內涵予以展現，過去學者必須利用有限的時間和耗費精神去檢閱浩如煙海的古籍，如今藉由資訊科技的協助，不僅事半功倍，達到擴增古籍整理的層面，更提升古籍資源的運用。

2.1.3 校勘、校讎學

古籍數位化的過程面臨版本眾多及內容錯誤之問題，為了增加古籍之正確性及可信度，校勘的工作便顯得格外重要。

「校勘」是整理文獻的基礎工作，就是改正書面材料上由於種種原因而形成的字句篇章上的錯誤，使之恢復或接近本來面目¹⁶。校勘與校對不同，校對是改正文字的錯誤，確保典籍中文字的異同、是非，以求得符合文章原貌；校勘比校對多了個判斷分析的過程，古人也稱「校讎」。研究校勘的學問稱「校讎學」，是研究與總結各種校勘規律、內容的一門科學¹⁷。

西方與中國的校勘事業都有悠久的歷史。西方的校勘學源自古希臘人對《荷馬史詩》的校勘，最先校勘新約文字的應該是狄奧多斯(Theodotus)，在二世紀即用經文鑑別的方法以圖從新約「原典」中找到支持其教義的根據。後來，天主教武加大版本聖經的翻譯編撰者耶柔米(Jerome, St.347-420)也很注意經文的細節鑑別，充分注意傳抄本時所發生的各種錯誤。

1675年，校勘學的方法整理古典經文則從德國學者格利斯巴赫(Jakob Griesbach)、貝克(Immanuel Bekker)和拉赫曼(Karl Lachmann)開始的。到了二十世紀，西方主要通行的經文校勘學理論或方法有：地域傳統理論、理性鑑別法、抄本分類法、及推測性修正¹⁸。

¹⁶ 程千帆、徐有富，1998，校讎廣義·校勘篇，頁3，山東，齊魯書店。

¹⁷ 李勝文，2002，古籍版本與校勘，頁97，荷澤師專學報，第24卷1期。

¹⁸ 蔡聚鎖，2010，圖書情報學視野中的新約校勘法形成淺析，頁119，圖書館論壇，2010年第8期。

至於中國古籍的校勘學，胡適於《元典章校補釋例序》中有以下解釋：

「校勘之學起於文件傳寫的不易避免錯誤。文件越古，傳寫的次數越多，錯誤的機會也越多。校勘學的任務是要改正這些傳寫的錯誤，恢復一個文件的本來面目，或使他和原本相差最微¹⁹。

程千帆、徐有富《校讎廣義》曰：「除對書面材料發生錯誤的規律進行不斷的總結外，人們還對校勘的作用、方法、據以校勘的資料、校勘者應具備的條件、校勘成果的處理形式等，都進行過不斷地探索，從而在校勘實踐的基礎上逐步形成了校勘學²⁰」。書中亦提及了「校勘學」的相關內容，包括「校勘的資料」、「從事校勘所應具備的知識」、「校勘的方法」、「校勘成果的處理形式」…等。

管錫華在《漢語古籍校勘學》則對「校勘學」作了更詳細的解釋：「『校勘學』則是現代在歷代校勘實踐以及對這種實踐的理論總結的基礎上產生出來的一門科學。它祇以古籍的校勘為自己的研究對象，其目的和任務是總結歷代學者校勘古籍的法則和規律，以指導校勘的實踐。…而校勘學研究校勘則是深入的、系統的，他是一門獨立的科學²¹」。

也就是說校勘學的工作就是要改正古籍傳抄中所產生的錯誤，並且恢復文件本來的面貌。

¹⁹ 胡適，1997，《元典章校補釋例》序，頁1，《校勘學釋例》，1997年7月。

²⁰ 程千帆、徐有富，1998，校讎廣義·校勘篇，頁22，山東，齊魯書店。

²¹ 管錫華，2003，漢語古籍校勘學，頁7，成都，巴蜀書店。

2.1.4 校勘的方法

古籍之所以要校勘，是因為內容有錯誤。清末孫清謙在《劉向校勘纂微》中說：「書之待校讎者，為能訂其脫誤也。」古籍中的錯誤情況有四個方面：一誤、二脫、三衍（多文）、四亂。產生錯誤原因有兩個方面：

1. 無意之錯。這種錯數量較大，範圍較廣，又分四種情況：字形相近，字音相同或相近寫錯；竹、木簡上文字模糊，刊刻時刪去，又不加省字符號，因此出現文字脫落情況；正文、注文相混，將注文誤當正文，因此出現衍文情況；竹簡寫書免不了有線斷簡亂，刊刻時發生內容次序混亂的情況。
2. 有意之錯。主要兩方面原因：避諱。這是中國特有的現象，由於政治、社會等原因，避諱歷代都有，但嚴格程度不同。避諱的方法有三種：空字、同義字代替、缺筆。清代有文字獄。在修《四庫全書》時，凡書中「胡、虜、戎、狄」等字都要改為相應的字。書商為了謀利，故意把一些小說、戲劇中的字加以篡改，投世俗之所好²²。

面臨古籍內容中的錯誤時，要如何進行校勘工作，在清代校勘家之間有過一段爭論，直至1931年陳恒著作《元典章校補釋例》一書（後改名為《校勘學釋例》），使校勘方法更具實踐性，也反映了歷史的進步。

²² 李勝文，2002，古籍版本與校勘，頁 97，荷澤師專學報，第 24 卷 1 期。

下列為參考陳垣《校勘學釋例》書中所論「校法四例」，並根據洪湛侯《中國文獻學新編》分別為「對校法」、「本校法」、「他校法」、「理校法」及「校法的綜合運用」等五個部份提出說明。

1.對校法

「對校法」是校勘的基本方法，可據以了解祖本或別本之本來面目，此法最為簡便、穩當，純屬機械法。陳垣論此法曰：即以同書之祖本或別本對讀，遇不同之處，則注于其旁。劉向《別錄》所謂「一人持本，一人讀書，若怨家相對者」，即此法也²³。

2.本校法

「本校法」即以本書校本書，根據本書的上下文意、相同或相近的句法、詞例來校勘本書的錯誤。陳垣論此法曰：「本校法」者，以本書前後互證，而抉摘其異同，則知其中之謬誤²⁴。

3.他校法

「他校法」即參考其他相關書籍校勘，陳垣曰：「他校法」者，以他書校本書。凡其書有采自前人者，以前人之書校之；有為後人所引用者，以後人之書校之，其史料有為同時之書所並載者，可以同時之書校之²⁵。

²³ 陳垣，1997，校勘學釋例，頁 118，上海，上海書店出版社。

²⁴ 陳垣，1997，校勘學釋例，頁 119，上海，上海書店出版社。

²⁵ 陳垣，1997，校勘學釋例，頁 120，上海，上海書店出版社。

4.理校法

「理校法」是運用分析、歸納、類推等考證手段，據理推斷古書譌誤的校勘法。另外清代學者段玉裁針對「理校法」亦提出：「校書之難，非照本改字不訛不漏之難，定其是非之難。遇無古本可據，或數本互異，而無所適從之時，則須用此法。此法須通識為之，否則鹵莽滅裂，以不誤為誤，而糾紛愈甚矣。故最高妙者此法，最危險者亦此法²⁶」。

5.校法的綜合運用

若能綜合運用上述的數種校勘法，將能全面性運用校勘資料，得出更令人信服的結論。洪湛侯曰：在實踐中往往幾種校勘法結合使用，效果更好，因為單用一種校勘法，有時不一定能作出結論，即使作出結論，也不一定可靠。遇到比較複雜的情況，就必須與其他校勘方法，參互運用，才能確切地判斷出異文的正誤與是非²⁷。

前述「理校法」因須透過校勘者長年累積之經驗，加以推理、判斷始可改正，屬讀書心法，電腦技術尚無法完全取代，故本研究採用陳垣「校法四例」中「對校法」、「本校法」及「他校法」三種校勘方法做為系統「校勘模組」設計基礎，並於第三章說明。

²⁶ 陳垣，1997，校勘學釋例，頁 121，上海，上海書店出版社。

²⁷ 洪湛侯，1994，中國文獻學新編，頁 179，浙江，杭州大學出版社。

2.1.5 古籍自動校勘

當古籍文件資料繁多，各版本在實施校勘時，人工作業難免出錯或疏漏，而古籍自動校勘是指透過特定撰寫的程式，利用電腦快速運算、記憶的特點來發現不同古籍版本之間文字或詞句的差異，並利用系統所設計的工具將差異處實施標記，提供研究者校勘，藉由此自動校勘的概念可以達到下列目標²⁸：

1. 節省時間，提高效率。
2. 減少誤失，提高準確率。
3. 可將校勘的段落實施排序。
4. 校勘資料可實施全面性的判斷。

而自動校勘的對象的類別²⁹，分述如下：

1. 錯文：意指文字出現錯誤，或因時代背景不同產生的文字差異，如古今字、異體字或避諱字。
2. 脫文：文件內容出現漏字或漏句的現象。
3. 衍文：文件內容比原文多出文字的現象，產生原因包括刻意增加，豐富閱讀性，或無意過失的增加。

依據上述文獻探討與校勘的概念，參考「自動校勘的核心計算方法」流程如圖 2-5，另外古籍自動校勘系統模組設計，如圖 2-6，雖為單機操作版本，

²⁸ 周學文、江荻，2007，元朝秘史的計算機自動校勘方法，頁 142，語言文字應用，2007 年第 3 期。

²⁹ 常娥、侯漢清、曹玲，2007，古籍自動校勘的研究和實現，頁 84，中文信息學報，第 21 卷第 2 期。

但可做為本研究系統開發的基礎方向，並拓展至網際網路運用。

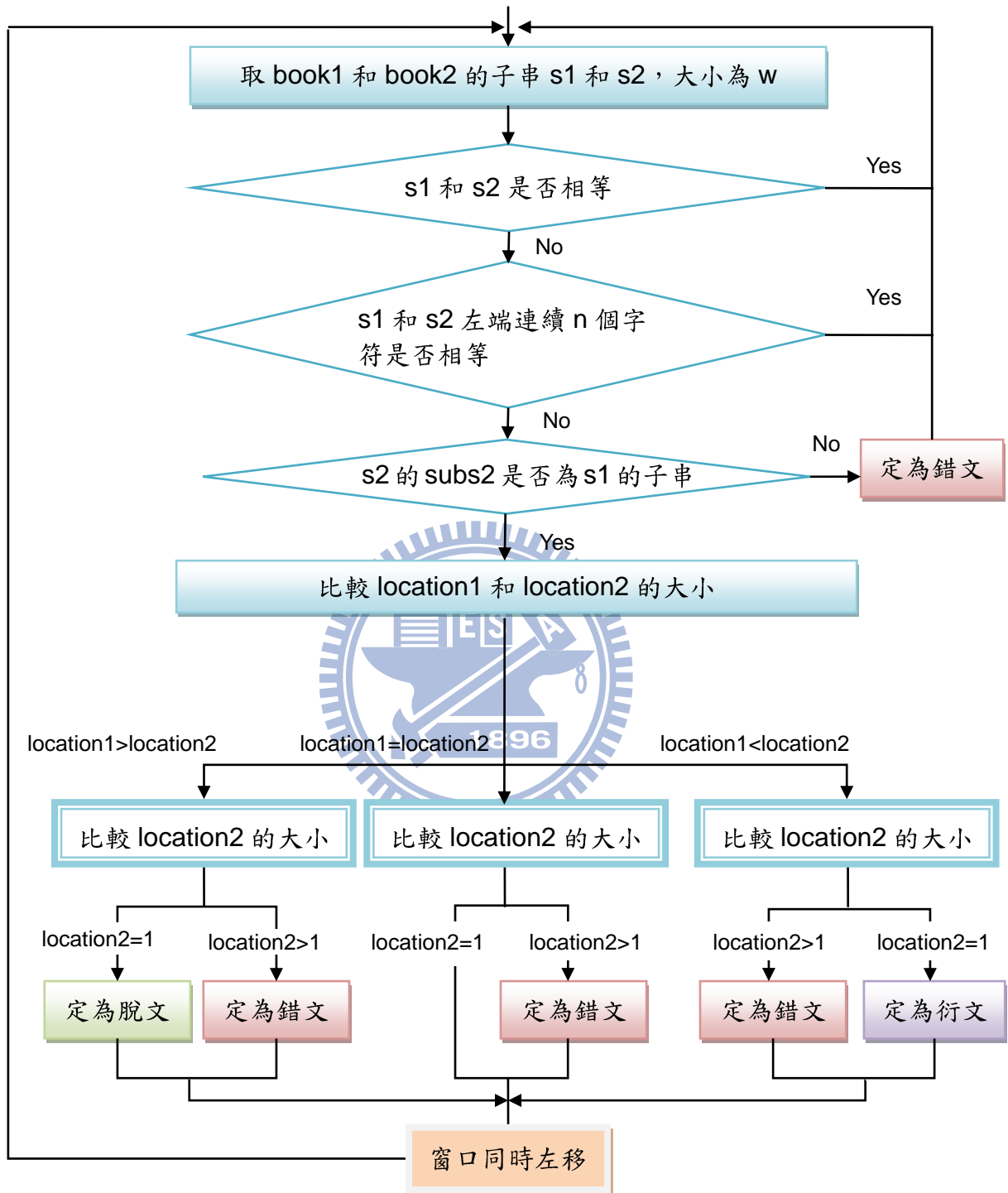


圖 2-5：自動校勘核心計算方法流程圖

資料來源：古籍自動校勘的研究和實現，常娥等，2007

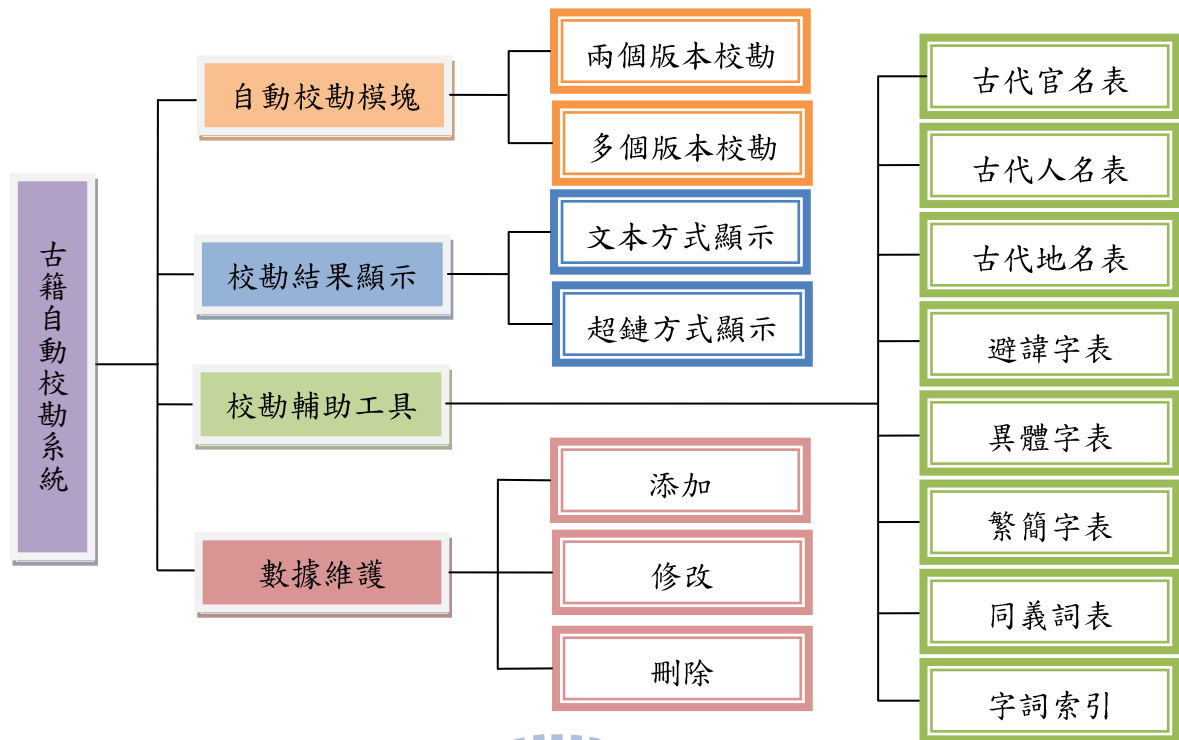


圖 2-6：古籍自動校勘系統模組圖

資料來源：古籍自動校勘的研究和實現，常娥等，2007

2.2 系統功能說明

在本系統的規範應用中，每一篇古籍包含了內容註釋及校勘結果，其中每一個內容註釋代表一個關鍵字詞，使用系統檢索時主要是針對校勘及註釋的資料，而系統主要的工作有：

1. 將文字資料轉換為網頁式處理的資料。
2. 將系統中的資料進行分類檢索。
3. 提供自動校勘的功能，校驗不同版本間文字的差異及參考文獻。

由於使用者有不同的需求層級，系統提供不同需求使用者權限的功能介

面，除了可讓使用者便於運用外，亦可避免使用者需求混淆而造成系統錯誤：

1. 系統管理者：維持系統的運作，及修正網頁或資料庫產生的錯誤。
2. 校勘研究者：可以匯入古籍文件資料，並可運用系統中的「校勘工具」（「對校法」、「本校法」與「他校法」）建立校勘資訊，以及使用「註釋工具」針對文件內容及差異文字實施註釋。
3. 一般使用者：無將古籍文件資料匯入系統的權限，僅能使用系統中的「校勘工具」、「註釋工具」及「個人收藏」等功能，主要分享及提供個人建議，並顯示於系統中給予研究者參考。

2.3 使用的程式工具

本研究建置之古籍校勘、閱讀平台所使用的程式工具計有 AppServ、PHP 及 MySQL，分別於本節介紹。

2.3.1 AppServ

AppServ 整合了 Apache、PHP、MySQL、phpMyAdmin 的伺服器安裝程式，除了能夠快速的在 Windows 或 Linux 環境中安裝之外，只要執行單次的安裝動作，Apache 便會自動完成相關的設定，無須再個別調整設定，另外，由於在網路上取得容易，而且是免費版本，更適合用來建置測試環境。

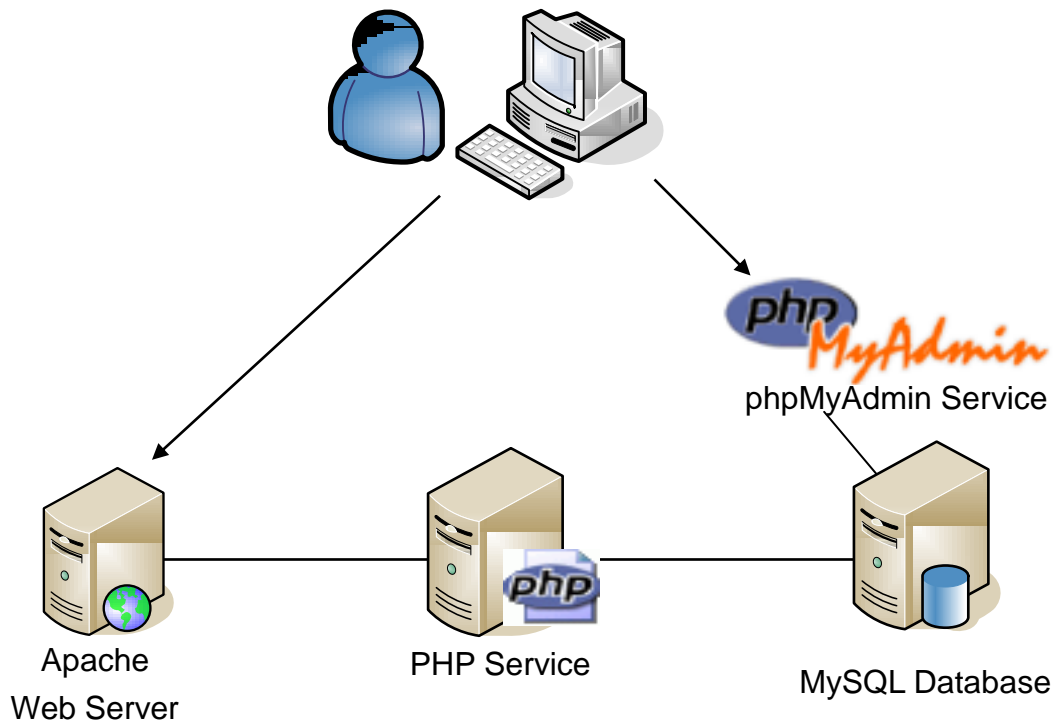


圖 2-7：Apache Service、PHP、MySQL、User 使用關係圖

AppServ 中的軟體與之間的關係如圖 2-7，簡要說明如下：

1. Apache HTTP Server

主要負責將網頁傳送給 PHP 處理，與網站相關的設定，如置放網站程式的主目錄，即是在 Apache HTTP Server 伺服器軟體中設定。

2. PHP

伺服器程式語言，用來解析伺服器軟體所傳送的程式碼，解析之後所產生的 HTML 網頁傳到瀏覽器，如 IE、FireFox、Chrome 等，再由瀏覽器解析 HTML 網頁，最後才是使用者所看到的網頁畫面。有關 PHP 程式碼語言在 2.3.2 一節中將更詳細的介紹。

Source of: /downloads.php

```
<?php // vim: et
// $Id$
$_SERVER['BASE_PAGE'] = 'downloads.php';
include_once $_SERVER['DOCUMENT_ROOT'] . '/include/prepend.inc';
include_once $_SERVER['DOCUMENT_ROOT'] . '/include/version.inc';

// Try to make this page non-cached
header_nocache();

$SIDEBAR_DATA = '
<h3>Binaries for other systems</h3>
<p>
We do not distribute UNIX/Linux binaries. Most Linux
distributions come with PHP these days, so if you do
not want to compile your own, go to your distribution\'s
download site. Binaries available on external servers:
'
```

圖 2-8：PHP 程式碼片段

資料來源：<http://www.php.net/source.php?url=/downloads.php>

3. MySQL

關聯式資料庫系統的伺服器。主要負責實際儲存網頁需要用到的資料，例如購物網站所需要儲存商品、客戶資訊以及交易的記錄等。並且透過 SQL 指令查詢或異動資料庫的內容，其特色將於 2.3.5 一節中詳加介紹。

```
Telnet - 192.168.0.200
Connect Edit Terminal Help

mysql> desc student;
+-----+-----+-----+-----+-----+-----+
| Field          | Type                | Null | Key | Default        | Extra          |
+-----+-----+-----+-----+-----+-----+
| roll_no        | int(10) unsigned    |      | PRI | NULL           | auto_increment|
| name           | varchar(20)         |      |     |                |               |
| specialization | varchar(6)          |      |     |                |               |
| dob            | date                |      |     | 0000-00-00    |               |
+-----+-----+-----+-----+-----+-----+
4 rows in set (0.00 sec)

mysql> show columns from student;
+-----+-----+-----+-----+-----+-----+
| Field          | Type                | Null | Key | Default        | Extra          |
+-----+-----+-----+-----+-----+-----+
| roll_no        | int(10) unsigned    |      | PRI | NULL           | auto_increment|
| name           | varchar(20)         |      |     |                |               |
| specialization | varchar(6)          |      |     |                |               |
| dob            | date                |      |     | 0000-00-00    |               |
+-----+-----+-----+-----+-----+-----+
4 rows in set (0.00 sec)

mysql> _
```

圖 2-9：MySQL 命令模式介面

資料來源：<http://www.tutorialsworld.com/sql/working-with-mysql.htm>

4. phpMyAdmin

由於原來的 MySQL 為指令操作模式的介面，對於習慣於 Web 介面的使用者而言，在管理及操作上不方便，故用 PHP 語言寫成 Web 操作介面的 MySQL 資料庫管理工具，除了可以方便建立、修改、刪除資料庫及資料表，更能大量處理資料的匯入與匯出。

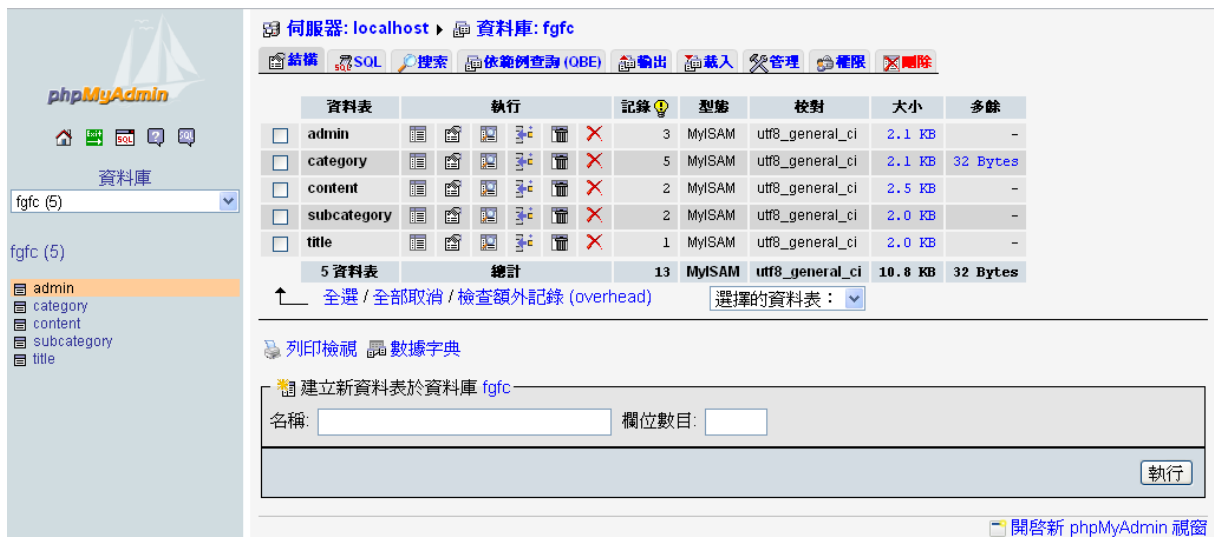


圖 2-10：phpMyAdmin 管理介面

資料來源：中國古籍文件校勘系統資料庫

2.3.2 PHP

PHP(PHP：Hypertext Preprocessor)是一種跨平台、嵌入式的伺服器端執行的描述語言，主要用途是在於處理動態網頁，也包含了命令列執行介面(command line interface)，或者產生圖形使用者介面(GUI)程式。最早是由拉斯姆斯·勒多夫在 1995 年發明，現在 PHP 的標準則由 PHP Group 和開放原始碼社群維護。

PHP 的應用範圍相當廣泛，大多執行在網頁伺服器上，透過執行 PHP 程式碼來產生使用者瀏覽的網頁，並且可以在多數的伺服器和作業系統上執行，而且使用 PHP 完全是免費的。根據 2007 年 4 月的統計資料，PHP 已經被安裝在超過二千萬個網站和一百萬台伺服器上^{30 31}。

³⁰ 維基百科，2011，PHP，「線上資料」，來源：<http://zh.wikipedia.org/wiki/PHP> 2011 年 10 月 30 日

³¹ PHP Usage Stats，「線上資料」，來源：<http://www.php.net/usage.php> 2011 年 10 月 30 日

Usage Stats for April 2007

PHP: 20,917,850 domains, 1,224,183 IP addresses

Source: Netcraft

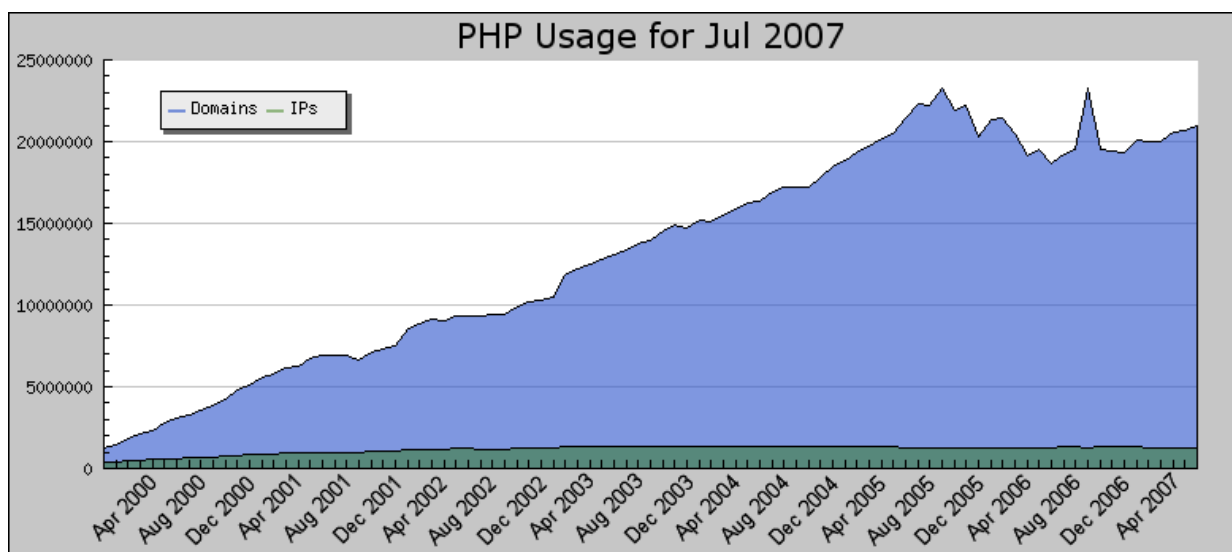


圖 2-11：PHP 使用趨勢狀況

資料來源：<http://www.php.net/usage.php>

由於網路軟體及硬體的快速發展，網路上的各種應用也已從單向的服務發展到互動式的服務，也就是資訊由伺服器單向提供給使用者，發展到使用者可以利用 CGI(Common Gateway Interface)來達到和伺服器互動的效果。為了要達到這種互動效果許多技術便發展出來，PHP 即是其中的一種 CGI 程式與 Web 伺服器的互動介面，根據此介面所撰寫的程式稱為 CGI 程式；CGI 程式必須在伺服器上執行。可以撰寫 CGI 的語言相當多，例如 C、Perl、ASP、Java、PHP 等，在網路上運作的關係如下³²：

³²凱文瑞克，2006，PHP5 與 MySQL5 入門學習指南，頁 1-2，台北，旗標出版。

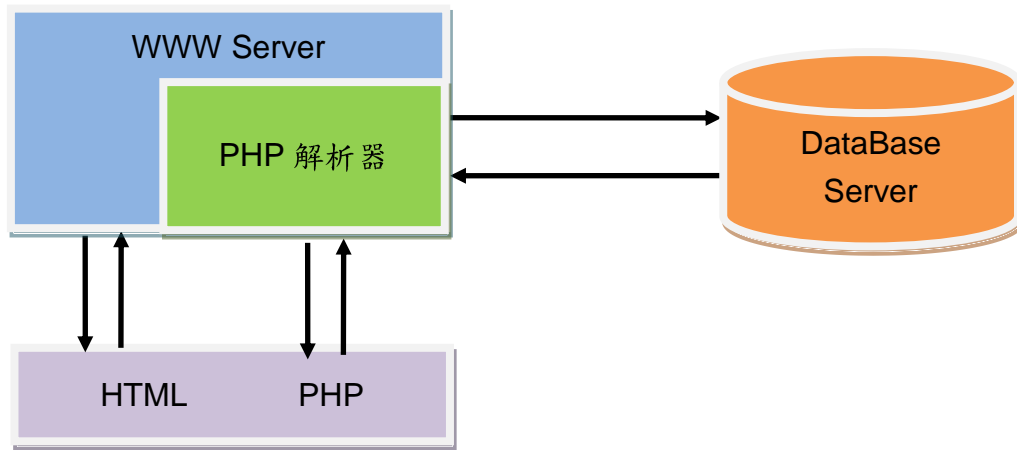


圖 2-12：CGI 關係圖

資料來源：PHP 5 與 MySQL 5 入門學習指南

使用者藉瀏覽器送出相關的訊息到伺服器，由伺服器上的 CGI 程式來負責解析使用者的要求再予以適當的回應，這些回應再送回給使用者。當然在必要的時候 CGI 程式必須與伺服器上的資料庫進行溝通，送出使用者所需要的資料。

另外，PHP 具有下列幾項特色³³：

1. 支援物件導向：在 PHP5 版本以後已能完整支援物件導向功能，更加強 XML、MySQL extension 等資料的存取。
2. 跨平台：PHP 可以在 UNIX 及 Windows 系列的作業系統上執行，並不會產生作業系統的相容性問題。
3. 嵌入式：PHP 是嵌入在 HTML 中執行的語言，具有高度的隱藏性。一般純 HTML 文件或 JavaScript 檔使用者可以利用瀏覽器檢視原始內容的功

³³ 凱文瑞克，2006，PHP5 與 MySQL5 入門學習指南，頁 1-3~1-5，台北，旗標出版。

能，就可以看到原始文件檔的內容，但是經過 PHP 處理後的文件檔，使用相同的方式並無法看到原始文件檔內容。

4. 開放原始碼：PHP 是 Open Source，也就是在 GPL(General Public License)下使用者可以任意的使用並不會受到限制。

5. 資料庫：PHP 支援的資料庫相當多，使得使用者可以很輕易的開發出網路資料庫的應用程式，支援的資料庫如下：

表 2-2：PHP 資料庫支援一覽表

Adabas D	Ingres	Oracle	Dbase
InterBase	Ovrimos	Empress	FrontBase
PostgreSQL	FilePro	mSQL	Solid
Hyperwave	Direct Ms-SQL	Sybase	IBM DB2
MySQL	Velocis	Informix	ODBC
Unix	dbm		

6. 通訊協定：PHP 本身提供了相當多的通訊協定服務，例如 IMAP、SNMP、NNTP、POP3 等，有了這些通訊協定的支援就可以開發相關的應用程式，如網頁式收發電子郵件。

7. 輸出模式：PHP 除了可以輸出 HTML 以外還可以輸出圖像、PDF、XHTML 和 XML 等。

8. 操作模式：PHP 提供三種操作模式。

(1) 伺服器端(Server-side)描述方式。這種方式必須要有 PHP 解析器，Web

伺服器及 Web 瀏覽器三種軟體才可以工作。使用者可以藉由 Web 瀏覽器經 Web 伺服器瀏覽含有 PHP 程式的網頁。

(2)命令操作模式。在此操作模式下僅需要 PHP 解析器即可。主要目的是讓 PHP 可以在作業系統下直接執行，就像是執行一個作業系統命令一樣。

(3)使用者端的應用程式。PHP 雖然是一個簡單的描述語言，但是 PHP 也可以設計 GUI 的應用程式。

2.3.3 MySQL

MySQL 是一個開放原始碼的關聯式資料庫管理系統(RDBMS)，可以用來管理並處理來自使用者資料的眾多需求，原開發者為瑞典的 MySQL AB 公司，由於效能高、成本低、可靠性好，已經成為流行的資料庫，並且被廣泛地應用在 Internet 上的中小型網站中。隨著 MySQL 的不斷成熟，它也逐漸用於更多大規模網站和應用，比如維基百科、Google 和 Facebook 等網站³⁴。所謂資料庫係指眾多可隨時取用的資料，以檔案形式，很有組織及效率的存放在電腦上，而這些資料以資料庫(database)－資料表(table)－資料錄(record)－欄位(column)這種階層式架構存放³⁵，如下圖：

³⁴ 維基百科，2011，MySQL，「線上資料」，來源：<http://zh.wikipedia.org/wiki/MySQL> 2011 年 10 月 30 日

³⁵ Mark Maslakowski 著，陳清豪、袁琿譯，2001，自學 MySQL 21 天課程，頁 4-5，台北，博碩文化。

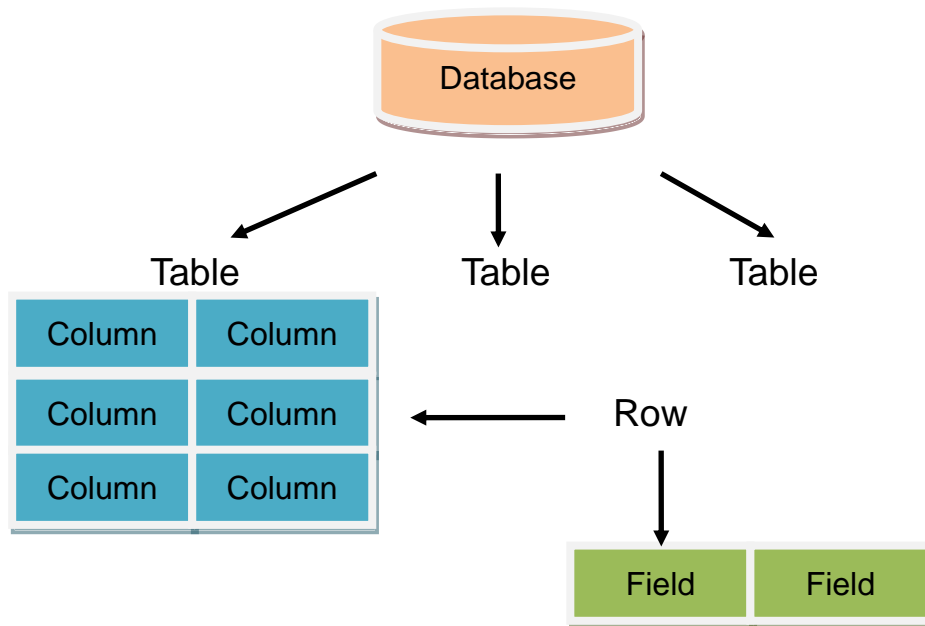


圖 2-13：資料庫階層式架構圖

資料來源：自學 MySQL 21 天課程

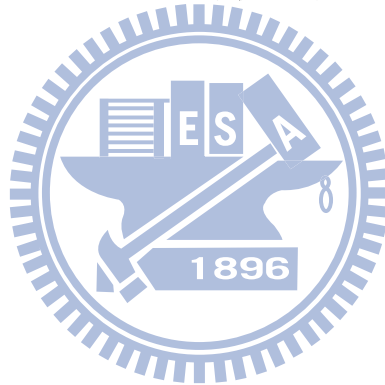
而 MySQL 的主要特色有^{36 37}：

1. 以 C 和 C++ 撰寫並使用了多種編譯器進行測試，保證原始碼的可移植性。
2. 可在 FreeBSD、HP-UX、Linux、Mac OS、Novell NetWare、NetBSD、OpenBSD、OS/2 Wrap、Solaris、Windows 等多種作業系統平台上執行。
3. 為提高可攜性，使用 GNU Automake、Autoconf 及 Libtool。
4. 提供 C、C++、C#、VB.NET、Delphi、Java、Perl、PHP、Python、Ruby 的 API。

³⁶ MySQLPress 著，范綱志譯，1995，MySQL 管理者手札-跟 MySQL AB 徹底研究 MySQL，頁 1-5，台北，上奇科技。

³⁷ 維基百科，2011，MySQL，「線上資料」，來源：<http://zh.wikipedia.org/wiki/MySQL> 2011 年 10 月 30 日

- 5.透過核心執行緒達到完全多重執行緒功能，若有多顆 CPU，輕易便能得到最高的效能。
- 6.優化的 SQL 查詢演算法，有效地提高查詢速度。
- 7.提供多語言支援，常見的編碼如中文的 GB 2312、BIG5，日文的 Shift JIS 等都可以用作數據表名和數據列名。
- 8.使用具索引壓縮的快速二元樹資料表(MyLSAM)。
- 9.記憶體內的雜湊(hash)資料表，可作為暫存資料表。
- 10.提供 TCP/IP、ODBC 和 JDBC 等多種資料庫連接途徑。



第三章 系統分析與設計

3.1 系統設計概念

本系統所設計的概念是基於協助校勘學者的理念，以古籍文件為基礎，讓使用者可以將校勘的文件比較出差異後加以註解，並可分享校勘的結果，減少古籍文件誤失，除了成為較具可信度的善本³⁸之外，亦可節省校勘時間，提高校勘的質與量。

此外，也可以透過此系統，讓古籍文件校勘後的結果與其他文件產生類似詞句的關聯，讓研究者可以透過點選的方式，短時間內找出文件或文件間的差異。

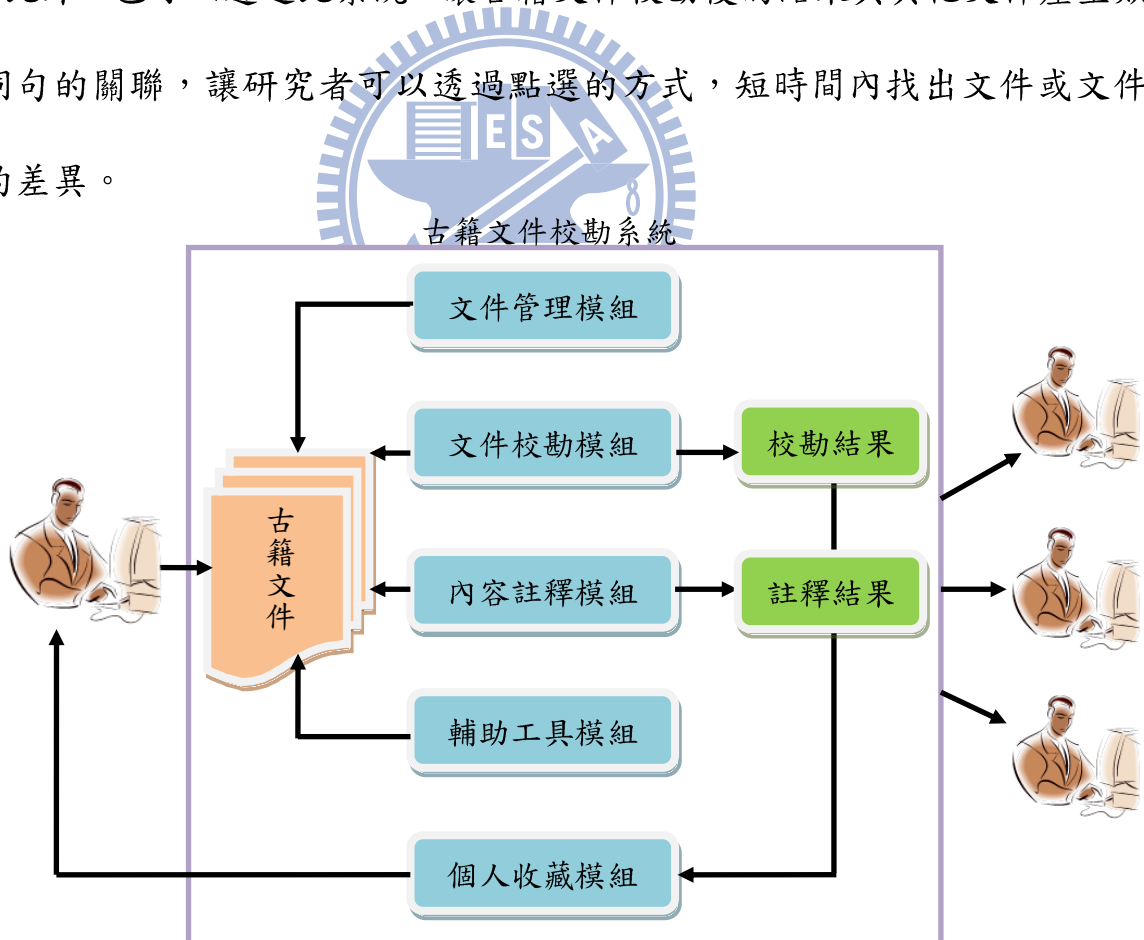


圖 3-1：系統設計概念圖

³⁸ 善本是指校勘嚴密，刻印精美的古籍。

3.2 設計架構

本系統於建構之初因考量到效率問題，基本上採用三層式（檢視層、商業邏輯層、資料存取層）架構來建立，並使用MVC設計模式，將模型(Model)，檢視(View)及邏輯控制器(Controller)等三種元件應用於三層式架構中，如圖3-2，除了讓使用者容易執行新增、修改和刪除等動作，而且使用不同的View來顯示一樣的資料外，亦可方便程式設計人員的撰寫，其設計模式分述如後：

1. Model元件

主要負責資料庫進行存取與管理的動作，以及程式演算法實作，當使用者提出資料變更時，通知View元件展示這些資料。

2. View元件

介面設計人員進行圖形化的設計，主要負責使用者所看到的畫面(User Interface)，以及展示Model元件的資料給使用者，讓使用者得以與系統進行互動，並通知Controller的動作。

3. Controller元件

主要負責使用者請求(Request)的處理，並且協調Model元件與View元件之間的動作，依照商業邏輯的需求指示Model動作。

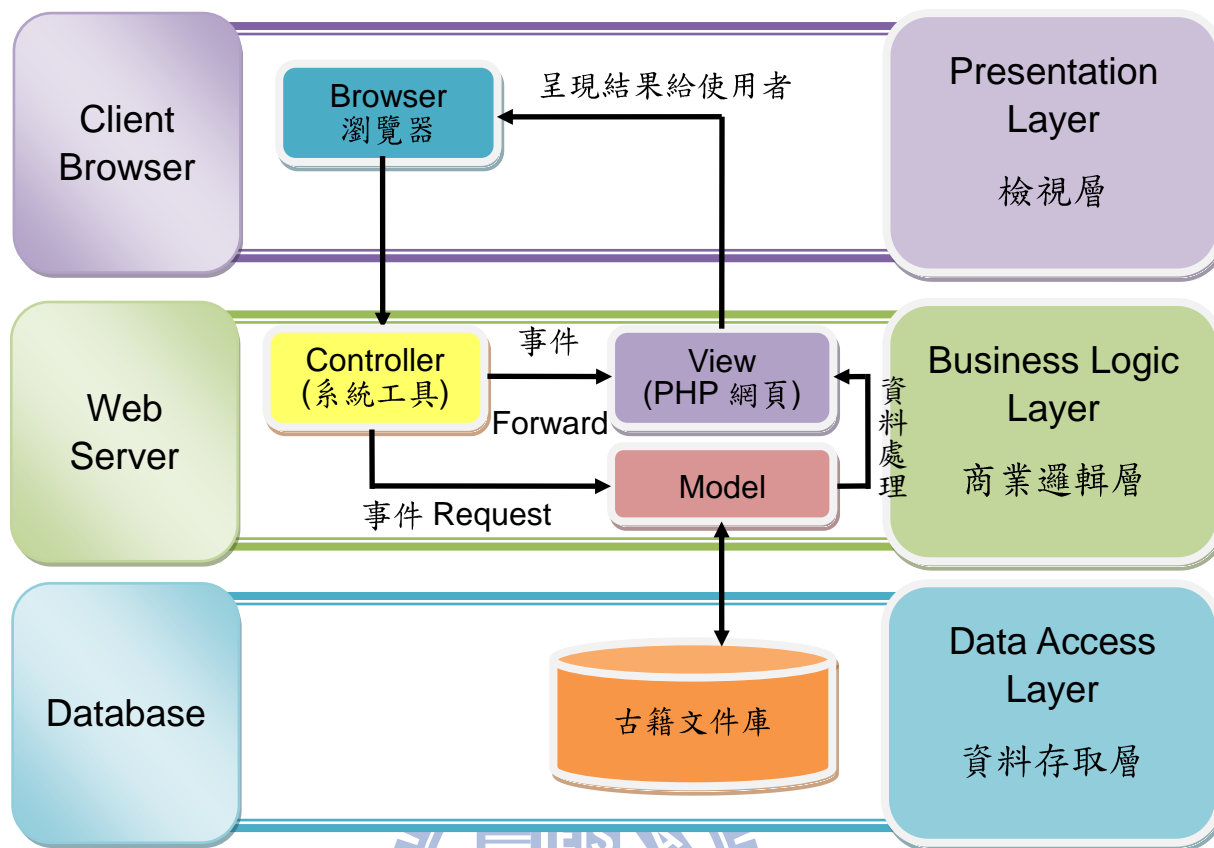


圖 3-2：MVC 系統設計架構圖

3.3 系統模組設計

系統設計主要分為五個模組，圖3-3為架構圖，各模組功能分述如下：

1. 古籍文件管理模組：使用者可透過此模組對文件進行新增、刪除、編輯等動作，其文件的類型為文字檔，將檔案上傳至系統後，轉換為網頁(HTML)格式，並提供使用者相關操作功能。
2. 古籍文件校勘模組：使用者可透過此模組將特定文件執行版本校勘，系統共提供了對校法、本校法、他校法的選擇，並藉由註釋及輔助工具模組，將校勘結果實施註釋。

- 3.古籍內容註釋模組：使用者可以透過此模組，將所選取的文字段落及校勘結果進行註釋動作。
- 4.輔助工具模組：使用者可透過此模組建立字詞庫，如同義字庫、避諱字庫，以便於文字校勘作業。
- 5.個人收藏模組：使用者透過此模組將有興趣之文件做個人化收藏、分享與管理。

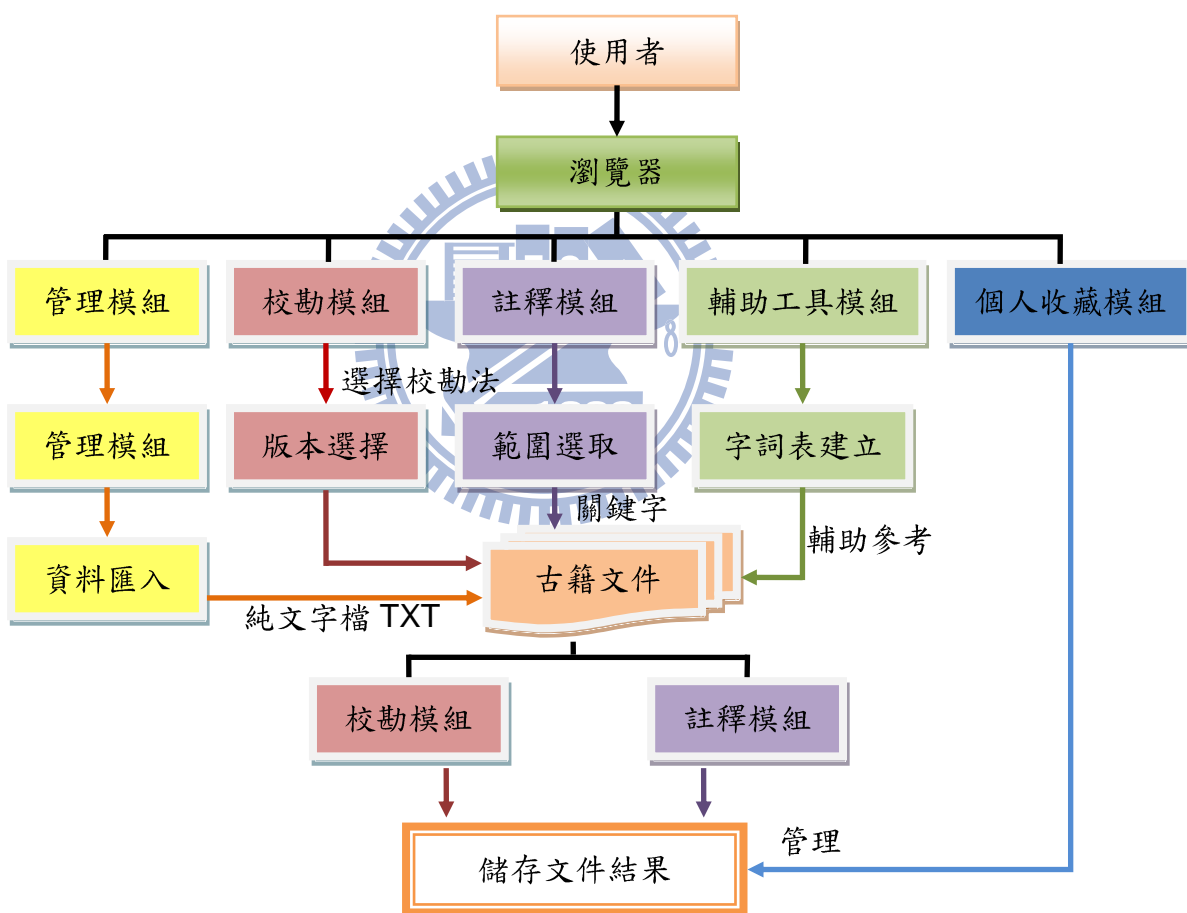


圖 3-3：系統模組設計架構圖

3.3.1 古籍文件管理模組

此模組計有下列三個功能，流程如圖3-4：

- 1.文件管理：共區分為古籍類別、古籍列表與古籍篇名，主要顯示系統中各文件的分類，並可分別針對文件名稱實施編輯與修改。
- 2.文件新增：將古籍依類別（如：儒家、道家）、列表（如：倫語、孟子）與篇名（如：學而、為政）建立。
- 3.文件上傳：使用者上傳純文字(.txt)，透過系統轉換為網頁檔(HTML)之格式後存入資料庫中，並顯示於古籍瀏覽樹狀圖，便可依類別瀏覽文件內容。



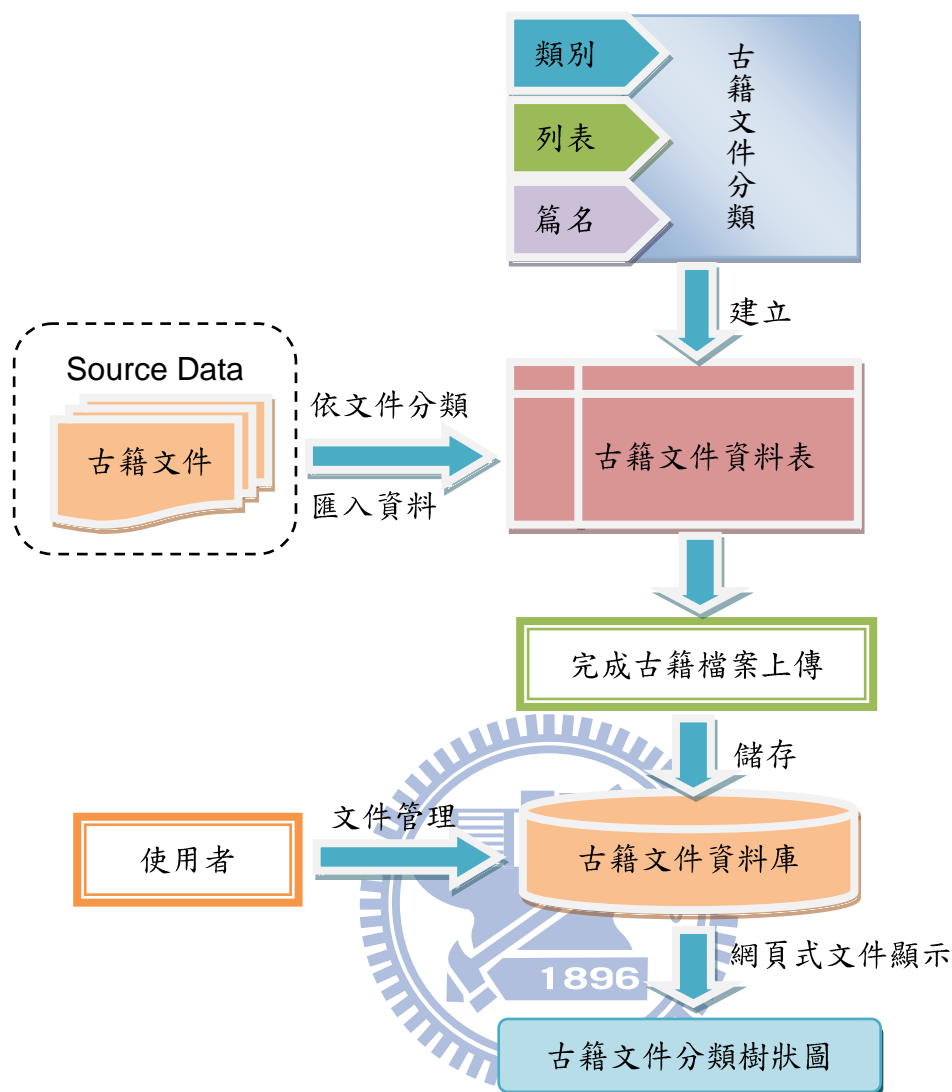


圖 3-4：古籍文件管理模組流程圖

3.3.2 古籍文件校勘模組

此模組共有「校勘方式選擇」及「校勘收藏」二種功能，流程如圖3-5，分述如下：

- 1.校勘方式選擇：系統提供了「對校法」、「本校法」與「他校法」三種校勘方法，使用者依實需點選採用的方式後，再選擇系統中列表、篇名顯示之古籍文件實施校勘。

另外，系統校勘方式「本校法、他校法」的演算法是以段落為基礎，將底本與校本逐次比對出相似段落後記錄，顯示校勘結果，而「對校法」則利用此相似段落取代之為特殊符號後，執行字串的切割，將底本與校本再執行比對，顯示相異的段落，如圖3-6。

2.校勘收藏：將校勘顯示結果加以註釋後，納入個人收藏或者實施分享，提供平台上的使用者參考。

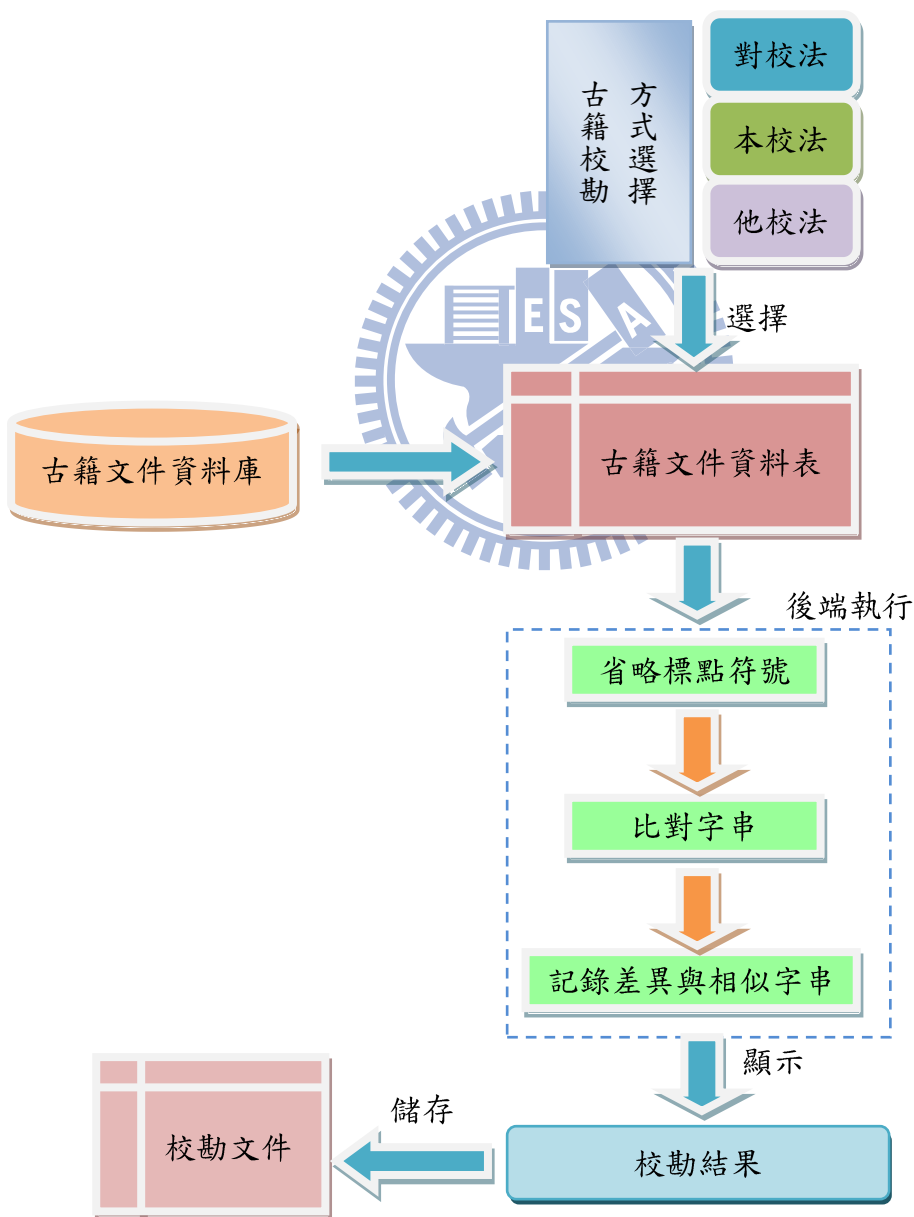


圖 3-5：古籍文件校勘模組（對校法、本校法、他校法）流程圖

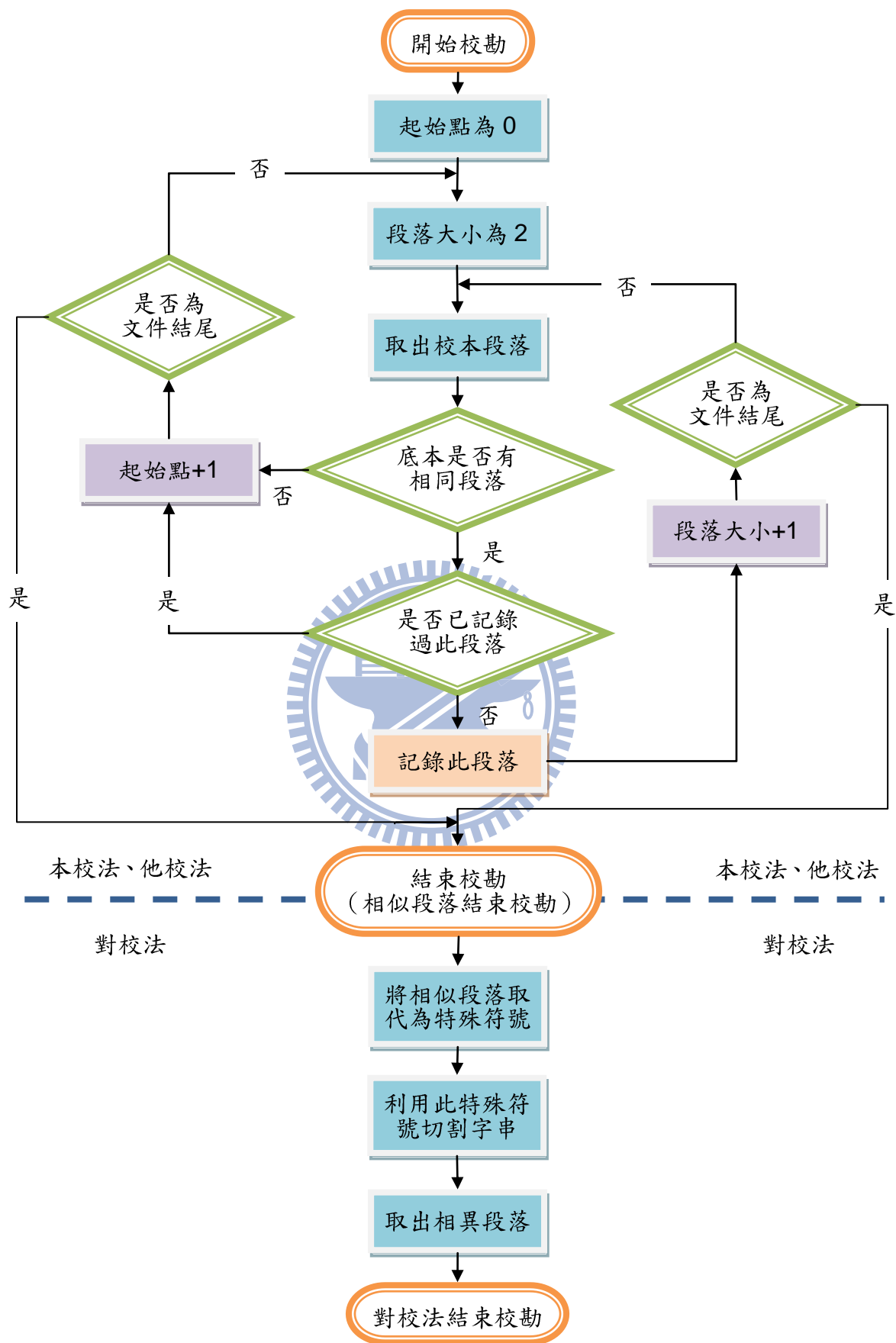


圖 3-6：對校法、本校法、他校法流程圖

3.3.3 古籍內容註釋模組

此模組可針對古籍文件內容執行註釋的動作；將系統之古籍文件中的段落或文字反白後，可自行加入註釋，或於文件校勘後所產生的差異處，執行註釋的動作，如圖3-7。

另外，文字或段落定位的演算法是以Javascript中TextRange函數撰寫，以TextRange中的Move方法模擬使用者的動作，如圖3-8³⁹：

1. 選取文字或段落範圍：字元(character)、字詞(word)、段落(sentence)。
2. 設置起始點與結束點。
3. 對選取的範圍執行動作。



³⁹林易賢，2011，「文件注釋系統設計與開發之研究」，碩士論文，頁 28，國立交通大學資訊學院數位圖書資訊組，新竹。

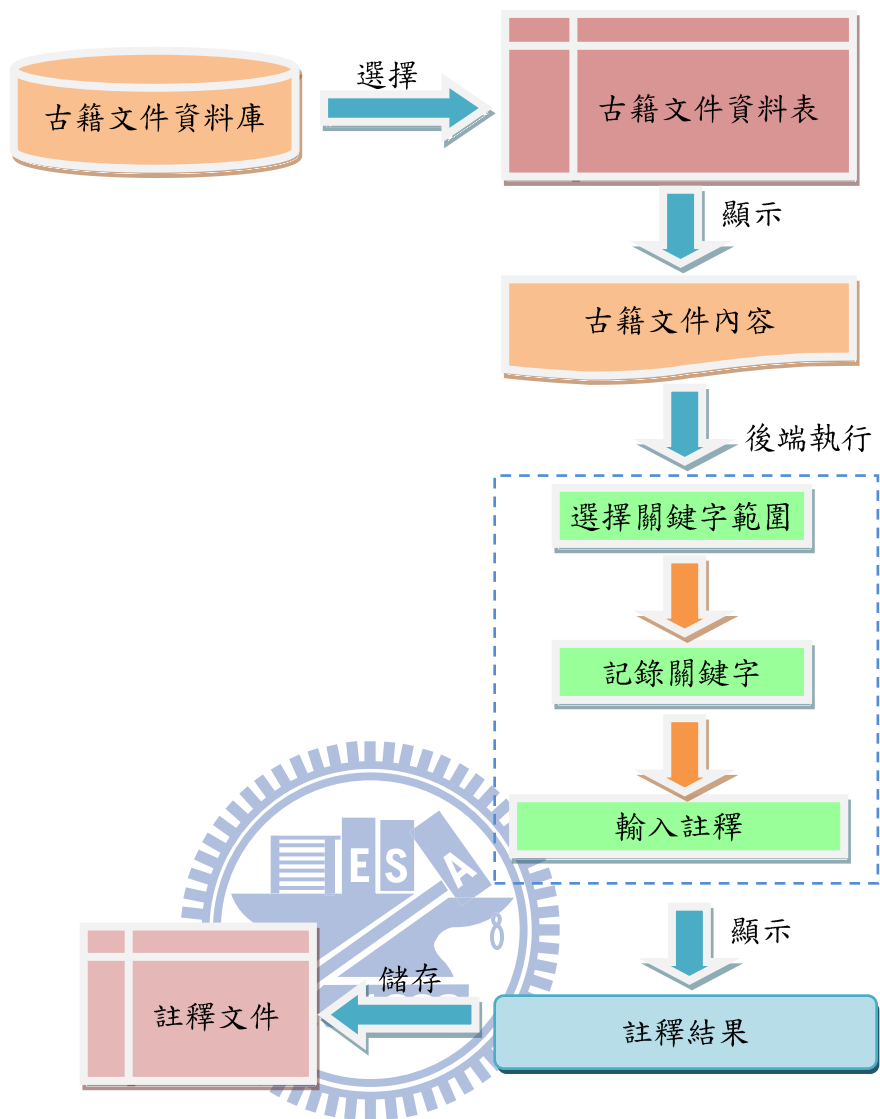


圖 3-7：古籍內容註釋模組流程圖

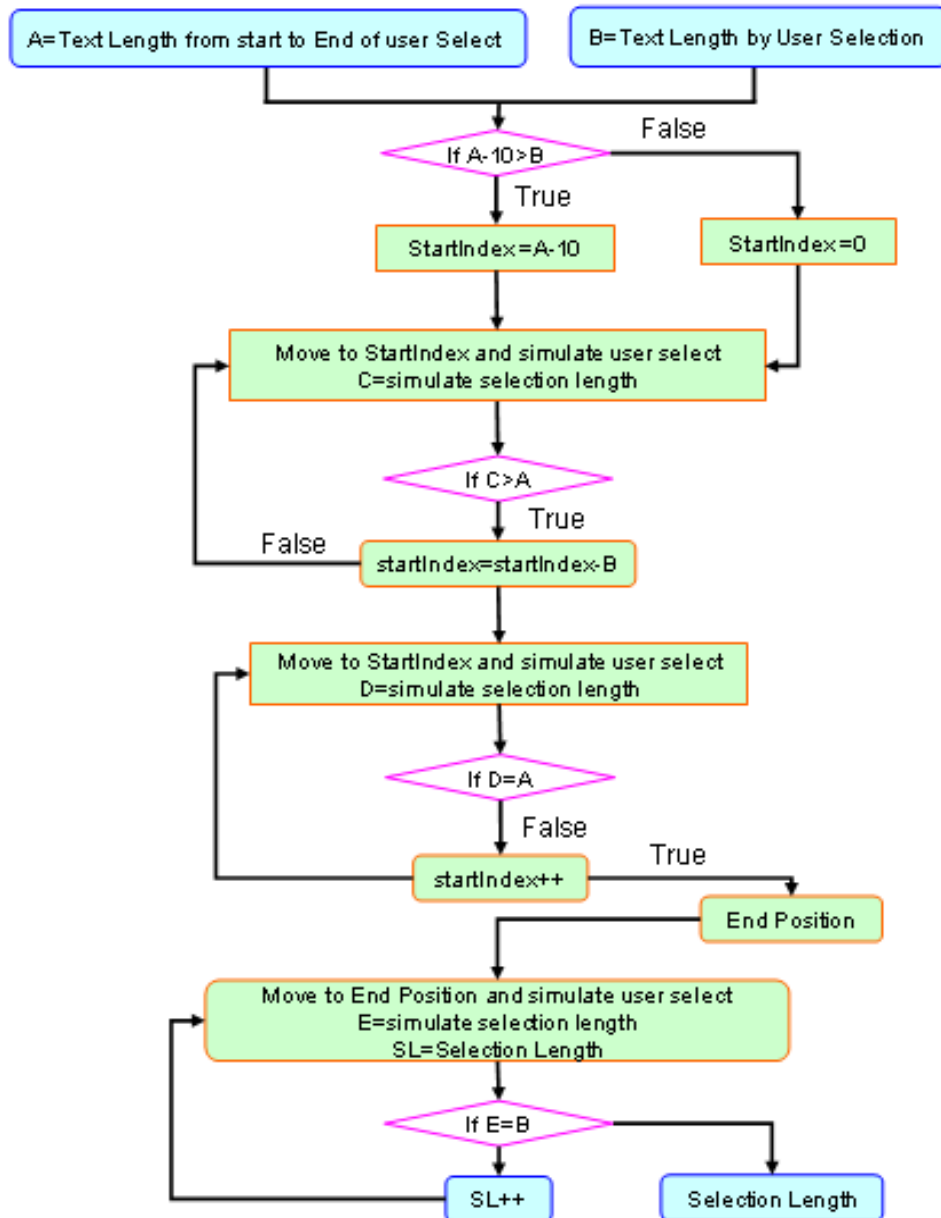


圖 3-8：文字定位流程圖

3.3.4 輔助工具模組

此模組主要標示古籍文件校勘後顯示差異的文字或字句；藉由此模組中的字庫，標示相關的文字或訊息，輔助使用者執行文字校勘的工作，如圖3-9。

古籍文件字彙、字詞博大精深且複雜，現有字庫無法完全滿足所有古

籍或現有通行的字彙，為了擴充系統的同義字或同義詞，讓使用者可以自行建立字庫資料，方便後續進行校勘工作，故發展此模組，其功能包含了可建立避諱字庫：中國歷史上各朝代皇帝之名諱、習俗或禁忌；建立同義字庫：如煙、烟…等。

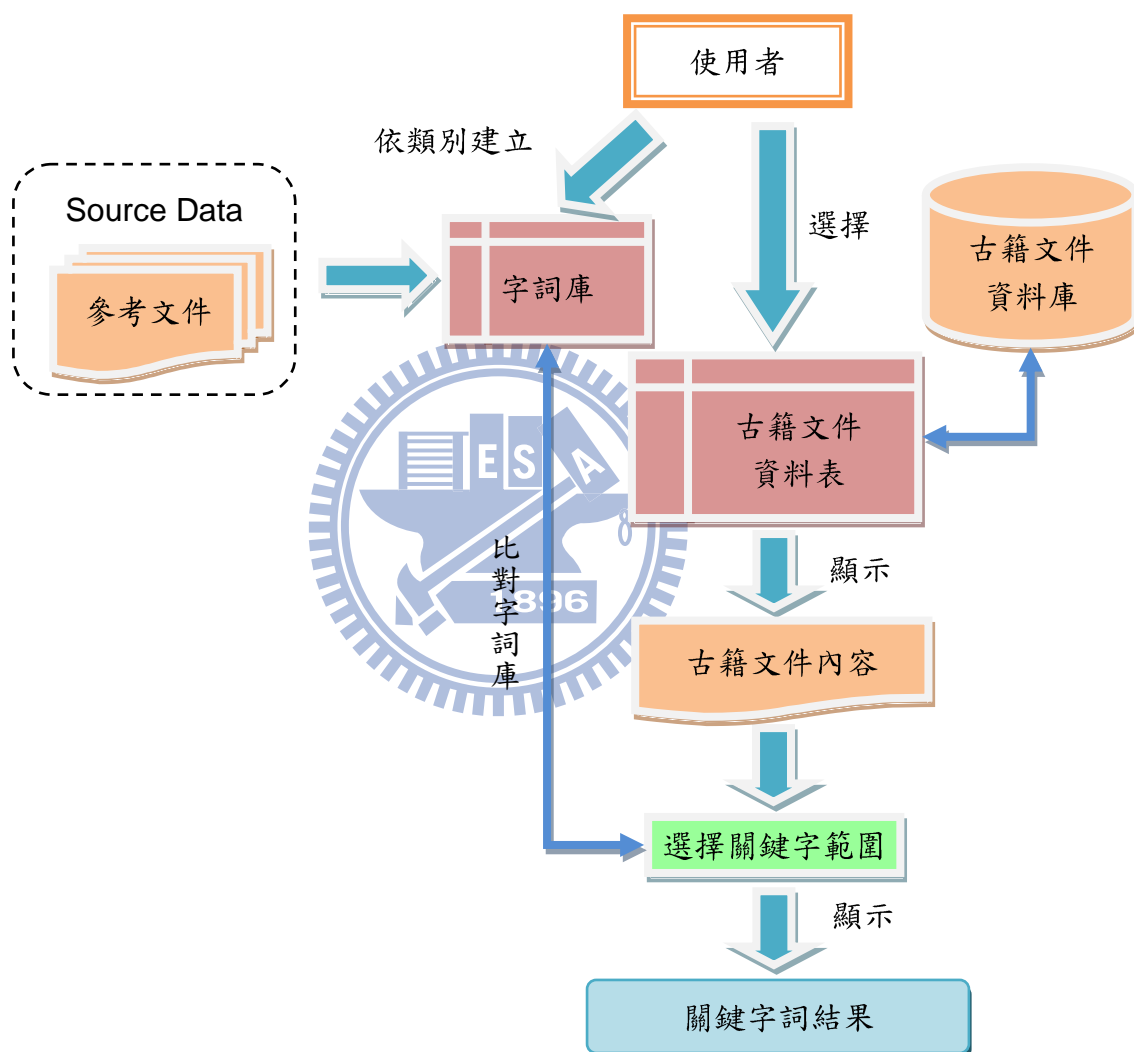


圖 3-9：輔助工具模組流程圖

3.3.5 個人收藏模組

使用者可以將文件內容註釋、校勘後的結果加入「個人收藏」，並透過此模組管理「個人收藏」，如圖3-10；如果系統將該古籍文件之註釋預設為

非公開分享的狀態，註釋者須透過此模組的分享功能進行資料公開的步驟，始可分享給其他使用者；若使用者對系統中之資料有興趣，但為其他使用者所註釋擁有，亦可加入個人收藏或刪除；若為個人所註釋的資料，並已公開分享，則無法執行刪除的動作，避免影響其他使用者所建立的相關資料連結。

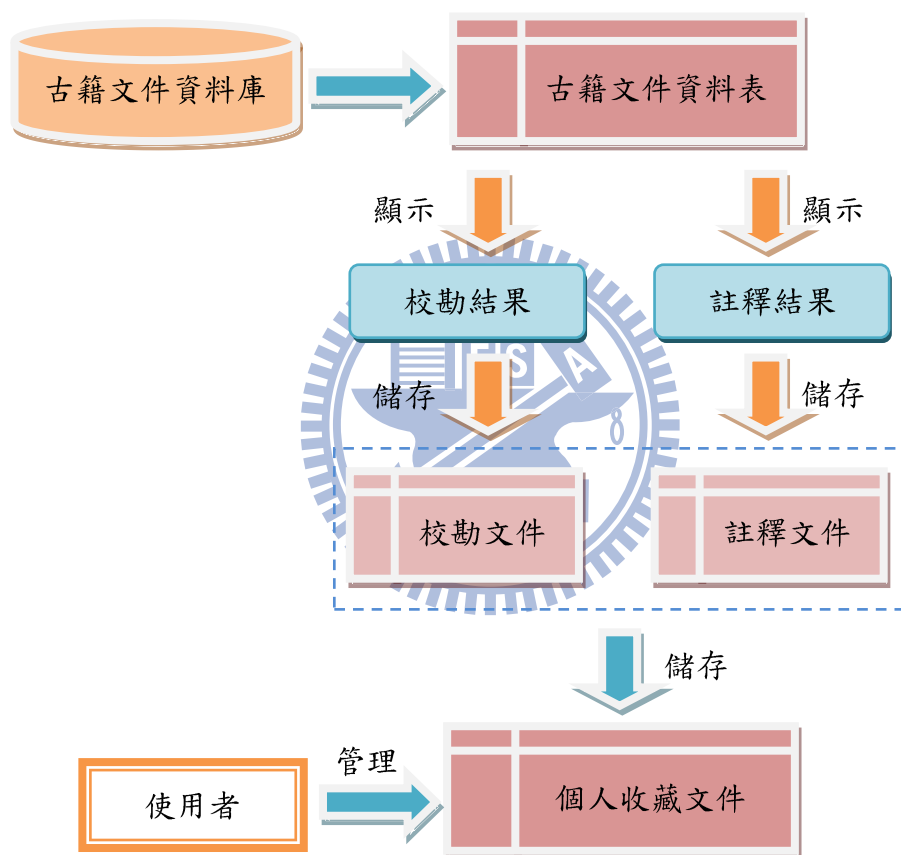


圖 3-10：個人收藏模組流程圖

3.4 資料表設計與欄位定義

根據本系統設計模組所需的資料內容、校勘方式與流程，總共定義10個資料表、相關欄位名稱與功能及資料表關聯，如圖3-11，詳細說明如下：

1. admin：校勘系統各階層使用者資訊。

表 3-1：校勘系統使用者資訊欄位一覽表

欄位	型態	功能描述
systemID	int(11)	系統自動產生序號
user	varchar(100)	使用者帳號資訊
pw	varchar(100)	使用者帳號密碼資訊
group	tinyint(4)	區分帳號階層（管理者、校勘研究者、一般使用者）

2. category：古籍類別名稱資訊。

表 3-2：古籍類別名稱資訊欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
name	varchar(100)	第一階類別名稱（如：儒家、道家、墨家、雜家）
description	text	類別描述

3. subcategory：古籍列表名稱。

表 3-3：古籍列表名稱資訊欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
name	varchar(100)	第二階列表名稱（如：老子、呂氏春秋、論語）
description	text	列表描述
categoryID	int(11)	與類別對應之序號

4. title：古籍篇名名稱。

表 3-4：古籍篇名名稱資訊欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
name	varchar(100)	第三階篇名名稱(如：老子、呂氏春秋、論語)
description	text	篇名描述
categoryID	int(11)	與類別對應之序號
subcategoryID	int(11)	與列表對應之序號
hierarchy	tinyint(4)	篇名階層設定

5. content：古籍內容資訊。

表 3-5：古籍內容資訊欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
description	text	版本內容描述
categoryID	int(11)	與類別對應之序號
subcategoryID	int(11)	與列表對應之序號
titleID	int(11)	與篇名對應之序號
content	test	古籍文件內容資料

6. notes：古籍註釋資訊。

表 3-6：古籍註釋資訊欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
notetitle	text	使用者所圈選的關鍵字範圍
notecontent	text	關鍵字中定義的註釋
author	varchar(200)	定義關鍵字註釋的使用者
contentID	varchar(100)	與古籍內容資訊對應之序號
share	tinyint(4)	是否分享給其他使用者參考

7. notes_compare：古籍校勘結果。

表 3-7：古籍校勘結果資訊欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
content1	tinyint(4)	選擇古籍內容為「底本」
content2	tinyint(4)	選擇古籍內容為「校本」
similar_str	varchar(500)	顯示相似段落
notes	text	相似段落註釋
author	varchar(100)	執行校勘註釋的使用者
share	tinyint(4)	是否分享給其他使用者參考

8. tools_title：字詞庫名稱。

表 3-8：字詞庫名稱欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
name	varchar(100)	字詞庫名稱
description	text	字詞庫功用描述

9. tools_detail：儲存字詞庫內容資訊。

表 3-9：儲存字詞庫內容資訊欄位一覽表

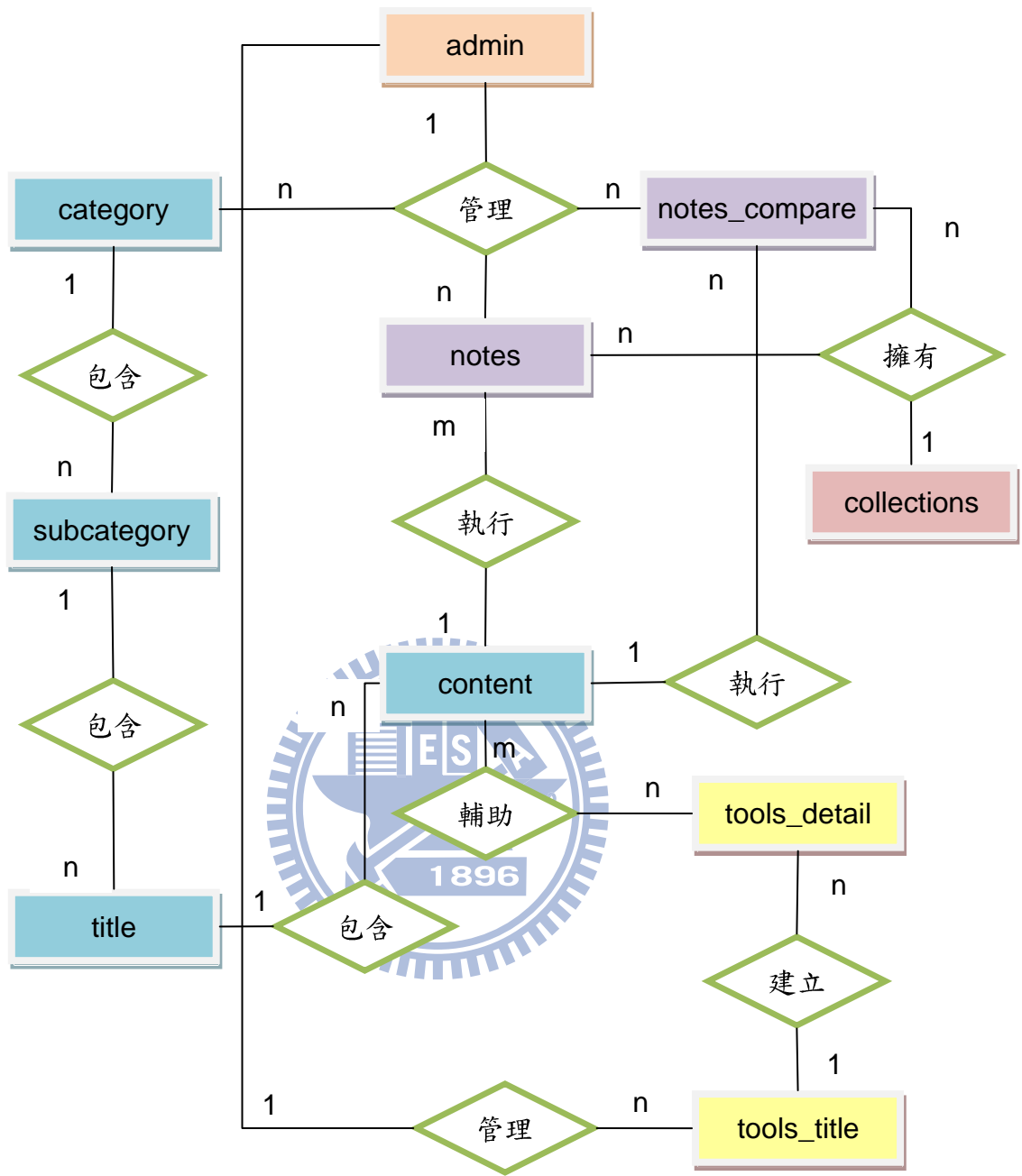
欄位	型態	功能描述
id	int(11)	系統自動產生序號
name1	varchar(100)	內容字詞
name2	varchar(100)	相對應字詞
description	text	字詞解釋
titleID	varchar(100)	字詞庫名稱序號
example	varchar(500)	相關字詞解釋例句
source	varchar(200)	相關篇章來源

10. collections：個人收藏資訊。

表 3-10：個人收藏資訊欄位一覽表

欄位	型態	功能描述
id	int(11)	系統自動產生序號
collector	varchar(100)	收藏者帳號
notesID	varchar(100)	納入收藏的註釋
notes_compareID	varchar(100)	納入收藏的校勘





圖示說明：

1 : 1=在關聯中為1對1的關係，1 : n=在關聯中為1對多的關係

m : n=在關聯中為多對多的關係

圖 3-11：資料表關聯圖

第四章 系統功能展示

本章內容區分為三部份，4.1 節介紹系統實作的環境、4.2 節介紹系統實作結果、4.3 節為系統功能完成度與效益。

4.1.系統實作環境

本研究的實作環境如圖 4-1 所示，系統作業平台為 Window Server 2003，使用 AppServ 做為伺服器工具，內含 Apache Httpd(網頁伺服器)、MySQL (資料庫)、phpMyAdmin (資料庫管理工具) 等網站所需工具，本研究所使用的 PHP 網頁也是透過 Apache 引擎進行處理，並利用 AppServ 所提供的 MySQL 做為資料庫系統，系統開發工具則為 PHP、HTML、CSS 及 JavaScript。

古籍文件校勘系統利用 AppServ 將所有網站伺服器需要的套件整合之便利性，且沒有版權問題，而 AppServ 中的 Apache 網頁伺服器可以處理許多類型的網頁，如 PHP、HTML、CSS…等，讓開發過程更為便利順暢，其中更整合 MySQL 資料庫，讓網頁及資料庫的連結在安裝後即可使用，網頁與資料庫間的參數設定亦更有效率，節省開發時程。

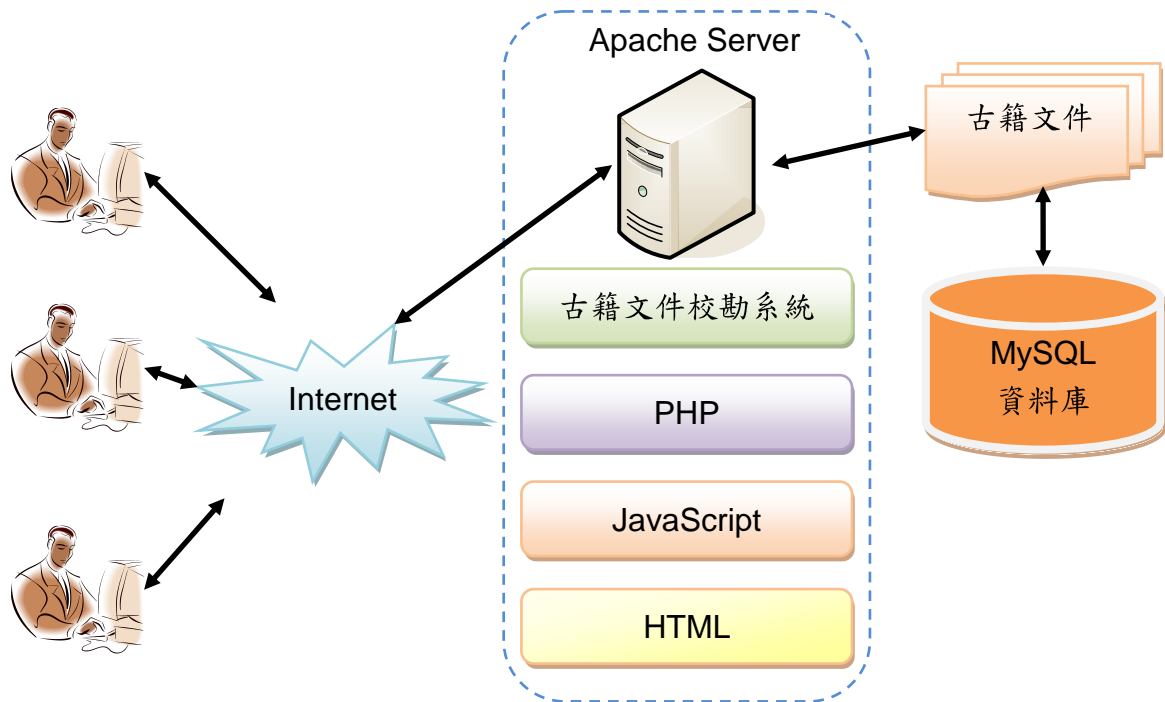


圖 4-1：系統實作環境圖

4.2. 系統實作結果

古籍文件校勘系統依前面章節之描述，主要分為：古籍文件管理模組、古籍文件校勘模組、古籍註釋模組、輔助工具模組、個人收藏模組等五個部份，如圖 4-2；其中有許多功能須記錄使用者相關訊息，必須經過註冊及身份認證後才可使用。主要群組分為「管理者」、「校勘研究者」及「一般使用者」；其中「管理者」能使用系統中所有功能與操作，主要職責是在系統或後端資料庫產生錯誤時加以修正，當然，若管理者對校勘有興趣，亦可加入校勘的行列，成為「校勘研究者」；「校勘研究者」為此系統主要的群體，他們對古籍文件及文字較為敏感，以往校勘古籍時，因需要翻閱大

量參考資料並做成紙本紀錄，費時又耗力，利用此系統則可以縮短時間，「校勘研究者」同時具有上傳古籍文件資料的權限；另一個群體為「一般使用者」，僅能使用「校勘工具」、「註釋工具」及「個人收藏」等三項功能，所有的使用者均須登入系統才能使用各項功能，未登入之使用者僅能閱覽古籍文件。古籍自動校勘系統平台頁面如圖 4-3。

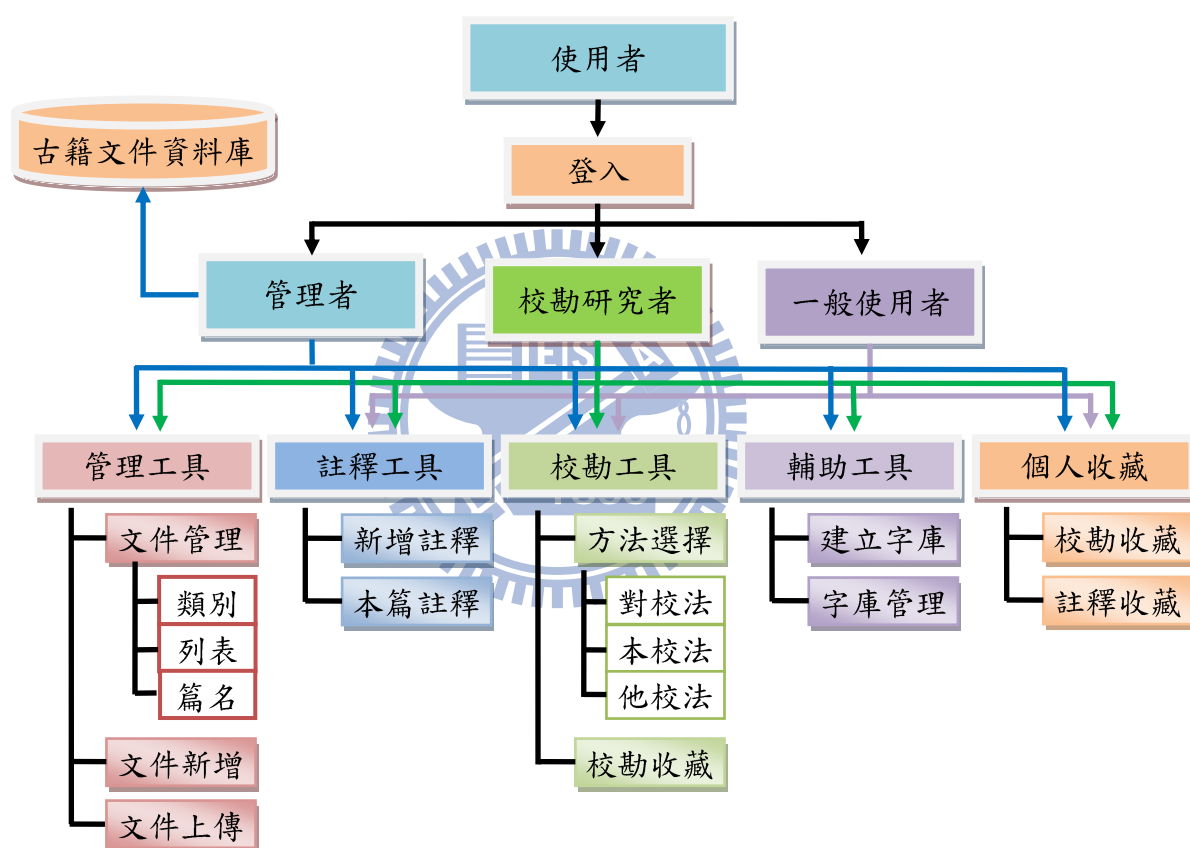


圖 4-2：系統功能架構圖

首頁	道家 > 老子 > 道經
管理工具	
註釋工具	道可道、非常道。 名可名，非常名。無名天地之始；有名萬物之母。故常無欲，以觀其妙；常有欲，以觀其微。此兩者，同出而異名，同謂之玄。玄之又玄，眾妙之門。
校勘工具	天下皆知美之為美，斯惡已。皆知善之為善，斯不善已。故有無相生，難易相成，長短相較，高下相傾，音聲和，前後相隨。是以聖人處無為之事，行不言之教；萬物作焉而不辭，生而不有。為而不恃，功成而弗居。夫弗居，是以不去。
輔助工具	不尚賢，使民不爭；不貴難得之貨，使民不為盜；不見可欲，使心不亂。是以聖人之治，虛其心，實其腹，弱其志，強其骨。常使民無知無欲。使夫1知者不敢為也。為無為，則無不治。
個人收藏	道沖而用之或不盈。淵兮似萬物之宗。挫其銳，解其紛，和其光，同其塵。湛兮似或存。吾不知誰之子，象帝先。
古籍瀏覽	天地不仁，以萬物為芻狗；聖人不仁，以百姓為芻狗。天地之間，其猶橐籥乎？虛而不屈，動而愈出。多言數
<ul style="list-style-type: none"> ☐ 儒家 ☐ 兵家 ☐ 法家 ☐ 雜家 ☐ 道家 	

圖 4-3：古籍文件自動校勘系統頁面

4.2.1 古籍文件管理功能

1. 文件類別新增及上傳

古籍文件自動校勘比對之前，將文件分類是相當重要的一件事，不僅讓古籍有系統條目的分門別類外，更可以讓使用者快速執行校勘之資料。在此可以利用系統中「管理工具」的「新增」功能建立古籍類別、古籍列表、古籍篇名以及版本描述，其中，版本描述主要區分同一篇名版本卻不同的說明，以方便使用者在校勘時能區分版本的差異性。

古籍文件類別建立完成後，再依序建立「古籍列表」與「古籍篇名」，例如：儒家、道家、法家、兵學、雜家等古籍類別，而儒家古籍中又有孔子、孟子、荀子等人之著作列表，孔子的論語中又區分為學而篇、為政篇…等篇名，依著作列表、篇名逐步建立，以利後續古籍文件校勘、搜尋與閱覽，如圖4-4、4-5。

我要建立 個階層
類別建立

列表建立

篇名建立

圖 4-4：古籍類別、列表、篇名「新增」功能



圖 4-5：古籍類別、列表、篇名建立完成顯示

完成「古籍文件類別」、「古籍列表」與「古籍篇名」建立後，再利用管理工具中「上傳」功能將文件內容以純文字(.txt)檔案上傳至相對應之古籍篇名中，如圖4-6，上傳後系統會經過程式轉換成HTML網頁，上傳成功後回覆訊息給使用者。

圖 4-6：古籍文件上傳功能

以下以建立「道家」古籍類別，古籍列表「文子」，古籍篇名「精誠」為例；使用者於「管理工具」中點選「新增」，並依欄位描述輸入所要建立的古籍類別、列表與篇名（如圖 4-7），逐步建立完成後，左側樹狀圖列產生「文子」的資料夾（如圖 4-8），再利用「上傳」功能（如圖 4-9），將文件內容傳至相對應的篇名，再執行點選即可閱覽所選擇的古籍文件（如圖 4-10），並可利用「註釋」、「校勘」及「輔助」等工具執行所需工作，此部份將於後面章節詳細說明。

圖 4-7：古籍列表建立



圖 4-8：產生建立的古籍列表資料夾



圖 4-9：古籍篇名文件內容上傳

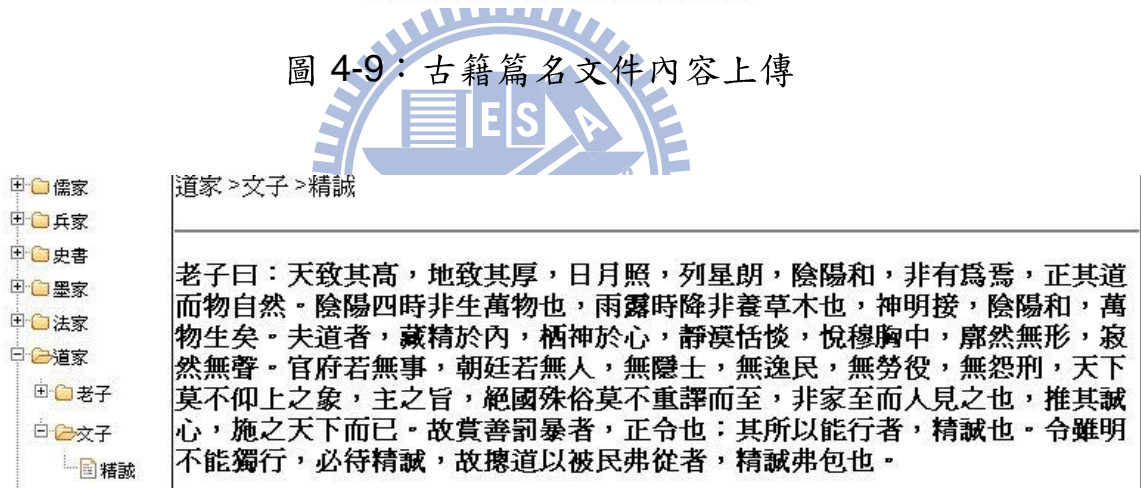


圖 4-10：顯示已上傳的文件內容

2. 古籍文件選擇與管理

「古籍文件選擇」不需身份認證，未登入的使用者點選頁面左側的樹狀圖就能開啟與閱覽文件內容，除了古籍文件內容之外，亦可觀看其他人註釋的內容，以提供參考建議，若要執行「校勘」與「註釋」功能則必須登入系統。

而點選管理工具中之「管理」功能後，顯示「古籍類別」、「古籍列表」與「古籍篇名」等三種項目（如圖4-11），每一項目皆可編輯與刪除名稱的動作，若點選「古籍篇名」除可以編輯與刪除名稱外，亦可編輯文件內容與文件名稱，但編輯權限僅限於「管理者」與「校勘研究者」；唯一要注意的是，因「古籍類別」為最上層之資料夾，其中包含了「古籍列表」與「古籍篇名」等第二、三層的資料夾，若對「古籍類別」資料夾執行刪除動作，除了刪除所屬列表及篇名外，相關註釋及校勘結果亦會一併刪除，故使用者在執行「刪除」之前必須審慎考慮，而系統亦會出現對話框提醒使用者確認是否刪除，另外，因「管理者」擁有最高權限，為避免誤刪「校勘研究者」所上傳的資料，在執行刪除動作時，亦會顯示對話框提醒。



圖 4-11：古籍「類別」、「列表」與「篇名」項目中編輯與刪除功能

4.2.2 古籍文件校勘功能

古籍文件校勘功能是利用電腦之記憶、檢索、程式比對及連結等特點，來校勘上傳至系統中的古籍文件資料；使用者登入後選擇古籍類別，再依古籍列表中點選不同版本的文件並利用「校勘工具」中的「對校法」、「本校法」及「他校法」等功能，將執行結果顯示出來後，利用註釋功能予以註解，如圖4-12；另外，具備修正文件內容權限的「管理者」與「校勘研究者」，盡量不更改文件中的內容，以保持文件的完整性，俾利後續研究者考證。下列分別介紹「對校法」、「本校法」及「他校法」等項目之功能。



圖 4-12：校勘功能選擇頁面

1. 對校法

是校勘的基本方法，可了解底本或別本的本來面目，純屬機械式的方法，依陳垣所述：「即以同書之祖本或別本對讀，遇不同之處，則注于其旁」，校勘方式如圖 4-13，即校勘出不同之文字或段落後在旁邊予以註解，以利後人查閱。

以「老子」「道經」第一篇馬王堆與四庫全書版本為例，如圖 4-14，執行完畢後以顏色顯示不同意義，黑色為相同段落，橘色為已註釋的關鍵字，紅色字體則為相異段落或文字，結果如圖 4-15。

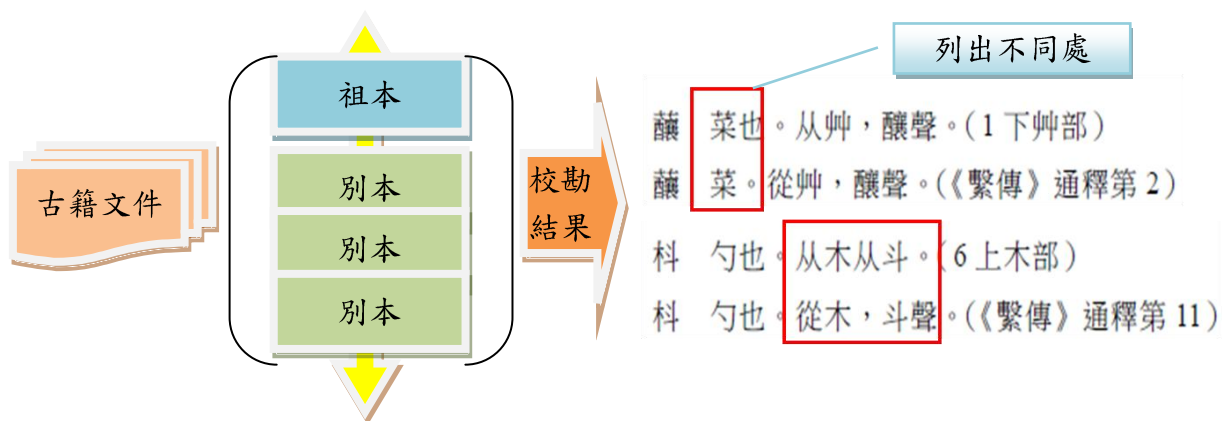


圖 4-13：對校法概念圖

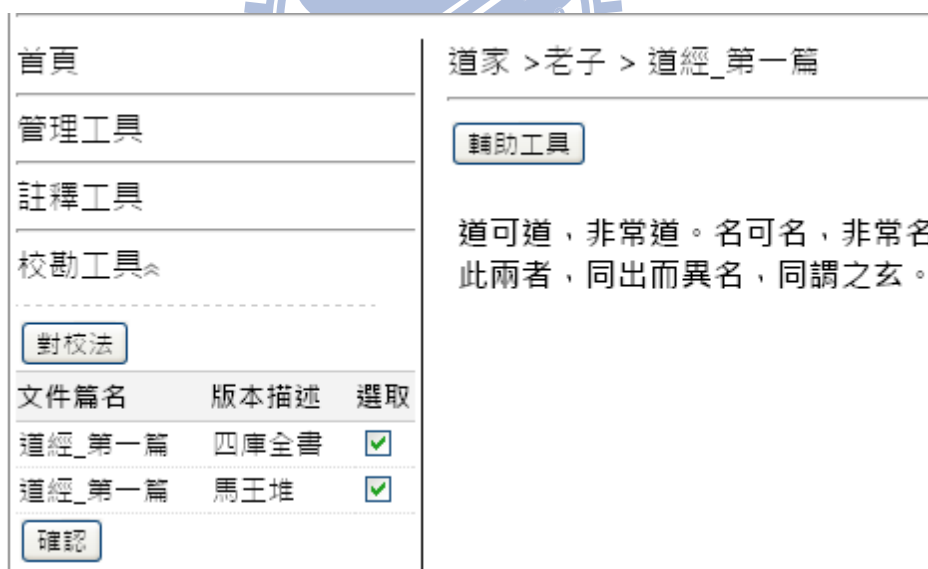


圖 4-14：對校法版本選擇

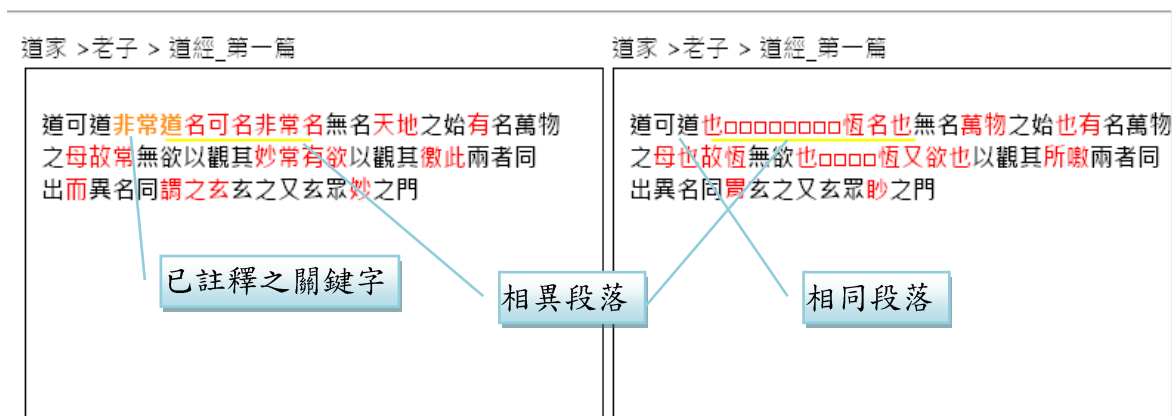


圖 4-15：對校法之顯示結果

2.本校法

即以本書校本書，根據本書的上下文意、相同或相近的句法、詞例來校勘本書的錯誤。陳垣曰：「本校法者，以本書前後互證，而抉摘其異同，則知其中之謬誤。」校勘方式如圖 4-16，即對文件本身之內容進行校勘比對。

以「呂氏春秋」一書為例，區分為孟春紀、仲春紀、季春紀…等 26 篇共 160 卷，若採用傳統校勘方式，需逐篇逐句才能找出相似段落，以「似順論_慎小」為底本透過校勘系統中的「本校法」選擇欲校勘的篇章（仲冬紀_長見、恃君覽_觀表），再選擇待校勘的篇名（仲冬紀_長見），如圖 4-17，能快速檢索出相似段落，再針對校勘結果予以註釋，如圖 4-18。

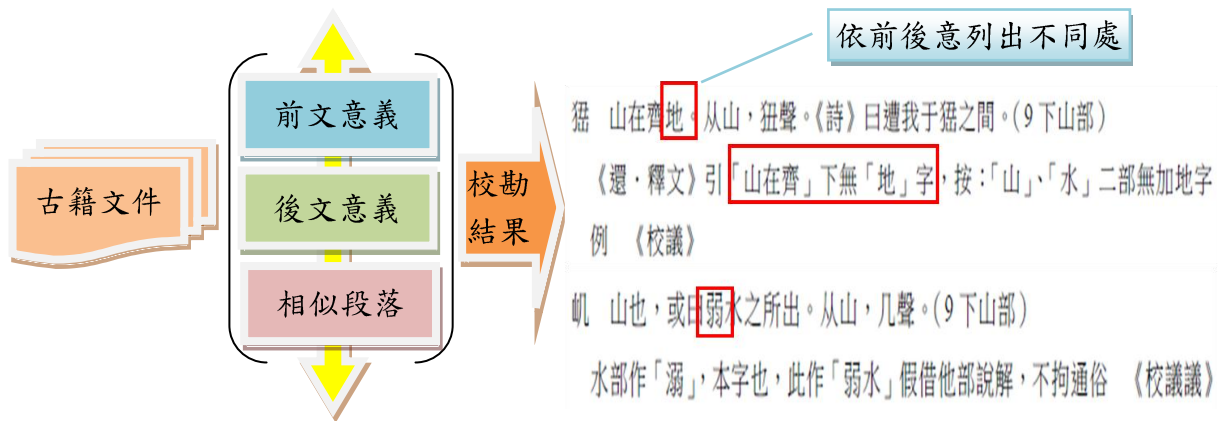


圖 4-16：本校法概念圖

<p>首頁</p> <p>管理工具</p> <p>註釋工具</p> <p>校勘工具</p> <p>對校法</p> <p>本校法</p> <p>文件篇名 <input type="checkbox"/></p> <p>仲秋紀_決勝 <input type="checkbox"/></p> <p>仲冬紀_長見 <input checked="" type="checkbox"/></p> <p>恃君竟_觀表 <input checked="" type="checkbox"/></p> <p>確認 共選取2篇</p> <p>待校勘結果</p> <table border="1"> <thead> <tr> <th>文件篇名</th> <th>相似段落</th> <th>選取</th> </tr> </thead> <tbody> <tr> <td>仲冬紀_長見</td> <td>後世</td> <td><input checked="" type="checkbox"/></td> </tr> <tr> <td>恃君竟_觀表</td> <td>矣人</td> <td><input type="checkbox"/></td> </tr> </tbody> </table> <p>進行校勘 共2筆待校勘</p>	文件篇名	相似段落	選取	仲冬紀_長見	後世	<input checked="" type="checkbox"/>	恃君竟_觀表	矣人	<input type="checkbox"/>	<p>雜家 > 呂氏春秋 > 似順論_慎小</p> <p>輔助工具</p> <p>上尊下卑。卑則不得以小觀上。尊則恣，恣則輕小物，下，下怨上矣。人臣之情，不能為所怨；人主之情，不巨防容蠖，而漂邑殺人；突洩一爍，而焚宮燒積；將失衛獻公戒孫林父、甯殖食。鴻集于園，虞人以告，公如子。二子不說，逐獻公，立公子黜。衛莊公立，欲逐石臼：“戎州也。”莊公曰：“我姬姓也。戎人安敢居國莊公，立公子起。此小物不審也。人之情不靈於山，而齊桓公即位，三年三言，而天下稱賢，群臣皆說。去肉吳起治西河，欲諭其信於民，夜日置表於南門之外，令日日晏矣，莫有償表者。民相謂曰：“此必不信。”有謁吳起。吳起自見而出，仕之長大夫。夜日又復立表，是之後，民信吳起之賞罰。賞罰信乎民，何事而不成，</p>
文件篇名	相似段落	選取								
仲冬紀_長見	後世	<input checked="" type="checkbox"/>								
恃君竟_觀表	矣人	<input type="checkbox"/>								

圖 4-17：選擇本校法的底本與待校勘的相似段落

校勘結果				
底本	校本	相似段落	注釋內容	管理
雜家 > 呂氏春秋 > 似順論_慎小	雜家 > 呂氏春秋 > 仲冬紀_長見	大夫	<input type="text"/>	<input type="checkbox"/>
雜家 > 呂氏春秋 > 似順論_慎小	雜家 > 呂氏春秋 > 仲冬紀_長見	三年	<input type="text"/>	<input type="checkbox"/>
雜家 > 呂氏春秋 > 似順論_慎小	雜家 > 呂氏春秋 > 仲冬紀_長見	為後世	<input type="text"/>	<input type="checkbox"/>
雜家 > 呂氏春秋 > 似順論_慎小	雜家 > 呂氏春秋 > 仲冬紀_長見	吳起治西河	<input type="text"/>	<input type="checkbox"/>
雜家 > 呂氏春秋 > 似順論_慎小	雜家 > 呂氏春秋 > 仲冬紀_長見	天下	<input type="text"/>	<input type="checkbox"/>

圖 4-18：本校法之顯示結果

3.他校法

即參考其他相關書籍校勘，陳垣曰：「他校法者，以他書校本書。凡其書有采自前人者，以前人之書校之；有為後人所引用者，以後人之書校之，其史料有為同時之書所並載者，可以同時之書校之。」校勘方式如圖 4-19，執行該校勘方法時須廣泛蒐集該類別文件並儲存於系統中，以增加古籍文件之間的參考性。

以「後漢書」中「列傳」為例，執行方式與上述「本校法」相同，不同之處在於選擇範圍較廣，因需要以「他書」校「本書」，故針對相關性之書籍選擇，如圖 4-20，顯示待結果結果之篇名後(孫子_軍形篇)實施校勘，如圖 4-21。

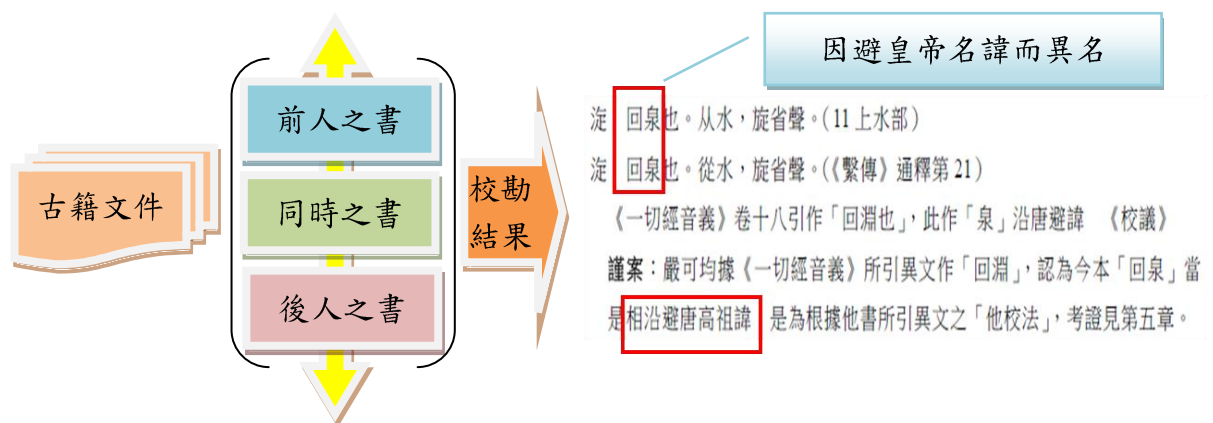


圖 4-19：他校法概念圖

<p>首頁</p> <p>管理工具</p> <p>註釋工具</p> <p>校勘工具</p> <p>對校法</p> <p>本校法</p> <p>他校法</p> <p>儒家</p> <p>古籍列表 <input checked="" type="checkbox"/></p> <p>論語 <input checked="" type="checkbox"/></p> <p>兵家</p> <p>古籍列表 <input checked="" type="checkbox"/></p> <p>孫子 <input checked="" type="checkbox"/></p>	<p>史書 > 後漢書 > 列傳</p> <p>輔助工具</p> <p>皇甫嵩字義真，安定朝那人，度遼將軍規之兄子也。父節，廩門太守。嵩少孝廉、茂才。太尉陳蕃、大將軍竇武連辟，並不到。靈帝公車徵為議郎，遷初，鉅鹿張角自稱“大賢良師”，奉事黃老道，畜養弟子，跪拜首過，符水之。角因遣弟子八人使於四方，以善道教化天下，轉相誑惑。十餘年間，眾幽、冀、荆、揚、兗、豫八州之人，莫不畢應。遂置三十六方。方猶將軍號渠帥。訛言“蒼天已死，黃天當立，歲在甲子，天下大吉”。以白土書京城：中平元年，大方馬元義等先收荆、揚數萬人，期會發於鄴。元義數往來京師以三月五日內外俱起。未及作亂，而張角弟子濟南唐周上書告之，於是車裂隸，使鉤盾令周斌將三府掾屬，案驗宮省直衛及百姓有事角道者，誅殺千餘已露，晨夜馳敕諸方，一時俱起。皆著黃巾為標幟，時人謂之“黃巾”，亦稱“天公將軍”，角弟寶稱“地公將軍”，寶弟梁稱“人公將軍”，所在燔多逃亡。旬日之間，天下嚮應，京師震動。</p> <p>詔敕州郡修理攻守，簡練器械，自函谷、大谷、廣城、伊闕、轅轅、旋門、</p>
---	--

圖 4-20：其他相關古籍選擇

校勘結果				
底本	校本	相似段落	注釋內容	管理
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	先為不可勝以待敵之可勝不可勝在	<input type="text"/>	<input type="checkbox"/>
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	不足	<input type="text"/>	<input type="checkbox"/>
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	有餘	<input type="text"/>	<input type="checkbox"/>
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	守者	<input type="text"/>	<input type="checkbox"/>
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	九地之下	<input type="text"/>	<input type="checkbox"/>
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	攻者	<input type="text"/>	<input type="checkbox"/>
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	九天之上	<input type="text"/>	<input type="checkbox"/>
史書 > 後漢書 > 列傳	兵家 > 孫子 > 軍形	日月	<input type="text"/>	<input type="checkbox"/>

圖 4-21：他校法之顯示結果

4.2.3 古籍內容註釋功能

1. 新增註釋

使用者選定某一篇古籍後，可以對古籍文件的內容進行註釋，選取所要執行註釋的字詞或段落範圍，而系統會將之當做關鍵字，連同註釋的內容一併存入資料庫中，提供使用搜尋運用；使用者完成註釋內容後，可選擇僅個人閱覽，或選擇「分享」，讓其他使用者亦可觀看到此註釋並加入收藏；對於同一篇古籍文件，透過不同使用者的詮釋，可以得到不同的見解，使文件內容更加豐富，如圖4-22所示。

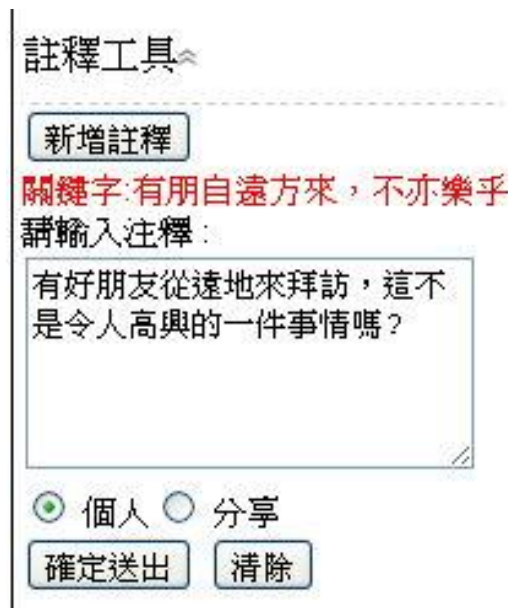


圖 4-22：古籍註釋工具「新增註釋」

2. 本篇註釋

使用者可選取「註釋工具」中的「本篇註釋」，查看古籍文件內容中有哪些字詞或段落為個人或其他使用者所分享的註釋，並且可將其他人分享的註釋加入「收藏」。選擇文件後，可看到有註釋的文字或段落以不同顏色標示，如圖 4-23；另外註釋的「編輯」與「刪除」功能則可以修改個人的註釋，或加入「收藏」及「分享」給其他使用者，如圖 4-24。

儒家 > 論語 > 學而

子曰：“學而時習之，不亦說乎？有用自遠方來，不亦樂乎？人不知而不慍，不亦君子乎？”

有子曰：“其爲人也孝弟，而好犯上者，鮮矣；不好犯上，而好作亂者，未之有也。君子務本，本立而道生。孝弟也者，其爲仁之本與！”

圖 4-23：古籍文件中具有註釋的範圍

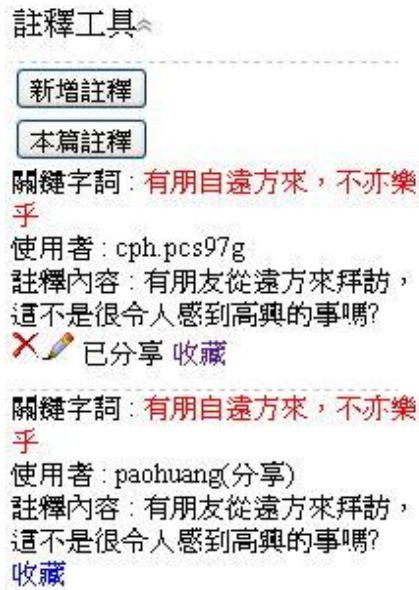


圖 4-24：古籍註釋工具「本篇註釋」

4.2.4 輔助工具功能

由於古籍文件中常有許多用字與現今不同，但具有同音同義、異字同義、同字不同義等現象，為幫助使用者能更快速了解校勘文字或詞句的差異，故建立輔助工具功能，允許「管理者」及「校勘研究者」自行建立輔助字庫，如：同義字庫、避諱字庫等，讓使用者更能了解差異文字的意義，增加校勘的正確性，如圖 4-25、4-26、4-27 所示。

輔助工具建立

輔助表名稱: 請輸入名稱

輔助表功能: 請輸入功能描述

建立

圖 4-25：輔助工具表建立畫面

輔助表名稱	功能描述	管理
同義字表	請輸入功能描述	管理 + ✎ ✕
逆義字表	請輸入功能描述	管理 + ✎ ✕
古代地名	請輸入功能描述	管理 + ✎ ✕
古代官名	請輸入功能描述	管理 + ✎ ✕

圖 4-26：輔助工具建立與管理畫面

儒家 > 論語 > 學而

輔助工具

開啟輔助工具功能

子曰：“學而時習之，不亦說乎？有朋自遠方來，不亦樂乎？人不知而不愠，不亦君子乎？”

有子曰：“其為人也孝弟，而好犯上者，鮮矣；不好犯上，而好作亂者，未之有也。君子務本，本立而道生。孝弟也者，其為仁之本與！”

子曰：“巧言令色，鮮矣仁！”

曾子曰：“吾日三省吾身：為人謀而不忠乎？與朋友交而不信乎？傳不習乎？”

關鍵字:說

名稱	描述
悅	快樂

詳細資料

名稱	說
名稱	悅
描述	快樂
例句	有朋自遠方來，不亦說乎
篇章	論語_學而

圖 4-27：開啟輔助工具功能與關鍵字詞的解釋

4.2.5 個人收藏功能

使用者在登入後可以管理個人收藏，目前系統提供管理「註釋收藏」與「校勘收藏」的功能，使用者選擇古籍類別、古籍列表，再點選搜尋之後，系統會過濾掉其他文件，只將個人收藏的註釋載入列表中，避免增加使用者尋找文件困難度，如圖 4-28 所示。

另外，在註釋（校勘）收藏中使用者可以點選所註釋的字詞或段落內容，返回原始完整的文件，再加以修正，其預設值為未分享狀態，當然也可以分享或關閉註釋（校勘結果），或刪除個人已收藏的註釋（校勘結果），

如果為已經公開分享的註釋（校勘），為避免有其他使用者已納入個人收藏，造成資訊遺失，系統不提供刪除功能。



圖 4-28：個人工具「註釋收藏」顯示畫面

4.3 系統功能完成度與效益

各系統功能完成度與可達到的效益如表 4-1 所示。

表 4-1：系統功能完成概況表

系統功能名稱		預期效益	完成概況
帳號登入、註冊		提供使用者登入系統，及註冊新帳號	●
管理工具	管理	「類別」、「列表」及「篇名」的編輯與刪除	●
	新增	「類別」、「列表」及「篇名」階層的建立	●
	上傳	提供古籍文件資料依建立的項目匯入系統	●
註釋工具	新增註釋	可將選取的文字範圍實施註解	●
	本篇註釋	顯示目前文件有哪些註釋，並可執行刪除、修改、分享與收藏等動作	●
校勘工具	對校法	針對同一種古籍不同版本，比對出差異部份後，可進行註釋、分享	△ 可比對出差異段落
	本校法	於同一本古籍中，不同章節或篇名之內容，比對出哪些段落相互引用，建立文件之間的關係，並可進行註釋與分享	△ 可比對出差異段落
	他校法	同本校法，但範圍較大，不單單限於同一本古籍內容	同上
輔助工具	新增	動態建立的輔助表，可幫助使用者實施校勘	●
	管理	管理建立的輔助表，並提供編輯等功能	●
個人收藏	註釋收藏	可將個人有興趣的註釋或校勘結果實施收藏，便於日後快速查找	●
	校勘收藏		●

圖示說明：●完成；△技術或方法仍須克服及改進。

第五章 結論與建議

本章節就整篇論文提出未來可能的研究方向；5.1 節主要提出本研究的結論及對於使用者的貢獻，5.2 節為本論文研究過程中之限制因素，以及從系統的角度，提出不足或是可精進擴充的方法，做為未來研究的方向。

5.1 研究結論與貢獻

到目前為止台灣鮮少有對古籍進行自動校勘之研究，此項工作除了存在一定的技術與難度之外，對於「古籍」所涉及之學術範疇又十分廣泛，如版本學、目錄學、考證學、音韻學…等，另外又有如何設計精進校勘系統等問題，顯示出古籍文件自動校勘之研究與工作，需要專業團隊長期投入與支持才能漸顯效益。

本論文針對現行的古籍電子文獻系統作了初步說明，發現大部份的系統皆是提供使用者閱覽，並經專家、學者整理過之古籍文件內容，但是其整理過程中仍舊依研究者個人意識、師承、學術深淺…等背景，將大量紙本資料逐本逐句的翻檢出不同之處，不僅耗時費力，亦提高錯誤的機率。

經測試結果顯示，系統使用自動校勘的技術方法是可行的，不僅可以快速找到各古籍版本間之異同，也能夠對不同版本進行註釋，對於推廣、發揚、甚至教學研究，皆有極大助益。

總結來說，本研究主要的貢獻有：

- 1.完成古籍文件半自動式的校勘系統，將「對校法」、「本校法」及「他校法」的概念納入於系統中，並以直覺的方式提供項目點選，提供便利的平台讓校勘研究者可以更快速的獲得校勘結果。
- 2.可對校勘結果加入不同見解，並且透過網際網路分享給其他使用者，達到知識分享的理念，使古籍文件內容的詮釋能獲得更具深度與廣度的迴響，使之成為可靠度較高的電子古籍文獻。
- 3.提供使用者可以為古籍文件建立內容、段落及註釋字詞之功能，使其他使用者可納為參考，並與文件內容產生關聯，讓文件內容融入更多的想法。
- 4.可透過系統中「輔助工具」功能自行建立字（詞）庫，取得同義字詞來進行校勘結果的擴展。

5.2 建議

總結目前的研究心得與操作結果，發現系統尚有許多空間與議題可提供後續研究者進一步的研究與修正，未來發展與應用可朝以下方向進行：

1.結合圖書館檢索系統

將校勘比對的概念運用在學術資料的檢索，讓使用者可以在圖書館的檢索系統中，有可以快速比對不同版本書籍內容差異之選項，縮短搜

尋所需版本的時間。

2. 輔助工具的增加與精進

現有的輔助工具雖可暫時滿足系統簡單的判斷，但面對博大精深的古籍文件，字詞庫仍顯不足，若能提供更多的字詞庫選項，將更能清楚標出文件中具有時代意義的關鍵字（某些用語、用詞只有在特定時代出現，如古代地名、官職或避諱字庫等）。

3. 多種檔案格式的上傳

本研究只以純文字檔(.txt)上傳至系統，然後轉成網頁文件進行校勘與註釋的處理，雖說純文字檔案為較基礎與普遍的檔案格式，未來若能於系統中納入其他上傳格式，定能增加系統的便利性，使系統功能更趨完整並發揮更大效益。

4. 加入數位學習的功能

本研究的目的是除了提供使用者尋找資訊的便利性外，也讓使用者學習古文能更快速且上手，雖然目前已有許多文獻肯定註釋工具對學習上的助益，但因投入時間不足的關係未能將數位學習相關功能建置完成未來完成數位學習環境的建置更能顯現系統對使用者的助益。

以上已面臨但仍無法解決的問題，有待未來持續研究努力。

參考文獻

1. 互動百科，古籍，<http://www.hudong.com/wiki/古籍#13>, available at 2012.6.14.
2. 曹之，中國古籍版本學，洪業文化事業有限公司，台北，1994。
3. 顧力仁，「資訊時代圖書館中國古籍組織與利用之探討」，國立臺灣大學，博士論文，2000。
4. 李致忠，古籍版本知識500問，北京圖書館出版社，北京，2001。
5. 謝清俊、林晰，中央研究院古籍全文資料庫的發展概要，http://cdp.sinica.edu.tw/paper/1997/19970301_1.htm, available at 2011.11.20.
6. 中央研究院，中央研究院漢籍電子文獻介紹，<http://hanji.sinica.edu.tw/index.html>, available at 2011.10.21.
7. 郝淑東、張亮、馮睿，「古籍數字化的發展概述」，情報探索，2007年第7期，2007。
8. 程千帆、徐有富，校讎廣義·校勘篇，齊魯書店，山東，1998。
9. 李勝文，「古籍版本與校勘」，荷澤師專學報，第24卷1期，2002。
10. 蔡聚鎖，「圖書情報學視野中的新約校勘法形成淺析」，圖書館論壇，2010年第8期，2010。
11. 胡適，「元典章校補釋例」序，校勘學釋例，1997年7月，1997。
12. 管錫華，漢語古籍校勘學，巴蜀書店，成都，2003。
13. 陳垣，校勘學釋例，上海書店出版社，上海，1997。
14. 洪湛侯，中國文獻學新編，杭州大學出版社，浙江，1994。
15. 周學文、江荻，「元朝秘史的計算機自動校勘方法」，語言文字應用，2007年第3期，2007。
16. 常娥、侯漢清、曹玲，「古籍自動校勘的研究和實現」，中文信息學報，第21卷第2期，2007。
17. 維基百科，PHP, <http://zh.wikipedia.org/wiki/PHP>, available at 2011.10.30.
18. PHP 官方網站，PHP Usage Stats, <http://www.php.net/usage.php>, available at 2011.10.30.
19. 凱文瑞克，PHP5 與 MySQL5 入門學習指南，旗標出版，台北，2006。
20. 維基百科，MySQL, <http://zh.wikipedia.org/wiki/MySQL>, available at 2011.10.30.

21. Mark Maslakowski 著，自學 MySQL 21 天課程，陳清豪、袁琿譯，博碩文化，台北，2001。
22. MySQLPress 著，MySQL 管理者手札-跟 MySQL AB 徹底研究 MySQL，范綱志譯，上奇科技，台北，1995。
23. 林易賢，「文件注釋系統設計與開發之研究」，國立交通大學，碩士論文，2011。
24. 謝清俊、莊德明，古籍校讀工具「中文文獻處理系統」的設計，中國古籍整理研究出版現代化國際會議，北京，1995。
25. 謝清俊、陳昭珍、莊德明、周亞民，電子佛典中處理中文版本的方法，中央研究院資訊科學研究所，新竹，1994。
26. 莊德明，以《心經》為例說明如何利用計算機處理佛經的多版本，中央研究院資訊科學研究所，新竹，1994。
27. R. Kawase, E. Herder, and W. Nejdl, "A comparison of paper-based and online annotations in the workplace," EC-TEL, pp. 240–253, Proceedings of the 4th European Conference on Technology Enhanced Learning. Springer-Verlag, 2009.
28. PL Patrick Rau, et al. "Developing web annotation tools for learners and instructors," Interacting with Computers, Volume 16, Issue 2, Pages 163–181, British Computer Society, April 2004.

