

古籍資料庫之人工標註與自動 標註功能使用分析

The Usage Analysis between Manual and Automatic
Annotation Function of Sinology Database

卓珈郁

Chia-Yu Cho

國立臺灣師範大學圖書資訊學研究所研究生

Graduate Student

Graduate Institute of Library and Information Studies

National Taiwan Normal University

【摘要 Abstract】

過往數位人文研究多由內容探勘層面，經由分析文本內容，提供研究者對於內容比對的基本分析資訊，較缺乏由使用者閱讀行為角度提供輔助閱讀工具。古籍內容的標註可以增加對於文本內容的詮釋與補充敘述，產生概念間的連結，亦可與其他學習者分享，形成知識交流，對於提升閱讀理解具有相當助益。

本研究設計是以建構在閱讀古籍特性所開發的兩種標註系統功能，分析 30 位受測者使用自動標註與人工標註兩種介面閱讀古籍，在閱讀成效、科技接受度與閱讀感受的差異。研究結果顯示：（1）使用不同標註系統對於閱讀理解成效並沒有顯著差異，但人工標註的科技接受度有用性明顯高於自動標註；（2）受測者對於自動標註的「添加新詞」與「查詢詞彙」功能，以及對人工標註的「查詢詞彙」與「新增標註」功能，有比較高的滿意度；（3）受測者認為自動標註的問題是自動斷詞不精確，人工標註是標註操作較為不易。最終對於古籍資料庫標註功能提出改善建

議，提供古籍資料庫發展系統功能之參考。

Previous studies have shown that digital humanities studies are mostly based on content exploration by analyzing text content, and lack of auxiliary reading tools from the perspective of user reading behavior. The annotation of ancient books can increase the interpretation and supplementary of the content, create a link between concepts, and share with other learners to form knowledge exchange, which is quite helpful for improving reading comprehension.

This study explores the functions of two annotation systems developed in ancient book database. 30 participants recruited for the study were those with humanistic background. This study analyzes the effects of reading ancient books on automatic and manual annotation system for reading effectiveness, technology acceptance and reading experience. The results of the study show that: (1) There is no significant difference in the effectiveness of reading comprehension using different annotation systems, but the usefulness of manual annotation system is significantly higher than automatic system; (2) the subjects have higher satisfaction with the "add new words" and "query vocabulary" functions of automatic annotation system, as well as the "query vocabulary" and "new labeling" functions of manual annotation system; (3) The subjects believe that the problem of automatic annotation is that automatic labelling vocabularies are inaccurate, and the manual annotation is not easy to label.

【關鍵詞 Keywords】

使用者調查；全文檢索；數位人文；介面設計；古籍資料庫
User Survey; Full-text Search; Digital Humanities; Interface Design; Ancient Book Database

壹、前言

數位掃描技術讓過往因保存考量，必須束之高閣的古籍文獻，藉由數位影像轉製，讓讀者得以親近文本內容；並藉由建置資料庫方式，提供研究者在浩瀚古籍中，檢索找尋與研究議題相關的材料。但資訊技術只是「工具」，研究工作的主導者還是「人」。即使資訊科技學

者與漢學研究者，皆期待未來的目標是電腦能更接近人類心靈的活動，但在此之前，仍必須經由漢學研究者的主導，建構更適合輔助漢學研究的數位研究環境，讓電腦在漢學研究的領域裡，具備分析大量文本功能、協助探查新研究主題，進而創造新研究成果的能力（羅鳳珠，1997；黃寬重，2012；項潔，2014）。

在數位時代，數位典藏所彙集的大量全文資料可作為研究的「新材料」，而數位人文研究應用資料探勘技術，作為人文學研究的「新工具」，能夠快速地從數百萬筆文獻史料中處理、分析，進而找出關連和連結；除了可以迅速地從大量文本中找出研究所需的「關鍵詞」分佈狀況外，甚至能從中發現大量資料內部隱含的關係，發掘出傳統人文學研究單靠人工無法處理的「新問題」，讓數位典藏資源不只是資訊來源，更晉升為積極的研究助理（項潔，2014）。數位人文研究的核心並非僅著重於資訊技術的發展，更重要的是能以使用者需求層面開發研究輔助工具，過往數位人文研究不乏由內容探勘層面，經由分析文本內容，提供研究者對於內容比對的基本分析資訊，資訊工具所分析的是文本內容，較缺乏由使用者閱讀行為角度探討輔助閱讀工具。

閱讀和理解古籍內容對於多數人而言，並不是一件容易的事，故而批注一直是中國文學鑑賞與評論的重要形式和有效理解內容的方法，其功能如同今日在文本進行的標註（annotation）內容。標註是一種為了評論或解釋，而增加的注解（American Dictionary of the English Language, 2017）。使用者透過標註將文本內容需要標示之處，留下閱讀紀錄，這些標註行為可以增加對於文本內容的詮釋與補充敘述，甚至可產生概念之間的連結，能為文本增添新的內容價值（Marshall, 1998）。標註內容不只是對於個人學習有強化記憶與理解的功能，這些標註的閱讀紀錄，亦可以與其他學習者分享，提供其他學習者吸收別人標註所闡述的閱讀知識，形成學習分享，對於提升閱讀理解相當有助益（陳志銘、韋祿恩、吳志豪，2011）。

標註從以往利用筆於紙本進行劃記或補充資訊的註記方式，開始發展到線上數位閱讀標註系統，透過線上數位閱讀標註系統支援數位文本閱讀時，可以進行的閱讀標註類型更廣，也能利用圖片、聲音及影片等新的媒體形式從事標註。實證研究顯示使用數位閱讀標註系統之實驗組學習者，對於閱讀理解與閱讀策略之運用，會優於採用紙本標註之控制組學習者（Chen & Chen, 2014）。但針對古籍文本而言，哪種標註方式對於使用者具有實質效益？使用者對於線上標註功能的

接受情況為何？為探究前述問題，本研究希望藉由分析不同標註功能的古籍文本閱讀介面，探討使用者使用自動或人工標註功能閱讀古籍，在閱讀成效、科技接受度與閱讀感受的差異，期許研究結果有助於古籍全文系統建置貼近使用需求的輔助閱讀功能。

具體而言，本研究目的在於：

- 一、探討閱讀者使用不同的古籍全文標註模式（自動標註與人工標註）對其閱讀理解、科技接受度與系統功能認可程度是否存在差異。
- 二、探討使用者對於兩種古籍線上標註系統（自動標註與人工標註）功能的看法，瞭解使用者對於不同標註功能的使用經驗及需求。

基於上述研究目的，本研究欲探究之研究問題包括：

- 一、不同的標註模式對於閱讀者的閱讀理解是否有差異？
- 二、不同的標註模式對於閱讀者的科技接受度是否有差異？
- 三、閱讀者對於不同標註模式的系統功能認可程度是否有差異？
- 四、使用者對於自動標註系統功能的使用經驗、操作問題以及新增功能需求有哪些？
- 五、使用者對於人工標註系統功能的使用經驗、操作問題以及新增功能需求有哪些？

貳、文獻探討

一、閱讀標註與標註行為

標註是鑑賞和批評的重要形式，也是一種有效的閱讀方法與策略。閱讀標註則是在閱讀過程中，於文章空白處寫下自己的心得、體會、評語或疑問，藉此幫助理解和評論文章。學習者會透過標註將文本內容需要標示之處，留下閱讀紀錄，標註方式可能是劃線、註記符號、註解文字或是寫下心得，這些標註行為可以增加對於文本內容的詮釋與補充敘述，甚至可產生概念之間的連結，能為文本增添新的內容價值（Marshall, 1998）。標註內容不只是對於個人反覆學習有強化記憶與理解的功能，這些標註的閱讀紀錄，亦可以與其他學習者分享，提供其他學習者藉由已標示之重點，吸收別人標註所闡述的閱讀知識，在學習過程中形成學習經驗分享，對於提升閱讀理解有相當助益（陳志銘、韋祿恩、吳志豪，2011；陳勇汀，2011）。

研究發現使用者喜歡透過不同的工具和顏色（如鉛筆、原子筆、螢光筆等）進行標註，而標註位置的不同，也可能有不同的功能，例

如在文本中簡短的螢光筆或圈選字詞，多是為了增強記憶或是給自己提示標註位置，但若是大範圍就有可能是追溯閱讀困難段落的一種方式。通常，個人標註大多被用來記憶、思考或瞭解文本，因此，當個人進行學習或是產生批判性思考時，會容易產生閱讀標註行為。故而，標註可以被視為主動性閱讀行為，當產生閱讀標註行為時，有助於提升閱讀的思考層次，也能提升閱讀的學習成效與理解（Adler & van Doren, 1972; Marshall, 1997; Hasen, 2006; Cabanac, Chevalier, Chrisment, & Julien, 2007）。

數位時代來臨，標註從以往以紙本利用筆或標籤方式，進行劃記或補充資訊的註記方式，開始發展到線上數位閱讀標註系統，在中文環境也發展出一些閱讀標註系統，如：KALS（陳勇汀，2011）、CRAS-RAIDS（Chen & Chen, 2014）等。線上標註系統在閱讀數位文本時，可以進行的閱讀標註類型更廣，也能在更多新的媒體形式上從事標註。研究顯示使用數位標註系統之實驗組學習者，對於閱讀理解與閱讀策略之運用，會優於採用紙本標註之控制組學習者（Chen & Chen, 2014）。線上標註系統是將紙本的標註轉換為使用電腦標註，並利用資訊工具特性，增加標註功能。Wolfe（2002）指出線上閱讀標註系統的功能，包括：輸入（input）、介面（interface）、文本（basetext）、定位（anchor）、儲存（storage）、檢索和篩選（searching and filtering）及其他特殊（specialized behaviors）功能等。陳勇汀（2011）闡述線上標註系統重要的功能為：

- （一）輸入：標註工具能接受的輸入類型，包括鍵盤輸入的文字型標註，麥克風輸入的聲音形式標註，以滑鼠、繪圖板或是電子筆等工具之繪圖形式標註。
- （二）標註顯示介面：標註的內容可以從主要文本上另開視窗顯示，在多人閱讀環境中，為了能分享他人的標註，最好能夠支援討論回應，使用者較能輕易瀏覽他人對於標註的回應。
- （三）被標註的文字類型：可能是獨立的文字文件、網頁、聲音檔、影片檔、電子文件或是對於標註的標註。
- （四）標註範圍的定位：允許使用者標註的範圍，可能是在一段文字、指定段落、整篇文章或獨立的像素（pixel）位置。
- （五）儲存位置：儲存位置可以分成兩種，一種是與主要文本儲存在相同的位置；另一種則是將文本跟標註分別獨立儲存。如果是分開儲存在資料庫中，標註內容可擁有更多彈性應用，能用於檢索、資訊過濾與合作分享。

(六) 檢索與過濾：讀者可以透過標註的各項資訊找到已經儲存的標註。

(七) 其他特殊功能：例如提供自動通知功能，當文件標註被更新時，系統可自動寄發電子郵件告知原標註者或作者標註有更新的訊息。

前述線上標註功能相較於傳統紙本標註，可提供更多的資訊分享與促進閱讀的成效，尤其合作標註的功能可累積標註內容，提升標註知識的完整度。Sato、Goto、Kimura 與 Maeda (2016) 將合作標註用於歷史文獻，支援不同研究者對於同一文本進行標註，讓人文學者可以藉由協作方式累積知識。

線上合作標註方式，除了可提供知識分享和促進研究的便利性之外，Renzel、Lottko 與 Klamma (2010) 的研究發現採線上合作標註的方式，可以激發創新的研究思維，發展出更有助於研究社群的文本語境。但不同標註者產生的標註內容可能是不同思維的角度，如果將合作標註的內容，增加標註分類或是輔助加上閱讀者的評分機制，可以提升對於標註內容品質的過濾，實證研究證明會更有助於閱讀理解成效 (Jan, Chen, & Huang, 2016; Chen & You, 2018)。

將標註方式用於古籍的閱讀由來已久，由學者對於古籍批注可協助後人對於文本內容的理解。然而時空轉異，並非所有古籍均有後人批注，或其批注時間已久，對於今人閱讀也產生閱讀理解有隔閡。因此，在資訊時代藉由線上標註系統，促進現代人對於古籍的學習。

Chen 與 Tsay (2017) 提出一種新穎的合作標註系統，將文本、圖片、語音和視訊四種類型標註，嵌入 HTML 網頁中，讓人文學者能夠在網頁上協作並添加和管理標註內容，這種協作合作標註以及提供討論的機制，對於人文學者研究古籍有很大的幫助。

Chen、Chen 與 Liu (2018) 藉由 Linked Data 的概念匯集來自不同資料庫的資源，並加以整合後，替文本進行自動註解，發展出古籍文本的「文本自動標註系統」，並比較「文本自動標註系統」與「MARKUS 文本半自動標註系統」在支援人文學者進行文本資料解讀之閱讀成效與科技接受度是否具有顯著差異。研究結果顯示文本自動標註系統閱讀介面，比較簡潔明瞭，比較兩個系統類似功能後發現，文本自動標註系統在查詢詞彙功能、連結到來源網站功能及新增標註功能，較為直覺易用，但半自動標註系統可供使用者分享標註內容的優點，而受到使用者青睞。

前述研究顯示，當文史研究者使用自動比對網路相關知識庫產生

之古籍內容自動標註時，尚有對於標註知識鏈結資料庫來源與標註內容深度的質疑。因此，基於前述文獻分析基礎，本研究調整 Chen、Chen 與 Liu（2018）的文本標註方式，由古籍閱讀特性為前提，將線上標註系統改為資料庫提供古籍全文自動標註，以及由使用者進行人工標註的兩種不同標註模式，評估何者較能符合人文研究者的閱讀需求。

二、研究者使用古籍資料庫相關研究

數位時代讓人文研究在短短的三十年內產生巨大的變化，從資料的取得方式到研究的進行，已經很難在沒有網路的環境裡進行。過往人文學者是依靠閱讀大量的紙本文獻耙梳整理學識成果，如果在資訊時代裡，面對浩瀚資訊，仍依循從前的研究方式，可能窮其力而難以盡閱文獻。因此，如何以資訊技術幫助人文學者進行研究，是近年備受關注的議題。

尤其，學術研究需要時時刻刻關注當下的趨勢變化，具備數位資源的使用能力已是不可抵擋的趨勢，現今研究者是否具備數位資源的掌握能力將影響其對於新知的掌握力（黃寬重，2012）。然而，人文學者進行研究時，仍偏重使用傳統的研究工具，如紙本文獻及工具書，對於資訊資源的接受度常常不如自然科學領域及社會科學領域研究者。Stone（1982）曾指出人文學者大都排斥電腦科技，但對於電腦的使用，仍有隨時間逐漸增加的趨勢。Horner 與 Thirlwall（1988）調查美國大學各領域學者使用電子資料庫的情形，也顯示人文學者較不熱衷於使用電子資料庫，而且比較不清楚如何進行文獻的檢索（轉引自吳明德、黃文琪、陳世娟，2006）。Wiberley（1991）訪談資深的人文學者，發現他們大都是獨立進行研究，不輕易依賴館員查找資料；並且會利用電子資料庫查詢資料的學者都是為了省卻交通時間，而非善用資料庫的檢索功能。

人文學者也有因自身參與資料庫的檢索調查，提出使用者角度的看法。王桂蘭（2007）就認為中研院的「二十五史資料庫」在文字上有缺字及錯字的問題，建議使用者應另行核對紙本文獻，以提高引用的正確率，而異體字的問題應提供學者造字檔使用；檢索上則建議資料庫應建立人名權威檔，並且增加古籍中附圖的圖檔，使古籍資料庫更完整。陳金木（2008）進行「四部叢刊電子全文檢索版」的檢驗，提出了幾點使用者應思考的問題：資料庫是否選擇最好的古籍版本？資料庫建構是否運用校對版本的處理技術？機讀過程是否有失真誤讀？

對於古籍資資料的運用，郭明芳（2011）根據使用中研院「古籍全文資料庫」的經驗，提出四項改進建議：

- （一）檢索結果呈現希望增加查詢結果排序選擇，並增加檢索結果篩選功能；
- （二）網路使用希望能免除流量限制，才能在線上大量翻閱資料；
- （三）圖像資料希望能完整呈現古籍圖表，而非只有文字內容；
- （四）功能鍵應清楚標示按鈕功能，才不會讓使用者誤操作。

廖益賢（2012）也針對「文心雕龍」資料庫進行檢索，提出古籍資料庫功能改善的意見：

- （一）資料庫應同時提供各家的注釋；
- （二）應提供版本相關資料，包含版本的序言、作者小傳、校勘記、目錄、校記等；
- （三）異體字的問題，可設置選取異體字的功能，或有圖文加以對照，若是採用通用字則須加以標示；
- （四）資料庫應能預測使用者需求，並提供使用者表達意見的管道。

此外，在資訊相關領域的研究，通常採用調查研究法分析人文學者使用資料庫之情形與問題，陳燦珠（1998）曾以觀察、訪談及問卷法，調查國文系學生使用中研院「古籍全文資料庫」文字版和 WWW 版的情形，發現不論是哪種版本，當使用者遇到困難而利用系統的求助功能時，常會遇到畫面提示不足、系統指引過於複雜或是用詞抽象難懂等問題。而系統畫面所設置的各連結按鈕未標示功能，也讓使用者感到困惑，受訪學生希望能增加跨資料庫檢索及利用資料從屬關係檢索的功能。

吳明德、黃文琪與陳世娟（2006）採用深度訪談法分析學者使用資料庫的目的是為研究及教學，偏重人物查詢、瞭解觀念發展和以確定資料有無為目的，檢索方式主要以關鍵字查詢為主，以瀏覽功能為輔。受測者認為關鍵字查詢比進階查詢可尋獲更多相關資料，瀏覽功能則主要用於確認發展脈絡。

綜合上述研究觀點，影響全文資料庫的使用因素有資料庫收錄內容與檢索功能問題。在資料庫收錄內容上，人文學者重視資料庫收錄的版本、有否注釋、內容完整性以及文字的正确性；在檢索功能上，系統有檢索提示不足、無權威控制、查詢結果呈現不夠親和、功能鍵標示、從屬關係瀏覽、相關連結以及使用流量管制等問題。對於那些使用資料庫的人文學者而言，自然是關心系統界面的功能設計、在乎系統功能的易用性、希望降低操作過程的困難等。因此，進行使用者

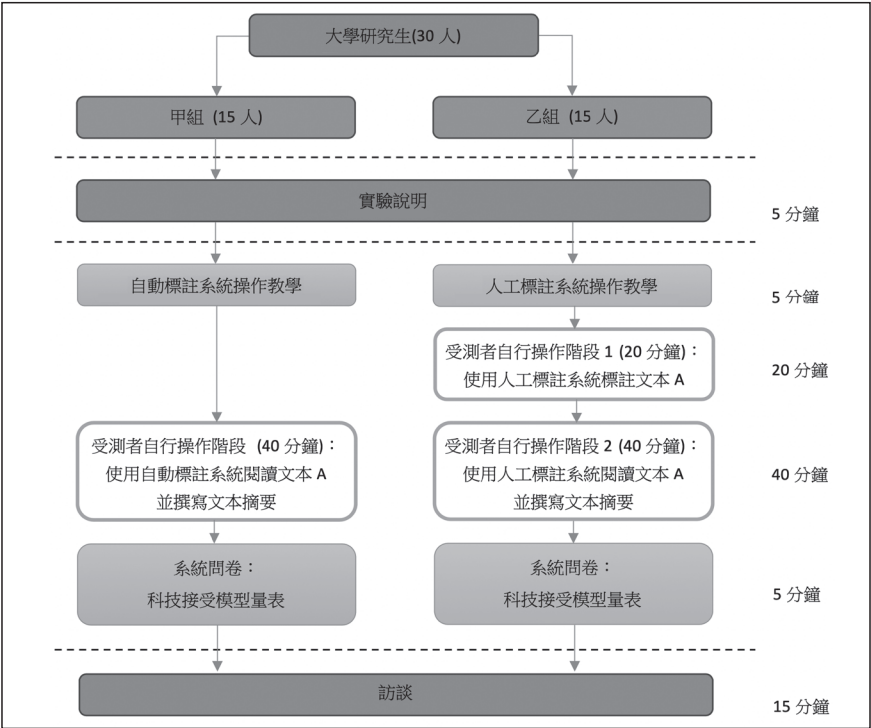
意見調查，顯然是在開發古籍系統介面功能之際，非常值得探討的議題。

參、研究設計與實施

一、研究設計

本研究目的為探究自動標註與人工標註對於古籍使用者之閱讀理解、科技接受度之影響以及使用感受差異。研究對象設定為具備解讀明代古籍文本並有撰寫讀後摘要能力的文史背景大學生和研究生，以張貼公告的方式，招募有意願配合實驗的北部大學中國文學、歷史學相關系所學生 30 人，進行文本標註與閱讀之實驗測試。

本研究採用實驗研究法，將受測者均分為甲、乙兩組，其中一組採用自動標註系統，另一組則採用人工標註系統，解讀相同內容之明人文集文本，實驗時間共 90 分鐘，流程如圖 1 所示：



實驗開始會先進行施測過程與系統使用操作說明，讓兩組受測者具備相同之系統操作能力。為了瞭解受測者是否能有效解讀文本內容，在標註系統上閱讀完畢後隨即撰寫讀後摘要，藉此分析閱讀理解成效差異。

在受測者完成文本閱讀與標註後，接著進行科技接受模型量表填寫，由量表問卷瞭解受測者對於兩種不同標註系統的認知有用性及易用性差異，並填答對於系統功能滿意度問卷。填答問卷後採用半結構式訪談，瞭解受測者對於兩種標註系統的使用意見，訪談問題重點在於瞭解：對於受測系統各項功能的意見、系統操作過程發現的問題、對於標註系統改善的建議。

本研究設計雖以比較人工標註和自動標註兩類系統之使用意見，但有可能受限於實驗階段系統功能之完善度與完整度，不能完全代表兩類標註系統用於解讀古籍的優劣，所得結論亦以採用工具為限，而無法概括論斷所有標註系統。

二、研究工具

(一) 標註系統

本研究於系統發展之際進行兩種標註系統功能之比較，目的在於提供系統功能設計參考。比較之兩種標註系統皆為國立政治大學陳志銘教授研究團隊所發展，系統功能說明如下：

1. 自動標註系統採用「明人文集自動標註系統」

自動標註系統是以國立政治大學受教育部委託研究「國家圖書館古籍數位人文平台建置計畫」為基礎所發展之古籍數位人文研究平台的自動標註功能，藉由 Linked Data 的概念匯集來自不同資料庫的資源，並加以整合後，替文本進行自動注解，讓使用者在解讀文本時能夠即時參照其他資料庫的資源，並提供文本標註閱讀介面，以利於人文學者進行資料的解讀。自動標註功能主要有：（1）系統鏈結比對資料庫會自動呈現有意義詞彙的不同顏色標註；（2）針對使用者感興趣詞彙可以選擇連結來源資料庫瀏覽內容。



圖 2 自動標註系統呈現標註概念

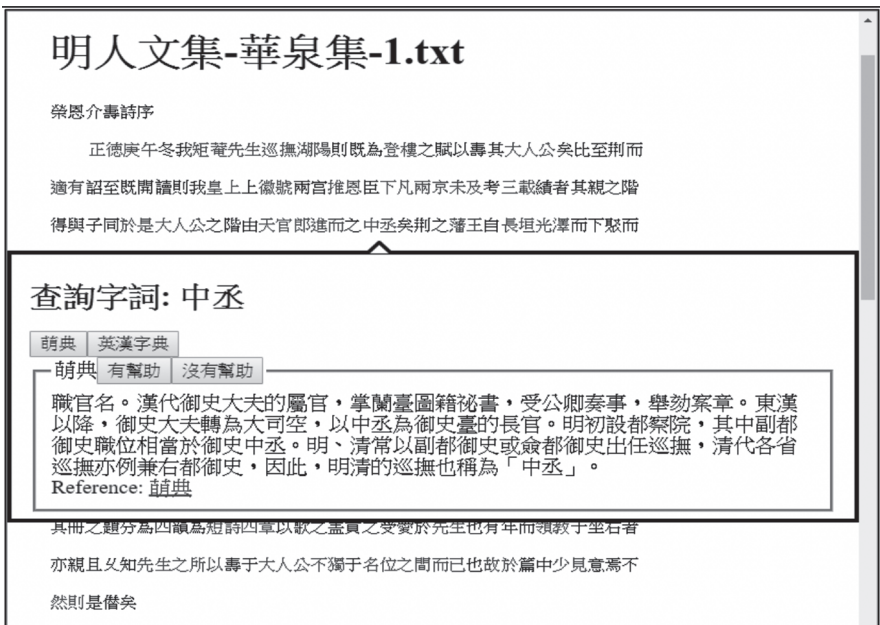


圖 3 自動標註系統呈現鏈結之標註內容參考

2. 人工標註系統採用「閱讀知識標註學習系統（KALS）的人工標註工具」

KALS 標註工具是一種附掛在 HTML 網頁上的標註工具，可以剖析、調整 HTML 的內文，將原本靜態的網頁加上「標註」的功能。使用者可透過 KALS 標註工具，為網頁上的文字新增標註，並填寫補充的筆記，也可以瀏覽其他人的標註、針對其他人的標註進行討論，或是將其他人的標註加入「喜愛清單」，給予鼓勵。本研究使用之人工標註系統係依本研究需求將 KALS 標註工具介面，改成針對明人文本特性的標註類型與標註來源連結，調整後標註工具介面功能有：（1）使用者可於介面選擇在文本進行人工標註，選擇標註類型；（2）根據標註需求選擇參考資料進行內容標註。



圖 4 人工標註系統提供之標註類型

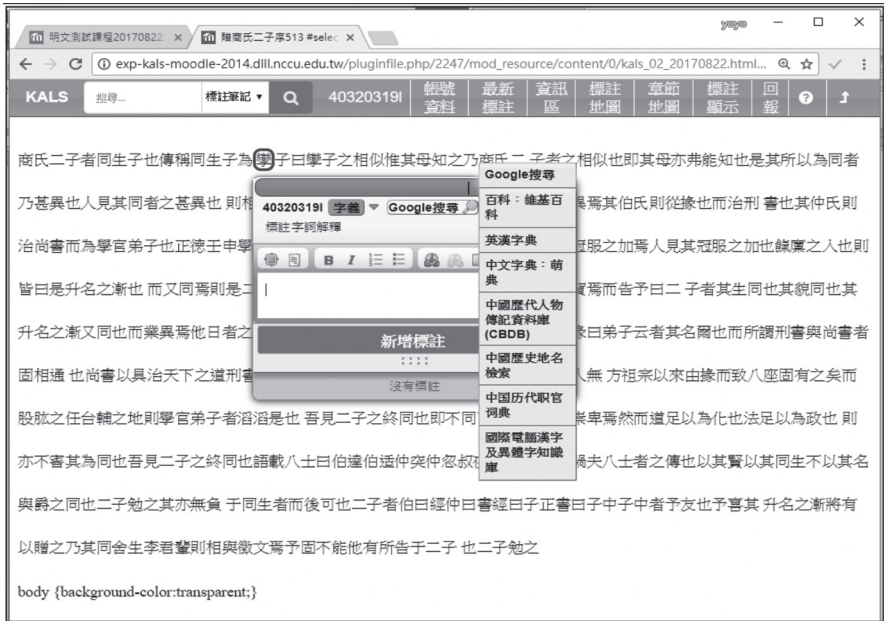


圖 5 人工標註系統參考來源資料進行標註

(二) 科技接受模型量表

本研究希望藉由科技接受模型，從使用者的認知有用性與認知易用性兩個構面，瞭解使用者對於使用線上標註系統工具的態度及意願。參考 Hwang、Yang 與 Wang（2013）所編製之科技接受模型量表，並修改語句以符合本研究需求，計分方式採用李克特五點量表。此一量表包含兩個構面，分別為「系統認知有用性」，共 6 題；「系統認知易用性」，共 6 題，總計 12 題。在量表信度上，「系統認知有用性」構面之 Cronbach's α 值係數為 0.95；「系統認知易用性」構面之 Cronbach's α 值係數為 0.94，皆具有良好的信度。此外為了瞭解受測者的使用過程感受，增加「系統使用感受」題組，共 4 題。

三、資料分析

(一) 閱讀理解分析

兩組受測者針對文本閱讀後撰寫摘要，摘要為閱讀理解之具體成果，由兩位中文系教師分別依其內容完整度給予 1 至 10 分後取其平均數。將閱讀理解分數進行描述統計，以分數區間呈現人數及其所占總

人數百分比，並計算整體之平均數與標準差，整理出閱讀理解分布情形。對於不同標註系統受測者閱讀理解之差異，以獨立樣本 t 檢定，檢驗兩組閱讀理解分數是否達顯著差異。

（二）科技接受度問卷量表資料分析

根據科技接受模型量表問卷，填答由非常不同意（1 分）、不同意（2 分）、沒意見（3 分）、同意（4 分）、非常同意（5 分）分為 5 等量表，分別計算各個題項之平均數與標準差，進一步以獨立樣本 t 檢定，檢驗使用不同標註系統之受測者對於科技接受度之差異，

（三）不同系統功能感受滿意度

根據不同標註系統受測者填答系統功能滿意度問卷，採用非常沒幫助（1 分）到非常有幫助（5 分）共 5 等量表計分，統計受測者認為自動標註系統和人工標註系統各項功能的滿意度平均數與標準差。

（四）訪談資料分析

研究者於訪談後將錄音檔謄錄為逐字稿，並確認繕打內容之正確性，反覆閱讀逐字稿，將訪談內容重點標示出核心概念，註明所有與研究目的及研究問題相關之語意概念編碼，將所有受訪逐字稿逐一標示概念編碼，最後將相同層面概念提出，作為整理文稿之基礎。接著，將訪談錄音文字紀錄加上研究面向註解，從中歸納出幾項重點意見，並將相同意見條理分析，節錄符合重點之訪談語句內容。

為忠於質性研究要求，本研究引錄受測者意見以逐字稿原樣呈現，採不同字體置於引號「」中，以示區別，最後以括號（）標示受測者代碼，以資明確引用出處。受測者代碼依使用之標註系統編號，第一個字碼為 A 代表使用自動標註系統，K 代表使用人工標註系統，後兩碼為受測者編序流水號。

肆、研究結果與分析

一、閱讀理解分析

（一）閱讀理解分數分布情形

受測者以標註系統閱讀古籍文本後，將閱讀所得撰寫摘要，摘要內容由兩位中文系教授分別評分後取平均數（評分滿分 10 分），分析

受測者閱讀理解平均數為 6.63 分，標準差 2.09。如以中間數 5 為界，理解分數達 5 分以上者，共有 25 人，分數分布較為集中在 5 分到 9 分之間（表 1）。

表 1
閱讀理解分數人數分布統計表

分數區間	人數 (n=30)	百分比 (%)	平均數	標準差
1-1.9 分	0	0	6.63	2.09
2-2.9 分	1	3.3		
3-3.9 分	2	6.7		
4-4.9 分	2	6.7		
5-5.9 分	4	13.3		
6-6.9 分	4	13.3		
7-7.9 分	5	16.7		
8-8.9 分	4	13.3		
9-9.9 分	8	26.7		
10 分	0	0		

(二) 不同標註系統閱讀理解分數之差異

比較不同標註系統受測者閱讀理解之差異，自動標註系統閱讀理解平均數為 6.27，人工標註系統為 7.00，檢驗兩組閱讀理解是否達顯著差異，以獨立樣本 t 檢定，當 $F=1.759$ ， $p=0.195>0.05$ ，在「假設變異數相等」時， $t=-0.958$ ， $p=0.346>0.05$ ，未達顯著差異，可知使用不同標註系統之閱讀理解分數並未達顯著差異（表 2）。

表 2
不同標註系統之閱讀理解分數差異分析表

系統	平均數 (M)	標準差 (SD)	t 值	p 值
自動標註	6.27	2.40	-0.958	0.346
人工標註	7.00	1.73		

二、科技接受度分析

(一) 科技接受度問卷分析

受測者填答科技接受度之問卷結果分析（表 3），呈現整體的易用性平均數（3.7）高於有用性平均數（3.0）和使用感受平均數（3.0）。

在認知有用性方面，受測者給予「4. 這樣的系統可以幫助我在需要時獲得有用的資訊」較為高分，而且認知歧異低，但對於「3. 這樣的系統所提供的標註功能讓我的文本解讀過程更為順暢」、「5. 這樣的系統可以讓我更正確的解讀文本內容」的平均數最低；代表受測者雖然系統認同功能有助於獲取資訊，但對於是否能有效解讀文本內容，仍然有些保留的態度。

在認知易用性方面，呈現「10. 操作過程對我來說沒有甚麼困難」之平均數最高，而「7. 在使用這個系統的過程中，並不需要花費過多的時間精力」平均數最低。可解釋受測者雖然認同系統操作能容易理解，但仍需要花時間學習和適應。

在使用感受方面，「14. 我能夠適應使用此系統解讀文本」為最高分，而「15. 使用此系統解讀文本符合我的閱讀習慣」與「16. 我會想繼續使用此系統功能來解讀文本」同列為最低分，代表受測者對於未來再次使用的意願是有遲疑的。

（二）不同標註系統之科技接受度差異

如果比較使用自動標註與人工標註系統之受測者對於科技接受度之差異，可發現在認知有用性的面向，人工標註系統受測者對於所有題項的平均數，皆高於使用自動標註的受測者，顯示受測者感受自動標註系統功能與解讀古籍需求仍有差距，使用者認為以人工標註方式閱讀古籍文本更為有用。

在認知易用性的面向上，不同標註系統受測者對於認知易用性的看法頗有出入，其中人工標註平均數高於自動標註的題項為「7. 我在這個系統的過程並不需要花費過多的時間精力」、「8. 本次使用此系統的活動對我而言是清楚且容易理解的」，其餘題項之平均數，皆為自動標註高於人工標註，顯示出受測者在系統易用性上，多數認為自動標註優於人工標註。

在使用感受的面向上，皆為自動標註系統之平均數高於人工標註系統。因此，根據平均數分析，呈現受測者認為人工標註系統的有用性高於自動標註系統，但易用性和使用感受，則是自動標註高於人工標註系統。

進一步將科技接受度各題項進行不同系統認知差異之顯著性比較，在「3. 這樣的系統所提供的標註功能讓我的文本解讀過程更為順暢」、「11. 我覺得這樣的系統介面很容易使用」、「15. 使用此系統

解讀文本符合我的閱讀習慣」此三項 p 值小於 0.05，在「6. 使用這樣的系統解讀文本比使用一般的數位典藏資料庫更有效果」此項 p 值小於 0.001，代表受測者對於不同標註系統中此四項認知有達到顯著差異（表 3）。

表 3
不同標註系統之科技接受度差異分析表

題目		整體 (n=30)		自動標註系統 (n=15)		人工標註系統 (n=15)		獨立樣本 <i>t</i> 檢定		
		平均數	標準差	平均數	標準差	平均數	標準差	<i>t</i> 值	<i>p</i> 值	
認知有用性	1. 我覺得使用這樣的系統讓文本的解讀更豐富	3.2	3.0	1.19	2.9	1.10	3.5	1.25	-1.398	0.173
	2. 我覺得使用這樣的系統對於我解讀文本很有幫助	2.9		1.08	2.6	1.12	3.3	0.96	-1.748	0.091
	3. 這樣的系統所提供的標註功能讓我的文本解讀過程更為順暢	2.7		1.16	2.2	0.86	3.1	1.25	-2.386	0.024*
	4. 這樣的系統可以幫助我在需要時獲得有用的資訊	3.4		0.96	3.1	0.92	3.6	0.99	-1.344	0.19
	5. 這樣的系統可以讓我更正確的解讀文本內容	2.7		1.02	2.4	1.12	3.0	0.85	-1.655	0.109
	6. 使用這樣的系統解讀文本比使用一般的數位典藏資料庫更有效果	3.2		1.18	2.5	0.92	3.9	0.99	-4.020	0.0004***

認知易用性	7. 我在使用這樣系統的過程並不需要花費過多的時間精力	2.8	3.7	1.07	2.4	0.91	3.1	1.13	-1.962	0.06
	8. 本次使用此系統的過程對我而言是清楚且容易理解的	3.7		1.06	3.3	1.18	4.0	0.85	-1.784	0.085
	9. 我很快便瞭解此系統的操作方式	4.1		0.97	4.2	1.01	4.1	0.96	0.370	0.714
	10. 操作過程對我來說沒有甚麼困難	4.2		0.95	4.3	1.05	4.0	0.85	0.960	0.345
	11. 我覺得這樣的系統介面很容易使用	3.7		1.15	4.1	1.25	3.3	0.88	2.197	0.036*
	12. 整體而言本次的系統是容易使用的	3.9		0.90	4.1	1.10	3.7	0.62	1.228	0.23
使用感受	13. 利用這種數位方式閱讀能符合我的研究習慣	3.0	3.0	1.15	3.1	1.13	2.9	1.19	0.631	0.533
	14. 我能夠適應使用此系統解讀文本	3.4		0.93	3.5	0.99	3.3	0.88	0.778	0.443
	15. 使用此系統解讀文本符合我的閱讀習慣	2.8		0.97	3.1	0.99	2.4	0.83	2.200	0.036*
	16. 我會想繼續使用此系統功能來解讀文本	2.8		1.07	2.8	1.01	2.7	1.16	0.167	0.868

* $p<.05$ ** $p<.01$ *** $p<.001$

三、不同系統功能感受滿意度

本研究所比較之不同系統功能滿意度是針對採用之研究工具特性，自動標註系統是藉由 Linked Data 的概念，系統鏈結比對不同資料

庫的資源，自動呈現有意義詞彙的不同顏色標註，替文本進行自動註解，讓使用者在解讀文本時能夠即時參照其他資料庫的資源，同時也可選擇連結來源資料庫瀏覽內容，並對來源內容進行評分，針對使用者感興趣詞彙若無斷詞標註則可以反白查詢或添加新詞。自動標註系統功能有：自動斷詞、自動標註、資料來源評分、連結到來源網站、查詢詞彙、添加新詞。人工標註系統透過 KALS 標註工具介面，使用者可為網頁上的文字新增標註，使用者可選擇標註類型，根據標註需求選擇參考資料進行內容標註，也可以填寫補充的筆記，或瀏覽其他人的標註、針對其他人的標註進行討論。人工標註系統功能有：標註分類、新增標註、編輯標註、查詢詞彙、標註地圖。

根據不同標註系統受測者填答系統功能滿意度問卷，採用非常有幫助（5分）、有幫助（4分）、沒意見（3分）、沒有幫助（2分）、非常沒幫助（1分）計分，統計出受測者認為自動標註系統最有幫助的功能是「添加新詞」（3.9）、「查詢詞彙」（3.5），最沒有幫助的功能是「自動斷詞」（2.5）、「連結到來源網站」（2.9）。在人工標註系統方面，受測者覺得最有幫助的功能是「查詢詞彙」（4.4）、「新增標註」（4.0），最沒有幫助的功能是「標註地圖」（3.1）（表4）。

表 4
不同標註系統功能認同度差異分析表

題目		非常沒幫助	沒有幫助	沒意見	有幫助	非常有幫助	平均數	標準差
自動標註系統 (n=15)	自動斷詞	0	10	2	3	0	2.5	0.83
	自動標註	0	5	3	7	0	3.1	0.92
	資料來源 評分	2	2	5	5	1	3.1	1.16
	連結到來源網站	2	4	3	5	1	2.9	1.22
	查詢詞彙	1	2	2	8	2	3.5	1.13
	添加新詞	0	1	3	8	3	3.9	0.83
人工標註系統 (n=15)	標註分類	1	1	4	6	3	3.6	1.12
	新增標註	0	1	2	8	4	4.0	0.85
	編輯標註	0	1	4	6	4	3.9	0.92
	查詢詞彙	0	0	2	5	8	4.4	0.74
	標註地圖	1	2	6	6	0	3.1	0.92

四、訪談紀錄分析

訪談問題聚焦於瞭解使用不同標註系統之受測者在使用過程中，對於系統功能設計優劣的看法，將訪談意見分為三層面分析，一是說明受測者針對使用的系統各項功能認為有用和沒有用的原因；二是對於使用的系統介面操作難易度看法；三是根據使用經驗提出未來新增功能或是改善建議。整理訪談意見如下：

(一) 自動標註系統

1. 自動標註各項功能使用經驗

受測者認為自動標註系統最有幫助的功能，主要是「添加新詞」和「查詢詞彙」。認為「添加新詞」有用的地方是能及時反饋給系統斷詞的建議，典型的看法如受測者 A06 的說明是：「有些詞彙其實是有意義的，不過系統並沒有把它斷出來，像是職官名一般人比較不熟悉，……或是地名有古今的差異，所以需要讓使用者可以添加新詞。」不過由於此功能是將受測者添加新詞的意見提供系統詞庫做參考，受測者在使用時較無法直接且明顯的感受到標註之後系統是否有紀錄，有可能導致重複操作，或是對標註的新詞正確性有疑慮，如受測者 A11 表示：「添加後他沒有立即的反應在資料上，我發現有些詞我好像重複添加了，……如果可以立刻有辦法辨識出，它這個詞我剛才已經標註過的話，我覺得這樣比較好。」所以，在標註新詞這部分建議可進一步讓使用者分辨出標註紀錄以及編輯修正的功能。

認為「查詢詞彙」有幫助的原因是當自動標註沒有標示出來，而閱讀理解有疑惑時，可以增加查詢連結，讓使用者能針對自己需要加強的部分，進行連結查詢和釋疑，故而能獲得受測者肯定。

而受測者認為最沒有幫助的系統功能，主要是「自動斷詞」與「連結到來源網站」。對於自動斷詞的功能，部分受測者認為系統提供此項設計的立意很好，但是由於機器比對標示，會產生標示過多或標示不正確問題，反而會干擾閱讀，甚至造成閱讀理解困難，需要調整自動斷詞所比對的詞庫來源與提供使用者參與修正的功能，才能讓自動斷詞的設計有助於輔助閱讀的效果。典型的意見如受測者 A02、A05 表示：「機器有時候會斷錯詞組，反而造成在解讀文本上的干擾，我覺得這個功能可能斷詞還需要進一步修正，才會有幫助」、「有些斷詞斷的不是我想要的，如果我覺得這個字詞沒幫助，希望可以讓藍色字變成黑色字那樣沒有標註的樣子，會更清楚一點」；至於「連結到

來源網站」，受測者認為因為資料庫來源不符合需求，所以對此功能的評價不高，還有部分受測者表示參考自動標註的註解內容已經足夠，並不需要進一步連結到來源網站，因此對於此一功能的態度保留。如受測者 A07 的看法是：「來源不是全然有幫助的，裡面有些資源我覺得不應該連結到這個詞典，就我使用上看文本的需求，比較會用到的只有萌典還可以。」

綜合前述分析，在目前自動標註系統提供的功能中，使用者認為「自動標註」和「查詢詞彙」有用的原因，是前者可藉由比對常用詞庫提供主動解釋，後者可讓使用者自主增加需要加強的解釋，都能促進使用者對於內容的閱讀理解，顯然使用者迫切需求的是對於文本內容釋義的功能。至於「自動斷詞」、「連結來源網站」、「添加新詞」等功能，較屬於輔助工具性質，能幫助使用者延伸閱讀或是與別人分享自己的知識，然而「資料來源評分」功能，對使用者而言似乎像是給系統回饋，顯然對於閱讀理解無直接關聯，而較不受使用者青睞。

2. 自動標註系統操作過程問題

對於使用自動標註系統的受測者認為系統操作的難易度，多數受測者表達系統操作簡單且直觀，沒有太多困惑和不清楚的地方。如受測者 A06 陳述意見：「我沒覺得特別不好操作的地方，……整體介面設計使用起來可以很快上手。」但也有部分受測者認為系統操作上仍有一些問題，例如：自動斷詞反而會妨礙正確判讀、添加新詞的效果不顯著，以及資料來源和呈現方式需要改善等，佐證受測者 A04 的意見：「因為有自動斷詞的問題，導致我在閱讀的時候，我認為這個字應該與前面那個字連結在一起，而它卻與後面的字相連時，我很難立即反應出原本我認為的那個詞的意思。」A12 則認為「添加新詞比較困難一點，因為系統應該可以設計成將添加出來的詞用其他顏色，或是另外一種方式標示出來，讓我知道這是我自己加的詞。」可知目前自動標註系統在整體操作感受上，沒有太大問題，但自動斷詞不精確反而變成自動標註的敗筆。

3. 希望自動標註未來增加的功能

(1) 增加句讀功能

有相當多受測者希望自動標註能在古籍全文自動標示句讀，幫助閱讀時更加通順解讀文意，引述受測者意見如下：「希望它可以有句讀功能，我們在看古書時會自己去點句讀，可是在數位文本上還不行，所有字句全都連在一起，有時候會很難閱讀，容易眼花，所以希望有個可以讓我自己去點它稍微分開一些，類似句讀的功能（A02）」。

這樣會比較符合古籍文本閱讀習慣。

(2) 改善資料庫來源

受測者認為目前自動標註所選擇的連結資料庫，僅有少部分有輔助文章解讀的功效，多數的資料庫在閱讀時所提供的資訊是沒有幫助的，因此認為應增加更多較為專業的資料庫，引述相關意見如下：「不管是大家熟悉的古文和不熟悉的古文，資料來源可以選比較學術的，即使它的詞彙艱深一點也沒關係，不需要像萌典那麼通俗性，萌典雖然也不錯，但在引經據典的部分，它的解釋就比較受限，所以我推薦要使用比較學術性的資料庫，提高解釋古籍內容的程度（A09）」。

(3) 提供白話翻譯功能與補充資料

受測者希望可以增加白話翻譯或是賞析類型的標註，作為閱讀時的輔助，引述受測者意見如下：「它可以像那個 Google，每次我們查詢一個詞，右邊不是都會有連結到小百科之類的，我覺得那樣蠻好的，比如說這篇文章是在描述送葬歌曲，它可以告訴我說輓詩是甚麼，那就可以在旁邊寫一些輓詩的由來，就是增加一些小百科，或是外面的連結資源等（A01）」。

(4) 區別標籤內容的差異

受測者表示在標註過程中，單一的標註顏色不容易分辨詞的屬性，希望標註時能以不同顏色顯示不同類別的標註。引述典型的意見為：「同一行的斷詞有沒有可能使用不同的顏色，這樣看下去我就知道這一行可能有人之前已經標過，或是可能是不同時期被標起來的東西，……斷成兩個字或三個字連接的部分不能一樣顏色，因為一樣顏色的話，我的視覺效果會覺得他是一整句（A14）」。

(5) 開放人工編輯

針對自動斷詞有時在文本的判讀上，會產生標註不夠精確的問題，受測者認為可以增加由人工添加標註或給予修正意見方式補強系統自動標註問題，將能使系統自動標註功能更加完善。引述相關建議如下：「我會希望針對斷句做很大的修正，或是在添加新詞的功能上，像是藉由人工添加來補足自動斷詞的錯誤（A04）」。

綜合上述分析，受測者認為自動標註系統操作容易，即使是第一次使用系統的使用者，也能很快的理解如何操作，目前受測者對於自動標註系統的意見偏向於對現有功能的改善，希望修正自動斷詞的正確性、能連結更有用的資料庫、能區別標註差異、增加人工編輯修正的功能。

(二) 人工標註系統

1. 人工標註各項功能使用經驗

使用人工標註系統之受測者認為，人工標註系統最有幫助的功能，主要是「查詢詞彙」、「新增標註」和「編輯標註」，尤其是「查詢詞彙」能幫助閱讀古文者針對需要解釋的詞彙進行超連結查詢，在閱讀古籍過程可以一邊閱讀一邊查，查到的結果也可以加入標註，只是希望目前提供查詢連結的資料庫，可以有更多的選擇性，或是能考慮使用者常用的資料庫增加新的連結。如受測者 K06 表示：「可以直接查詢是好的，不過如果可以的話，希望能手動加其他字典進去，因為我不習慣用萌典，我習慣用漢典。這次閱讀過程中，我比較常用萌典和 Google 搜尋，但是我有時候要查資料的時候不會用系統內附的資料庫，而是另外開一個網頁來查漢典。」

「新增標註」是將自己查詢所得或理解的知識加上標註內容與人分享，認為此項功能有用的意見是認為閱讀文言文往往有艱深文句，需要別人閱讀經驗的標註內容幫忙理解，如受測者 K12 表示：「查詢詞彙的功能提供很多找到的資源，可以去知道他的意思是什麼，然後一旦拿到那資源以後還要在新增標註的地方做一些筆記，才不會忘記說剛剛找到的是什麼」，亦即新增標註是查詢資源之後自然產生的必備功能。

至於「編輯標註」功能是提供閱讀過程對於別人已標註的內容增加評論或解釋，如同閱讀作筆記，除了可以看到別人作的筆記，甚至可以附上自己的筆記內容。多數受測者認為此項功能有助於閱讀理解，如受測者 KL10 表示：「因為他就像是做筆記一樣，你只要點上去你就可以看到你之前做的筆記，但那個過程用滑鼠動一動就可以了，我覺得閱讀古文的時候是需要去做這些註解的。」

而受測者認為最沒有幫助的系統功能，主要是「標註分類」、「標註地圖」，此兩項系統功能設計原意是在輔助使用者，可以掌握文本內容已經有的前人標註類型與整體分布，但顯然對於受測者而言重要的是在於閱讀過程對於文本內容的直接閱讀理解，提供此項輔助功能設計雖然有助於文本內容分析，但對於當下閱讀過程的理解並無太大實質意義。而目前標註分類專指性不足可能也是讓受測者感受不理想的原因。引述代表性意見如受測者 K07 的看法：「前人的標註可能會有我不認同的，我可能就不會標他，或是標註分類我也覺得沒太大的幫助，我站在中文專業角度上來說，我覺得這對我們來說沒有太大的意義，因為有些詞在我們專業中間就很明顯的，我們一看我們就知道

他是哪些詞，例如人名、地名、職官名這些是很明顯的，我們就不需要這個功能。」

整體而言，對於目前人工標註系統提供的功能中，「新增標註」、「編輯標註」和「查詢詞彙」對於受測者而言，可以查詢內容和增加標註，對於閱讀古籍內容有直接幫忙。但目前人工標註所設計之「標註分類」、「標註地圖」對於閱讀理解過程較無直接關聯，而較不受受測者認同。

2. 人工標註系統操作過程問題

使用人工標註系統的受測者對於系統操作的難易度，均屬於系統操作過程的操作問題，例如：在選擇查詢詞彙時的標示不容易操作、標註屬性的下拉式選單不容易點選、標註選錯要取消不容易、提供連結的資料庫不實用（認為古文不用查 YouTube、Google），較典型的意見如受測者 K06 所述：「標註時會習慣點螢幕，但是點錯了的時候，系統會自動框起來又跑出訊息，這時候要取消選取就比較麻煩，這是操作上比較困擾的地方」。K07 說明：「如果那個點選只能換成鼠標拖動直接點選的話會更方便，像是那個 MARK 筆直接劃上去就可以了，現在如果我點一個字我要點兩下，就是一個詞頭到最後一個字，標註設計應該要更容易去標示，不然我覺得這樣就很容易看錯字。」

也有受測者認為目前標註內容的深度不夠，僅是貼上查詢所得，這樣的筆記內容與自己透過查詢的內容是一樣的，所以才會讓受測者認為看到的內容與查詢所去不遠，何需此項新增功能，也意味著目前編輯標註的內容都屬於淺層內容標註，只是解釋性內容，並沒有系統使用者綜合知識所得的標註，未來要增加標註內容深度應該有賴於使用者將閱讀所得寫出有深度的筆記內容，而非僅是貼上查詢所得的文字。如受測者 K04 表示：「看別人的標註感覺與自己查詢的內容沒什麼差別。」

3. 希望人工標註未來增加的功能

(1) 增加句讀功能

古文缺乏標點斷句，是現代人閱讀古文最大的障礙，因此，有不少受測者希望系統能在古籍全文加上句讀，才能大幅降低閱讀古文的困難。引述受測者 K07 意見：「我覺得斷句比標註更重要，寧願對斷句有更精確的判斷，不管是專業人員或非專業人員來說，斷句斷錯的話，意思差別很大，而且中間會造成很大的理解誤區，如果能夠增加句讀功能會比較好一點，因為我們經常會碰到引了一段話，原本以為是這個意思，但是因為閱讀斷在不一樣的地方，理解就差之千里。」

(2) 增加有用的資料庫來源

人工標註所提供的查詢詞彙功能，目前內建提供的查詢資料庫來源，與受測者期許顯然有出入，不乏有受測者提出希望有其他資料庫來源，或是能貼近一般閱讀古籍內容需求的資料庫，例如：漢典比萌典好用、Google 缺乏學術性、中研院漢籍電子資料庫應該加進來，未來對於連結資料庫部分必須進一步諮詢漢學研究者意見，納入更多在閱讀上有用的連結。引述相關建議 K02：「希望能提供更多的資料庫來源，提供的資料庫最好是真正跟查詢明代古籍內容有關的資料庫，不是一般性的字辭典」。K13 同樣認為「中研院的漢籍電子資料庫的特色有原文，也會有一些夾註或是考證等，中研院有把它做好分類，只是中研院的問題是缺少一個標註，以及可查詢地圖的功能，那我覺得這是 KALS 可以補充中研院漢籍電子資料庫的地方。」

(3) 改善標註操作過程

在進行人工標註過程，標註詞彙的圈選操作不容易，經常會造成使用者標註不精確，往往在前後多選一個字之後，標註的意思就不正確，因此，受測者反映未來可以針對標註的點選方式提供更友善的設計，讓標註者操作比較方便。引述 K03 相關意見：「系統算是完善，但系統介面點按選項時，不要常跑掉就好。」

(4) 增加類似文本的連結

受測者提到可以在標註系統提供該篇古文的註解或是白話翻譯文本，藉由匯集古文內容的相關閱讀資料，促進對於古文內容的理解。此外，類似的建議是由延伸閱讀參考的角度，建議可以提供類似主題的相關文本推薦連結，或是類似個人化商品推薦的機制，主動推薦給使用者符合研究主題或是閱讀興趣的相關篇章內容。引述 K05 相關意見：「如果這篇文章有今人做過註解或翻譯的版本的話，也可以額外附加進來供人觀看，方便參考。」

(5) 增加對於標註內容的評價

由於標註是使用者閱讀過程之筆記，如同一般筆記會有記載詳簡不同或是筆記錯誤的問題，未來系統可以增加對於既有的前人標註評分與增加標註修正過程，藉由評分將有用標註優先顯示，可以降低標註資訊無用或是幫助不大的問題。引述 K07 的意見：「有一個地名是廣陵，我知道這是地名而且我用了歷代地名的辭典去做檢索，發現其實廣陵這地方一直都有，但是它會產生歷史的變化，……那就要考慮，這篇文章的產生時代是什麼時候，才能正確的判斷到底是現在的哪個地方，這系統沒有先提供這篇文章是哪個朝代寫的，……對一個古文

進行標註他應該是沒那麼簡單的，需要學者對內容有深層的認識才能標註正確，單純簡單的標註對研究者沒太大的幫忙。」

綜上所述，人工標註系統雖然能藉由人的閱讀標註修正自動標註的錯誤，但是標註詞彙的圈選操作不容易，是目前系統介面不易用的問題，人工標註與自動標註都有受測者提到需要有古文自動標示句讀的需求，並且較為在乎有深度閱讀標註的參考價值，也希望能增加延伸閱讀或是不同註解版本的加入。

伍、結論與建議

一、不同標註系統對於閱讀古籍理解程度之影響

分析參與受測之 30 位受測者的閱讀理解成績，集中在 5 分到 9 分之間，且 9 分以上者達 26.7%，顯示本實驗研究受測結果之閱讀理解成績以分布在高分為多。而人工標註之閱讀理解平均分數（7.00）略高於自動標註系統（6.27），但以獨立樣本 t 檢定則未達顯著差異。此與 Chen、Chen 與 Liu（2018）對於自動標註和半自動標註分析結果一致，亦即使用不同標註系統對於閱讀古籍的理解成效沒有顯著差異。

二、不同標註系統對於閱讀古籍之科技接受度影響

- （一）標註系統的科技接受度為易用性高於有用性：不同於 Chen、Chen 與 Liu（2018）分析自動和半自動標註方式，發現易用性和有用性的認知沒有差異。本研究在科技接受度分析結果呈現易用性平均數（3.7）高於有用性的平均數（3.0）和使用感受的平均數（3.0）。在「認知易用性」所有題項的表現，雖然認同系統操作容易理解，但仍表示需要花時間學習和適應。在「認知有用性」所有題項的平均數統計，得分最高的是「系統可以幫助我在需要時獲得有用的資訊」，但「系統可以讓我更正確的解讀文本內容」卻是得分最低。顯示受測者認為標註系統功能雖有助於獲取資訊，但對於是否能有效解讀文本內容，則抱持保留的態度。
- （二）人工標註的科技有用性較佳，自動標註的易用性較佳：分析使用自動標註與人工標註系統之受測者對於科技接受度之差異，結果呈現人工標註的有用性高於自動標註，代表自動標註系統功能與解讀古籍需求仍有差距，使用者認為以人工標註方式閱讀古籍文

本更為有用。但易用性和使用感受，則是自動標註高於人工標註，在 Chen、Chen 與 Liu（2018）的研究同樣顯示自動標註的易用性高於半自動標註系統，亦即受測者認為人工標註雖然較為有用，但操作過程則是自動標註感受較佳。印證受測者訪談意見，可知受訪者認為自動標註系統現有的「自動斷詞」和「連結來源網站」的功能較沒有幫助，以致於影響自動標註系統的有用性。

- (三) 在兩種標註系統的科技接受度有明顯差異的是能否有效閱讀以及能否容易使用：在科技接受度所有題項中，認知有用性的「系統所提供的標註功能讓我的文本解讀過程更為順暢」、「使用這樣的系統解讀文本比使用一般的數位典藏資料庫更有效果」題項；在認知易用性的「這樣的系統介面很容易使用」題項，以及在使用感受的「使用此系統解讀文本符合我的閱讀習慣」等 4 個題項， p 值小於 0.05，代表受測者對於不同標註系統中此 4 項認知達到明顯差異。

三、不同標註系統功能感受的滿意度

自動標註系統最有幫助的功能是「添加新詞」和「查詢詞彙」，最沒有幫助的功能是「自動斷詞」和「連結到來源網站」。在人工標註系統方面，受測者認為最有幫助的功能是「查詢詞彙」和「新增標註」，沒有幫助的功能是「標註地圖」。顯示受測者對於自動標註系統的自動斷詞並不受用，反而轉而利用添加新詞功能，以補強自動斷詞不足之處。同樣在人工標註系統上，能夠新增標註反而很受歡迎，說明目前系統的自動斷詞和自動標註皆有不夠精確問題，允許人工加入輔助修正機制，才能滿足使用需求。

四、不同標註系統使用經驗及其使用問題

- (一) 使用者關注的是能否直接有助於內容理解的功能設計：自動標註系統的「自動標註」和「查詢詞彙」對使用者而言，能藉由比對常用詞庫提供主動解釋，也能讓使用者自主增加需要加強的解釋，對於一般閱讀理解需求有直接的幫助。而「自動斷詞」、「連結來源網站」、「添加新詞」等較屬於輔助工具的功能，能幫助使用者延伸閱讀或是與別人分享自己的知識，使用者也能肯定功能的必要性，但對於「資料來源評分」功能，使用者感受像是給系統回饋，對於閱讀理解並無直接關聯，顯然較不受使用者

青睞。而人工標註系統提供的「新增標註」、「編輯標註」和「查詢詞彙」對於受測者而言，可以查詢內容和增加標註，對於閱讀古籍內容有直接幫忙。但「標註分類」、「標註地圖」對於閱讀理解過程較無直接關聯，受測者感受功能用途不強。

(二) 自動標註主要的問題是「自動斷詞」不精確，人工標註是標註操作不易的問題：如同 Chen、Chen 與 Liu (2018) 對於自動標註提出的問題，本研究由系統自動標註產生的斷詞會有因為不正確反而干擾閱讀，甚至造成閱讀理解困難。由於單純由機器比對產生的斷詞結果，會產生標示太多（太多藍字干擾閱讀）或是標示錯誤，造成閱讀干擾，目前自動標註斷詞之正確率，雖侷限於本研究工具功能設計的完善程度，但也意味自動標註的斷詞正確性很重要。人工標註雖然降低自動斷詞不精確問題，但受測者在人工進行標註選詞過程，因滑鼠點選不容易操作，而產生不易用的感受。

(三) 增加句讀、改善連結資料庫是受測者對於未來期許的共同需求：無論是自動或是人工標註系統受測者均提到對於古籍閱讀需要有句讀斷句標示，也希望能檢討目前設定的連結資料庫需要更符合查詢古籍的特性。

五、建議

(一) 調整自動標註「自動斷詞」詞庫，並改善「添加新詞」能更直觀友善操作。

自動斷詞的功能設計立意良善，也能受到受測者肯定，若自動標註能夠提升斷詞正確性，將可提升使用者感受系統之有用性。但目前因機器比對詞彙標示，會產生標示過多或標示不正確問題，建議未來系統設計需要調整自動斷詞所比對的詞庫來源，讓斷詞顯示比較符合古籍文本內容出現的詞彙。此外，也需要併同強化讓使用者無礙使用「添加新詞」的功能參與修正，才能讓自動斷詞的設計，因為有使用者參與修正，能夠產生更有助於輔助閱讀的效果，目前使用者對於添加新詞的操作不夠熟悉，不免阻礙了對於自動斷詞後的人為修正意願，未來可以調整此部分操作設計，提升使用者閱讀過程修正標註的意願。

(二) 改善人工標註系統的「新增標註」與「編輯標註」操作的易用性。

有鑑於進行人工標註過程中，會有標註詞彙的圈選操作不容易、標註屬性的下拉式選單不易點選、標註選錯卻不容易取消等問題，經

常造成使用者標註不精確，往往在前後多選一個字之後，標註的意思就不正確，根據受測者反映，建議此標註系統未來可以針對標註的點選方式提供更友善的操作設計，讓使用者方便運用。

(三) 建議標註系統未來新增句讀標示功能，可協助使用者對於內容的理解。

有相當多受測者希望系統能在古籍全文加上標點，在閱讀時能更加通順的解讀文意，降低閱讀文言文的困難。如果要由系統自動進行古文句讀的標示，需要找出標示規則，標出句讀的判斷規則需要先由熟悉古文內容專家提供學科知識，或是將已經有正確標示句讀的古文進行規則分析。建議可先在標註系統新增由使用者進行句讀標示的功能，累積使用者標出句讀的判斷結果，經分析找出句讀字詞或是位置的規律，未來才有可能發展出自動標示句讀的功能。

(四) 鼓勵更多使用者參與閱讀標註，藉以累積並促進標註內容深度。

編輯標註是提供閱讀過程對於別人已標註的內容增加評論或解釋，也能自己新增標註內容，「編輯標註」功能如同閱讀作筆記，可以自己作筆記，也可以看到別人作的筆記，雖然此項功能非常有助於閱讀理解，但目前標註內容深度不夠，僅是貼上查詢所得，這樣的筆記內容與自己透過查詢的內容相去不遠，致使本項編輯標註的功能未能充分發揮。未來要增加標註的內容深度應該有賴於使用者將閱讀所得寫出有深度的筆記內容，因此需要鼓勵使用者願意貢獻所知提供標註內容，系統可以增加標註評分排名，讓更多使用者願意貢獻有用的標註內容，而非僅止於目前呈現的淺層資訊標註。

(五) 建議古籍標註系統可以導入相關文本與延伸閱讀的推薦連結設計。

受測者使用標註系統除了查詢資料庫與閱讀標註，也希望能在閱讀特定古文時，有該篇古文的註解或是白話翻譯本，也有類似的建議是由延伸閱讀參考的角度，建議可以提供類似主題的相關文本推薦連結，或是類似個人化商品推薦的機制，主動推薦給使用者符合研究主題或是閱讀興趣的相關篇章內容。這些期許標註系統能採用延伸閱讀的觀點，適度加入古籍閱讀的輔助功能，但是對於如何促進有參考價值的標註內容以及累積閱讀知識的分享，才是標註系統真正要關注和處理的問題。

誌謝

本研究為科技部補助大專學生研究計畫 MOST 106-2813-C-004-001-H「古籍全文資料庫使用分析及使用者意見調查」研究成果，感謝陳志銘教授指導研究。

參考文獻

- 王桂蘭（2007）。漢籍電子文獻「二十五史資料庫」評介。國文天地，23(2)，16-21。【Wang, G. L. (2007). Han ji dian zi wen xian er shi wu shi zi liao ku ping jie. *The World of Chinese Language and Literature*, 23(2), 16-21. (in Chinese)】
- 吳明德、黃文琪、陳世娟（2006）。人文學者使用中文古籍全文資料庫之研究。圖書資訊學刊，4（1/2），1-15。【Wu, M. D., Huang, W. C., & Chen, S. C. (2006) Humanities Scholars and Databases for Ancient Chinese Books. *Journal of Library and Information Studies*, 4(1/2), 1-15. (in Chinese)】
- 郭明芳（2011）。中央研究院漢籍全文資料庫。國文天地，10，16-20。【Guo, M. F. (2011). Zhong yang yan jiu yuan han ji quan wen zi liao ku. *The World of Chinese Language and Literature*, 10, 16-20. (in Chinese)】
- 陳志銘、韋祿恩、吳志豪（2011）。認知型態與標註品質對閱讀成效之影響與關聯研究：以數位閱讀標註系統為例。圖書與資訊學刊，3(1)，1-25。【Chen, C. M., Wei, L. E. Wu, J. H. (2011). Assessing the Effects of Cognitive Style and Annotation Quality to Reading Performance: A Study on a Digital Reading Annotation System. *Bulletin of Library and Information Science*, 3(1), 1-25. (in Chinese)】
- 陳金木（2008）。電子全文資料庫與學術研究－以《四部叢刊電子全文檢索版》為例。明道通識論叢，5，120-135。【Chen, C. M. (2008). Electronic Version of Text Retrieving of Collection and Academic Research: Take Electronic Version of Text Retrieving of Collection of the Four Traditional Divisions of Texts as an Illustrative Example. *MingDao Journal of General Education*, 5, 120-135. (in Chinese)】
- 陳勇汀（2011）。合作式閱讀標註之知識萃取機制研究（未出版碩士論文）。國立政治大學，台北市。【Chen, Y. T. (2011). A Study on Developing Knowledge Extraction Mechanisms from Cooperative

- Reading Annotation (Unpublished master's thesis). National Chengchi University, Taipei. (in Chinese)】
- 陳燦珠（1998）。使用者與中文全文檢索系統互動關係研究：以中央研究院「漢籍全文資料庫」的使用為例。圖書館學與資訊科學，24(2)，65-85。【Chen, T. C. (1998). A Case Study on Human-Machine Interaction for Handy Text Retrieval System. *Journal of Library and Information Science*, 24(2), 65-85. (in Chinese)】
- 項潔（2014）。數位資源對於歷史學研究的意義：一個數位人文研究者的觀察。國史研究通訊，7，4-9。【Hsiang, J. (2014). Shu wei zi yuan dui yu li shi xue yan jiu de yi yi: yi ge shu wei ren wen yan jiu zhe de guan cha. *Academia Historica Research Newsletter*, 7, 4-9. (in Chinese)】
- 黃寬重（2012）。數位時代人文研究的衝擊與蛻變。漢學研究通訊，31(1)，1-6。【Huang, K. C. (2012). The Impact and Metamorphosis of Humanities Research in the Digital Age. *Newsletter for Research in Chinese Studies*, 31(1), 1-6. (in Chinese)】
- 廖益賢（2012）。電子文獻全文檢索資料庫管窺－以「中央研究院漢籍電子文獻」資料庫所收《文心雕龍》為例。中國文化大學中文學報，25，285-303。【Liao, Y. H. (2012). A Study of Electronic Full-Text Databases -Using “Wen Xin Diao Lung” of “Scripta Sinica” as an Example. *Zhong guo wen hua da xue zhong wen xue bao*, 25, 285-303. (in Chinese)】
- 羅鳳珠（1997）。在網際網路建立漢學研究環境的重要性及可行性：就中國文學而論。漢學研究通訊，16(1)，取自 <http://cls.hs.yzu.edu.tw/present/han.htm> 【Lo, F. J. (1997). The Importance and Feasibility of Sinology on the Internet — A Focus on Chinese Literature. *Newsletter for Research in Chinese Studies*, 16(1) Retrieved from <http://cls.hs.yzu.edu.tw/present/han.htm>. (in Chinese)】
- Adler, M. J., & Van Doren, C. L. (1972). *How to Read a Book*. New York: Simon & Shuster.
- Annotation. (2017). American Dictionary of the English Language. Retrieved September 25, 2017 from <http://webstersdictionary1828.com/Dictionary/annotation>
- Cabanac, G., Chevalier, M., Chrisment, C., & Julien, C. (2007). Collective annotation: perspectives for information retrieval improvement. *In*

- Proceedings of the Eighth Conference on Information Retrieval and its Applications (RIAO'07)*. Retrieved from ftp://ftp.irit.fr/pub/IRIT/SIG/2007_RIAO_CCCJ.pdf
- Chen, C. M., & Chen, F. Y. (2014). Enhancing digital reading performance with a collaborative reading annotation system. *Computers & Education*, 77, 67-81.
- Chen, C. M., Chen, Y. T., & Liu, C. Y. (2018). Development and evaluation of an automatic text annotation system for supporting digital humanities research. *Library Hi Tech*, 37(3), doi:10.1108/LHT-10-2017-0219
- Chen, C. M., & Tsay, M. Y. (2017). Applications of collaborative annotation system in digital curation, crowdsourcing, and digital humanities. *The Electronic Library*, 35(6), 1122-1140.
- Chen, C. M., & You, Z. L. (2018). Community Detection with Opinion Leaders' Identification for Promoting Collaborative Problem-based Learning Performance. *British Journal of Educational Technology*, 50(4), 1846-1864.
- Hansen, F. A. (2006). Ubiquitous annotation systems: technologies and challenges. In *HYPERTEXT '06: Proceedings of the 17th conference on Hypertext and hypermedia* (pp.121-132). New York, NY: ACM Press.
- Horner, J., & Thirlwall, D. (1988). Online searching and the university researcher. *Journal of Academic Librarianship*, 14(4), 225-230.
- Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, 69, 121-130.
- Jan, J. C., Chen, C. M., & Huang, P. H. (2016). Enhancement of Digital Reading Performance by Using a Novel Web-based Collaborative Reading Annotation System with Two Quality Annotation Filtering Mechanisms. *International Journal of Human-Computer Studies*, 86, 81-93.
- Marshall, C. C. (1997). Annotation: from paper books to the digital library. In *Proceedings of the second ACM international Conference on Digital Libraries* (pp. 131-140). New York, NY: ACM Press.
- Marshall, C. C. (1998). Toward an ecology of hypertext annotation. In *Proceedings of the ninth ACM conference on Hypertext and*

- hypermedia: links, objects, time and space--structure in hypermedia systems* (pp. 40-49). New York, NY: ACM Press.
- Renzel, D., Cao, Y., Lottko, M., & Klamma, R. (2010). Collaborative Video Annotation for Multimedia Sharing between Experts and Amateurs. In *Proceedings of the 11th International Workshop of the Multimedia Metadata Community*, Barcelona, Spain (pp.7-14).
- Sato, T., Goto, M., Kimura, F., & Maeda, A. (2016). Developing a collaborative annotation system for historical documents by multiple humanities researchers. *International Journal of Computer Theory and Engineering*, 8(1), 88-93.
- Stone, S. (1982). Humanities scholars: Information needs and uses. *Journal of Documentation*, 38(4), 292-313.
- Wiberley, S. E. (1991). Habits of humanists: scholarly behavior and new information technologies. *Library Hi Tech*, 9(1), 17-21.
- Wolfe, J. (2002). Annotation technologies: A software and research review. *Computers and Composition*, 19(4), 471-497.

The Usage Analysis between Manual and Automatic Annotation Function of Sinology Database

Chia-Yu Cho

Graduate Student

Graduate Institute of Library and Information Studies
National Taiwan Normal University

Introduction

In this study, reading interfaces of ancient books were analyzed for investigating the differences of reading efficiency, technology acceptance and reading experiences when users read ancient books using reading interfaces with manual or automatic annotation functions. The goal was to understand user experiences and needs of using different annotation functions. The study results would help build a full-text system of ancient books that meets users' needs for assistive reading functions.

Research Methods

Participants were recruited in this study to have a test of using and annotating in interfaces with automatic or manual annotations. Participants were students of Chinese literature and history departments at universities in Taipei, who were capable of reading and interpreting ancient books of Ming Dynasty. Thirty participants were divided into two groups, using different annotation systems, reading the same text contents of Ming Dynasty. After reading with the annotation systems, the participants wrote a reading summary immediately, for revealing the differences of their reading and comprehension efficiencies. After the participants finished their reading and noting, they were asked to fill out an inventory of technology acceptance model and a satisfaction survey on system functions. After the test, participants were interviewed for revealing their experiences and opinions of using the annotation systems.

Results

1. Impacts of different annotation systems on comprehension of reading ancient books

The mean score of users' reading comprehension tests was 6.63 (of 10, the full marks), and the scores were distributed between 5 and 9, with 26.7% of participants obtaining a score over 9. It suggested that most participants had a relatively high score on reading comprehension. The mean score of reading comprehension using manual annotation systems (7.00) was slightly higher than the ones using automatic annotation systems, but with a non-significant difference.

2. Impacts of different annotation systems on technology acceptance of reading ancient books

- (1) In terms of overall technology acceptance, the accessibility was higher than the usefulness, with the mean of accessibility (3.7) higher than the mean of usefulness (3.0), in a 5-point Likert scale of technology acceptance. Among all the question items regarding "Perceived Ease", the highest scores fell on the item "The system helps me get useful information when I need", and the lowest score fell on "The system enables me to interpret correctly the text content". It suggested the participants considered the annotation functions helpful in obtaining information, but not so helpful in terms of helping interpret effectively the text content.
- (2) Manual annotation systems were considered with a better usefulness, while automatic annotation systems had a higher score on usability. The results of comparing technology acceptance between automatic and manual annotation systems, revealed that manual annotation systems were with a higher usefulness than automatic annotation systems, and users considered manual annotation systems were more useful for reading ancient books. However, in aspects of usability and user experiences, automatic annotation systems had a greater performance than manual annotation systems, that is, although participants considered

manual annotation systems more useful, users had better experiences of using automatic annotation systems in actual operations.

- (3) The biggest differences of technology acceptance between the two types of annotation systems lied in question items including “Is it helpful in effective reading” and “Is it easy to use”. Among all question items, participants had significant differences in four question items of technology acceptance, including “the provided annotation functions facilitate my reading and interpreting of text” and “using such systems for reading and interpreting text is more effective than using digital archive databases” in “Perceived Usefulness” aspect, “such system interface is easy to use” in “Perceived Usability”, and “using such systems for reading and interpreting text suits my reading habits” in “Perceived Experiences”.

3. Perceived Satisfaction of Different Annotation Systems

For satisfaction perception, a 5-point Likert scale ranging from “Not Useful At All” (1 point) to “Very Much Helpful” (5 point) was adopted. The most helpful functions of automatic annotation systems perceived by the participants were “new word adding” (3.9) and “phrase searching” (3.5), while the functions not so helpful were “automatic word segmentation”(2.5) and “ linking to source websites” (2.9). As to the manual annotation systems, participants considered the most helpful functions were “phrase searching” (4.4) and “adding new annotations” (4.0), while the function not so helpful was “annotations in maps” (3.1). It suggested that participants considered the automatic word segmentation of automatic annotation systems not so helpful, but they would use the function of adding new words to offset the deficiency of automatic word segmentation. On the other hand, in manual annotation systems, the function of adding new annotations was popular, suggesting that the automatic word segmentation and automatic annotations were not enough accurate. Only when the correction mechanism of adding new words manually was permitted could it meet users’ needs.

4. Experiences and Other Problems of Using Different Annotation Systems

The results of interviewing the participants for their using experiences revealed:

- (1) The most important function of annotation systems was to help comprehend the text. The automatic annotation and phrase searching of automatic annotation systems could provide automatic explanations by comparing commonly-used phrases, and could automatically add necessary augmented explanations. All of these facilitated directly the general reading comprehension. Automatic word segmentation, linking to source websites, and adding new words were functions as assistive tools, which helped in further reading or knowledge sharing, thus the participants also considered them necessary functions. However, the function of rating the credibility of data sources did not concern directly with reading comprehension, thus not considered useful by the participants. The functions of adding new annotations, editing annotations, and phrase searching provided in manual annotation systems allowed searching in content and adding annotations, thus helping directly in reading and interpreting ancient books. However, annotating categories and annotating in maps were relatively irrelevant to the process of reading comprehension, and were considered not useful by the participants.
- (2) The major problem of automatic word segmentation in automatic annotations was not accurate enough, while the problem of manual annotations was with an operational procedure that was not easy to use. Word segmentation generated in automatic annotations were not accurate enough, thus might interfere with reading, even cause difficulties in reading comprehension. Although manual annotations might reduce the problem of inaccurate word segmentations, participants perceived it not easy to use when using a mouse to select and click on phrases.

Discussions and Conclusions

Based on results of this study, a few suggestions are provided here for improving annotation systems in the future.

1. Word databases for automatic word segmentation comparisons should be adjusted for enhancing the accuracy of word segmentation, and improving the operational procedures of adding new words for encouraging uses of annotation modifications.
2. The usability of functions of adding new annotations and editing annotations should be improved.
3. The function of adding periods and commas should be incorporated into annotation systems, for facilitating users' comprehension of ancient books' contents.
4. More users should be encouraged to participate in reading and annotating, for increasing both the quantity and quality of annotation contents.
5. The function of recommendations for further reading and website linking should be incorporated into the annotation systems of ancient books.