

國立臺灣大學文學院圖書資訊學系

碩士論文

Department of Library and Information Science

College of Liberal Arts

National Taiwan University

Master Thesis

使用者查詢擴展行為研究：以 MeSH 為詞彙來源的建議詞彙工具為例

**A User Query Expansion Behavior Study:
Using MeSH as Term Suggestion Source**

洪承理

Cheng-Li Hung

指導教授：唐牧群 博士

Advisor: Muh-Chyun Tang, Ph.D.

中華民國 99 年 6 月

June, 2010

謝 辭

悠悠三載，倏忽即逝。轉眼已然畢業時分。寫論文的日子，戛然而止，餘下的只有一篇又一篇的文獻以及只能資源回收的廢紙。縱使對於這不事生產的學生生涯多有眷戀，於今也只能就此落幕。

論文的完成，首先要感謝唐教授牧群的循循善誘，殷勤指導，讓我能夠順利的完成論文的撰寫。猶記得當初考完資格考後的惶恐，若非玉信在系館的階梯上引薦唐老師，我相信我一定惶惶不可終日到研四。

感謝林教授維真及吳教授怡瑾和卜教授小蝶對於論文的評批與指教，使得論文能夠更趨於完善。

藍教授文欽的治學風格以及盧教授秀菊授課之嚴謹，強烈的震撼了我對於學習的態度，謝教授寶煖的課程也著實對於論文的寫作發揮了極大的影響，陳教授光華的幽默風趣亦重新點燃我對於學習的興趣。

感謝研究所同儕玉信、宛青、瑋妮、雅鴻、一帆、瑜君，雖然彼此之間沒有很熟稔，但是在越來越陌生的研究室中看見熟悉的面孔總是令人寬慰的，祝大家論文順利。坐在研究室裡的日子，大家就像是關在同一間牢房的獄友，彼此雖然不認識，但是天天看到同樣的面孔，總是有些熟悉。我個人熟識的陌生人就有7、8個以上。祝這些陌生人都可以順利畢業。

感謝爸媽無怨無悔的支持我攻讀研究所，同時容忍我宅在家中無所事事還整天打電動看漫畫。

感謝女友琇婷，五年半來不離不棄，互相磨練砥礪。感謝善意第三人東玫及其他政大圖檔所同學，讓我能夠自由出入百年樓，以及毫無尷尬的一起出遊，讓我在索然無味的研究所生活中找到了些許繼續向前的樂趣。此外，蒙大學同學俊翰、元興、宜均、婉婷不棄，偶爾會找我聊聊天，讓我感覺到自己的存在。

論文對我而言，曾經就像是薛西佛斯的巨石，終無完成的一日。在時間的淬煉及老師的薰陶之下，我終於完成了。這一天來得如此突然，如此缺乏真實感，直到現在我依然無法相信我已寫完了論文。但名為論文的戲碼確實就此告一段落，反倒令人不勝感慨。

真沒想到我也有寫完論文的一天，謝謝大家。

摘要

由於傳統的資訊檢索評估以回收率及精確率做為系統評估之指標，忽略了使用者在資訊檢索中所扮演的角色，因此近來產生了互動式資訊檢索評估。此外，由於使用者在檢索系統中多只會利用簡單關鍵字進行查詢，因此會得到大量的檢索結果而無法管理。因此本研究設計了一針對 PubMed 醫學資料庫的詞彙建議系統 MAP(Multiple Access to PubMed)，目的在於輔助使用者取得其所需的資訊。

本研究以實驗法及問卷法進行研究，同時設計詞彙建議系統並採用使用者的真實資訊需求做為檢索任務，以生物科學及醫學系所研究生為研究對象，總計 44 名受試者。並就使用者查詢行為及查詢擴展行為進行分析，同時比較不同介面過濾檢索結果之功能以及說明實驗產生之次序效應。

研究結果顯示：(一)主題熟悉度會影響使用者查詢行為，不熟悉主題任務其查詢相似度會低於熟悉主題任務。(二)不同介面會影響使用者查詢行為，MAP 系統能夠提供使用者大量的建議相關詞彙，因此使用者可從中利用之，而在 PubMed 介面中使用者查詢詞彙的刺激有限。(三)不同主題熟悉度與不同介面對使用者查詢行為產生之交互作用，MAP 介面在使用者較不熟悉檢索主題時可以發揮較大的作用(四)不同主題熟悉度未影響使用者查詢擴展行為，使用者在不同主題熟悉度下，其瀏覽及最終選用詞彙類別之比例並未呈現明顯之變化。(五)不同介面會影響使用者查詢擴展行為，原因在於 MAP 能夠提供全面性的領域內的相關詞彙，而 PubMed 則僅能提供使用者相關的檢索結果，造成使用者在詞彙選擇上的差異。(七)MAP 介面能夠輔助使用者取得在大量檢索結果中排序較低的相關文件。

關鍵字：互動式資訊檢索；查詢擴展；使用者行為；詞彙建議系統；醫學標題表

Abstract

It has been argued that traditional information retrieval evaluation is ill-equipped to address the need to validate the efficacy of today's highly interactive systems, which require users' active participation to be effective.

To answer the challenge of interactive information retrieval evaluation, a novel methodology was applied to test the effectiveness of MAP (Multiple Access to PubMed), a metadata-guided search interface for PubMed bibliographic search. The most distinctive aspect of our methodology is to use real users searching for real search requests on real system, instead of using assigned tasks common in traditional IR evaluation. To control the impact of individual search requests on search performance, a repeated measure design was adopted where users' search request served as its own control of variance. Comparisons of information behaviors between MAP and the regular PubMed interface were made.

The purpose of the study is to examine whether interfaces and topic familiarity might interfere user's search and term selecting behaviors.

Some major findings are as follows. 1. The participants were found to input more diverse terms and make more submissions were made using MAP. There were also differences in term category selection between the two interfaces. 2. Topic familiarity was also shown to influence users' query expansion behaviors. Differences were found between the categories from which terms were selected between MAP and PubMed. 3. MAP was shown to help users to uncover relevance document ranked much lower in the original dataset. 4. MAP is more effective in search situations where the users were less familiar with the topics.

Key words: interactive information retrieval; query expansion; user behavior; term suggestion; MeSH; Medical subject heading; PubMed

目次

謝辭	i
中文摘要.....	iii
英文摘要.....	v
目次	vii
表次	ix
圖次	xi
第一章 緒論	1
第一節 問題陳述.....	1
第二節 研究目的.....	3
第三節 研究問題.....	3
第四節 研究範圍與限制	3
第五節 研究預期貢獻	4
第六節 名詞解釋.....	4
第二章 文獻分析	7
第一節 互動資訊檢索評估.....	7
第二節 查詢擴展.....	10
第三節 使用者詞彙選擇行為.....	13
第三章 研究方法與步驟.....	19
第一節 研究方法與設計	20

第二節 研究對象與研究工具.....	25
第三節 資料蒐集與分析	26
第四節 研究實施與步驟	27
第四章 研究結果與發現.....	31
第一節 使用者基本資料	31
第二節 使用者檢索特性分析.....	36
第三節 使用者查詢擴展行為.....	47
第四節 檢索結果過濾分析.....	65
第五節 實驗之次序效應	68
第六節 使用者評論	70
第五章 結論與建議	75
第一節 結論.....	75
第二節 建議.....	79
第三節 進一步研究之建議.....	81
參考文獻.....	83
附錄一 前測問卷	93
附錄二 正式實驗問卷	99
附錄三 次序效應案例	105
附錄四 受試者檢索問題.....	109

表 次

表 3-1 實驗分組表.....	24
表 4-1 受測者基本資料表	32
表 4-2 學院分類表.....	32
表 4-3 受測者檢索背景資料	33
表 4-4 檢索主題特性.....	36
表 4-5 檢索主題特性相關程度.....	36
表 4-6 新舊查詢之檢索特性比較.....	38
表 4-7 不同介面之檢索特性比較.....	39
表 4-8 重複檢索與介面對查詢行為各變數之二因子變異數分析摘要	42
表 4-9 重複檢索與否及不同介面交互作用之查詢行為	43
表 4-10 詞彙分類表.....	47
表 4-11 不同主題熟悉度瀏覽詞彙之類別	48
表 4-12 不同主題熟悉度選用之詞彙類別	50
表 4-13 全新檢索之瀏覽及選用詞彙.....	51
表 4-14 重複檢索之瀏覽與選用詞彙.....	53
表 4-15 瀏覽建議詞彙次數分配表.....	54
表 4-16 不同介面選用詞彙之類別.....	56
表 4-17 PubMed 介面使用者瀏覽與最終查詢詞彙類別.....	57

表 4-18 MAP 介面使用者瀏覽與最終查詢詞彙類別	58
表 4-19 不同介面及主題熟悉度之瀏覽詞彙	60
表 4-20 不同介面及主題熟悉度之選用詞彙	61
表 4-21 不同主題熟悉度高度利用查詢擴展詞彙類別	63
表 4-22 不同介面高度利用查詢擴展詞彙類別	64
表 4-23 不同介面及不同主題熟悉度高度利用查詢擴展詞彙類別	64
表 4-24 不同介面之相關文獻於初始檢索結果之位置	66
表 4-25 實驗次序效應分組表	69
表 4-26 先 PubMed 介面後 MAP 介面之次序效應案例	70
表 4-27 先 MAP 介面後 PubMed 介面之次序效應案例	70
表 4-28 使用者對於 MAP 輔助功能之評估	73

圖 次

圖 3-1 MAP 系統頁面	21
圖 3-2 MAP 介面之建議詞彙呈現	22
圖 3-3 研究步驟圖	29
圖 4-1 受試者利用 PubMed 及 Google Scholar 頻率圖	34
圖 4-2 不同介面及不同熟悉度之相似度平均數圖	44
圖 4-3 不同介面及不同熟悉度之新增詞彙數平均數圖	44
圖 4-4 不同介面及不同熟悉度之移除詞彙數平均數圖	45
圖 4-5 不同介面及不同熟悉度之最終查詢字數平均數圖	45
圖 4-6 不同介面及不同熟悉度之瀏覽文件數平均數圖	46
圖 4-7 不同介面及不同熟悉度之查詢次數平均數圖	46
圖 4-8 不同主題熟悉度瀏覽之詞彙	49
圖 4-9 不同主題熟悉度選用之詞彙類別	51
圖 4-10 全新檢索之瀏覽與選用詞彙	52
圖 4-11 重複檢索之瀏覽與選用詞彙	53
圖 4-12 使用者不同介面之瀏覽之建議詞彙次數長條圖	55
圖 4-13 使用者不同介面之最終選擇之建議詞彙次數長條圖	56
圖 4-14 PubMed 介面瀏覽及選用之建議詞彙橫條圖	58
圖 4-15 MAP 介面瀏覽及最終選擇之建議詞彙橫條圖	59

圖 4-16 Pubmed 介面及不同熟悉度之瀏覽詞彙.....	60
圖 4- 17 MAP 介面及不同主題熟悉度之瀏覽詞彙	61
圖 4-18 PubMed 介面及不同主題熟悉度之選用詞彙.....	62
圖 4-19 MAP 介面及不同主題熟悉度之選用詞彙	63
圖 4-20 使用者 2 號之不同介面文獻排序	67
圖 4-21 使用者 30 號之不同介面文獻排序	68



第一章 緒論

第一節 問題陳述

資訊檢索自發展以來，主要目的是能夠有效的組織資訊並從而讓使用者檢索到其所需的資訊。在資訊檢索發展早期，系統的效能是以使用者所檢索出的所有相關文獻數量為基準，分為回收率(Recall ratio)以及精確率(Precision ratio)。回收率是指使用者所檢索出相關的文獻與系統中所有相關文獻的比值，精確率則是指使用者所檢索出文獻中相關文獻的比率。在此一系統導向之評估方法中，主要有兩點爭議：其一為相關文獻的判斷，早期相關文獻之判斷為二元形式，由專家或電腦從給定的主題判斷文獻是否相關，其忽略了相關判斷的複雜性；其二為回收率與精確率兩者之間存有反比關係，若欲提升回收率，則勢必犧牲精確率。因此資訊檢索系統並無法同時提升回收率以及精確率來改善其檢索效能(黃慕萱，民85)。此外，系統導向的評估方法最受人爭議的地方在於忽略了使用者在資訊檢索中所扮演的角色，其將資訊檢索視為查詢詞(Query)與文件集(Data collection)的匹配過程。然而使用者在資訊檢索中所扮演的角色並非只是單純輸入查詢關鍵字進行檢索，其透過與資訊檢索系統及檢索結果互動的過程，不斷的重塑其資訊需求，改變其資訊需求表述，此一複雜的變化在系統導向的評估方法中是被忽略的。因此，必須要針對使用者在利用資訊檢索系統的行為進行調查，以發展能夠輔助使用者在不同情境下檢索其所需資訊的系統(Xie, 2008)。

在資訊檢索的過程當中，使用者需求的表達可以說是最為困難的任務(Belkin, Oddy, & Brooks, 1982)。在當前發展的資訊檢索系統之下，使用者往往只用少數幾個關鍵字來描述其內在複雜的資訊需求，例如使用者在利用搜尋引擎時通常是以短且模糊的檢索詞進行檢索(Jansen, Spink, Bateman, & Saracevic, 1998)。而另一方面，大型資料庫中的文件則依照資料庫之特性著錄其相關的詮釋資料如題名、

作者、關鍵字、主題等，然而使用者往往只會運用若干關鍵字進行檢索，資料庫繁複的著錄欄位與使用者表達其資訊需求進行檢索的簡易關鍵字呈現產生對比，使用者的關鍵字檢索並無法完整利用資料庫中所有的欄位。兩者的不對稱關係往往讓使用者無法檢索到所需的資訊。

在文件檢索技術的演進之下，目前已發展出能夠以自然語言進行全文索引，使得一般的使用者可在不需要專家的協助下也能夠獨力進行檢索。然而，在面對如 PubMed 如此大型且內容資料異質性較高的資料庫時，一般使用者與檢索系統之間的互動往往會遭遇到新的困難。PubMed 的許多預設功能，如擴展(explode)、題名和摘要欄位的全文索引等，其目標都在於協助一般使用者獲得足夠的檢索結果。但這些提升索引詳盡性的功能也使得搜尋的精確率下滑。當精確率下降時，使用者必須面對數量龐大的檢索結果，若無有效的工具協助，使用者將無法有系統地縮小檢索範圍，而往往只能選擇倉促瀏覽檢索結果的前數十筆資料而已。

上述這種現象則造成了資訊超載(information overload)的後遺症，使得使用者與系統間的回饋(feedback)過程窒礙難行，無法有效進行溝通。過多的檢出文件至少造成了兩個人機互動的問題。其一是使用者所需的相關文件可能堆積在檢出文件集的深處，而使用者倉促之下瀏覽檢索結果並無法取得利用之。其次，由於使用者只能窺見檢索結果的冰山一角，往往無法針對其初始的關鍵字進行有效性的判斷。

本研究以 PubMed 資料庫之基本查詢介面為研究對象，並將詞彙建議系統加諸於 PubMed 之上。其目的在於了解在大型異質資料庫中，詞彙建議系統所扮演的角色和所能提供的輔助以及使用者之詞彙選擇行為。PubMed 資料庫為美國醫學圖書館提供作為 Medline 檢索的服務，其包含了自 1950 年代起的生物期刊文獻之引文資料庫，涵蓋醫藥，護理，牙科，獸醫，衛生保健，臨床醫學等學科。這些引文來自 Medline 及其他生醫期刊共 4600 種，總計 1,200 萬筆引文資料。

第二節 研究目的

本研究之目的主要是為了發展應用於大型資料庫如 PubMed 之詞彙建議系統，此外並透過分析使用者利用詞彙建議系統之行為以了解查詢與資訊需求之關係以及在何種情境下詞彙建議系統能夠最為有效的輔助使用者。

具體而言，本研究擬達成之研究目的為：

- 一、設計並測試以控制語言來協助使用者表達與改善其查詢(Query)的互動式介面。
- 二、探討搜尋情境與檢索行為之間的交互作用。
- 三、瞭解詞彙建議系統能夠發揮作用的檢索情境。
- 四、瞭解使用者利用詞彙建議系統的檢索行為。

第三節 研究問題

- 一、在不同介面下使用者之利用詞彙建議系統所選擇之詞彙其與初始查詢的語意關係是否有所不同。
- 二、不同介面下使用者的查詢擴展行為是否有所差異。
- 三、不同主題熟悉度是否會影響使用者查詢及查詢擴展行為。
- 四、詞彙建議系統是否能輔助使用者取得更多的相關文獻。
- 五、詞彙建議系統在何種情形下能夠提供使用者最大的效用。

第四節 研究範圍與限制

本研究的研究限制如下：

- 一、本研究所蒐集之參考文獻，以國內可取得之中、英文文獻為限。

- 二、本研究以生醫相關系所之學生為研究對象，無法推論至相關從業人員。
- 三、本研究以模擬的情境來取代真實的檢索情境進行實驗。
- 四、本研究以 PubMed 資料庫作為實驗之標的，並設計一詞彙建議系統與 PubMed 一般檢索介面比較使用者查詢擴展之行為。因此研究結果無法推論至其他類型資料庫之使用者查詢擴展行為。

第五節 研究預期貢獻

目前大多數查詢建議或是擴展的方法均仍處於實驗階段，尤其國內此一技術尚未能夠被業界廣泛的採納。當網路上的資源不斷成長，其內容將更為多元，使用者越需要有效的檢索介面來輔助其取得所需的資訊。查詢擴展的互動式的技術將有效提升檢索系統的效能，而 PubMed 是目前全球使用最頻繁醫學書目資料庫，其重要性自然不言而喻，因此以 PubMed 資料庫作為介面輔助對象，進而提供使用者詞彙建議功能。再者，本研究的結果也將深化檢索情境以及檢索工具之間關係的了解，進而對於資訊尋求行為理論的發展做出貢獻。

而透過分析使用者查詢擴展之行為，並比較不同介面、不同資訊需求下的差異，能夠進一步理解使用者操作系統上所利用之功能，從而作為設計詞彙建議系統之準則。

第六節 名詞解釋

一、查詢(Query)

使用者將己身的資訊需求利用數個關鍵字句來表達，並將其輸入資訊檢索系統，目的在於取得資訊系統中符合資訊需求的文件。而表達需求的關鍵字句即本研究所指稱的查詢。

二、相關回饋 (Relevance feedback)

相關回饋是利用使用者初始查詢中相關的部分來進行進一步查詢，為檢索系統輔助使用者改善其檢索結果之機制。其步驟為 1.使用者輸入初始查詢 2.系統傳回相關之文件 3.使用者點選相關或不相關之文件 4.系統重新計算使用者之資訊需求表徵 5.系統傳回改變的檢索結果。

相關回饋可分成顯性回饋、隱性回饋、自動回饋三種。顯性回饋是指透過使用者判斷初始查詢中相關部分，而後系統傳回修改之查詢結果。使用者會透過二元或給分的方式指出相關的程度。隱性回饋則是透過使用者行為如瀏覽及不瀏覽之文件，瀏覽時間，視窗下拉動作等來推測，而由系統自動進行相關回饋。自動回饋則是自動處理相關回饋的操作，因此使用者不必透過多餘的互動便可取得改善的資訊檢索效果。此方法是使用者進行一般檢索後取得相關文件之集合，系統假設前數筆之文件為相關，並自動以該假設進行相關回饋，取代了使用者判斷相關文件的步驟。(Baeza-Yates, 1999; Manning, 2008)

三、查詢擴展(Query Expansion)

查詢擴展意指重新組合原始查詢來改善檢索效能的過程，其牽涉到使用者輸入的評估及擴展查詢來取得更多的文獻。其範圍包含了搜尋同義詞、變形詞、拼字更正或建議以及原始查詢的重新給予權重等作業。(Efthimiadis, 1996)

相關回饋與查詢擴展之差異在於前者是由使用者點選相關之文件回饋至檢索系統，而後傳回相關之結果。後者則是透過系統提供之建議詞彙讓使用者增加查詢關鍵字以改善檢索結果。

查詢擴展可分為利用控制詞彙工具、人工索引典、自動產生索引典、查詢 log 檔分析等方法，分析使用者初始之查詢，而後系統傳回相關之建議詞彙供使用者利用點選之。此類方法不與使用者初始檢索結果進行互動，而是以分析資料庫中

所有記錄產生相關建議詞彙。查詢擴展亦可透過使用者進行初始查詢後，分析初始查詢結果集，從而產生相關建議詞彙。此方法則是與自動回饋相仿，均假設使用者初始檢索結果之前 N 筆文件均為相關之資訊而後進行分析。(Manning, 2008)

四、檢索詞彙選擇行為(Search Term Selection Behavior)

意指使用者在使用提供詞彙建議功能之系統時，在檢索的過程中，依據系統提供之詞彙對初始查詢進行修改，發生於其中的詞彙選擇行為，包含新增、刪除、替換等行為。

五、互動式資訊檢索(Interactive Information Retrieval)

發生在資訊檢索時的互動溝通過程，其牽涉到資訊尋求與檢索中主要的參與者，包含了搜尋者、社會組織情境、資訊設備、界面、資訊空間等(Ingwersen & Järvelin, 2005)。



第二章 文獻分析

第一節 互動資訊檢索評估

自從 60 年代 Cranfield 的實驗(Cleverdon, 1967)，乃至晚近的 Text Retrieval Conference (TREC)為資訊檢索的演算方式建立了一套以回收率和精確率等系統導向的指標為基礎的評估典範。在傳統資訊檢索中，僅比較了使用者輸入及系統輸出，其並未描繪出變動的互動過程。(Saracevic, 1997)。Belkin (1993)認為傳統的資訊檢索有兩個基本的假設：1.資訊需求是固定的，並且可以清楚描述；2.資訊尋求行為只是將需求透過檢索與已清楚描述的相關資訊進行連結。但是 Cranfield 典範並沒有辦法反應今日動態以及互動式的檢索過程，在評估檢索系統時，如何均衡反應真實的使用情境和建立客觀的評量標準這兩項要求已經成為一項日益受到重視的課題(Robertson & Hancock-Beaulieu, 1992)。

對於實驗是否反映使用者真實使用情境的反思主要源自學界對於使用者檢索行為的了解。在受到使用者導向，認知學派資訊檢索理論的影響之下(Belkin et al., 1982; Saracevic, 1997; Saracevic & Kantor, 1988a; Saracevic, Kantor, Chamis, & Trivison, 1988)，研究者開始注意到一些可能影響檢索過程的情境因素，近來學者對於整合資訊尋求行為以及資訊檢索的呼籲正反應了此一趨勢(Vakkari, 1999)。

使用者導向的資訊檢索理論最大的貢獻在於將使用者此一過去未受重視的要素納入評估的過程中(Dumais & Belkin, 2005; Robertson & Hancock-Beaulieu, 1992)。Robertson & Hancock-Beaulieu(1992)指出了三個互動式評估的基本原則如下：

1. 以潛在使用者作為測試的對象。

2. 使用動態的及個人化的資訊需求。
3. 採用多面向的以及動態的相關性判斷。

TREC 中的互動專題(interactive track)是以使用者為中心的評估模式的開端。在此一互動式評估的典範中，除了一般評估中所關注的搜尋結果之外，使用者的線上檢索行為以及其對檢索結果的影響也一併成為研究分析的對象。

互動式評估典範的形成是因應諸如相關回饋和瀏覽式的檢索工具此類互動性技術的發展。當非專家使用者的搜尋越來越頻繁，諸如介面好用性(usability)以及使用者心智模型(mental model)等議題也越來越受到重視。互動式的評估典範所帶來的影響主要有二，一為新的評估指標的需求，二為搜尋任務(task)的設計。

儘管互動式的檢索並不必然與回收率和精確率等傳統指標相衝突，但有越來越多的研究者認為有必要將系統是否真正能夠有效的幫助使用者完成他們手上的工作做為評估的指標之一(Hersh, Pentecost, & Hickam, 1996; Hersh & Hickam, 1998; Vakkari, 2000)。另外一個需要新的評估指標的原因是在網路環境中找出完整的相關文件(relevant set)有其實際困難，因而研究者開始嘗試以使用者的滿意程度或者是被搜尋文件的效用來作為評估指標。事實上，過去也有研究指出這些主觀的指標與回收率和精確率有顯著的相關(Belkin et al., 2003)。

Su (1992, 1994)透過分析 40 個使用者及 6 個透過專業人士檢索大型線上系統來辨識最適合的互動資訊檢索評估指標，其研究結果顯示檢索結果(search result)是最好的單一效能指標。使用者對於結果的滿意度及精確率與檢索結果明顯相關，但精確率與檢索成功之間並未明顯相關。對使用者而言，回收率較精確率更為重要，其原因在於：1. 高精確率並非代表高品質的檢索；2. 受到檢索任務的影響，使用者的目標並非是檢索的精確率，反而是盡可能取得較為完整的相關資訊亦即

高回收率的檢索結果來完成任務。

Hersh et al. (1996)發展了一套模擬使用者在工作環境中所可能面臨的資訊問題來做為受測者搜尋的任務，該系統是以它們能夠幫助受測者回答這些問題的程度來做為評估標準。

Su (2003)提供了一組綜合的互動式資訊檢索系統評估指標清單，共有 5 大面向包含了 16 種指標。其分為：1.相關；2.效率(時間、查詢數)；3.效用；4.使用者滿意度；5.連結性。

Ruthven(2007)指出互動資訊檢索包含了研究及支援多變的使用者取用及檢索系統。互動資訊檢索的研究可以分為 1.資訊尋求及檢索行為 2.發展與電子資源互動的新方法。前者提供了在互動資訊檢索中資訊取用決策的藍圖，後者則是對於新的科技能有進一步的認識以提供資訊取用。理想的互動資訊檢索是由終端使用者建立資訊需求的形式；而這些需求則輸入至能夠檢索到資訊物件或相關參考資訊的電子檢索系統；而使用者取用檢索結果後進而抽取出相關的物件或資訊。

Xie (2008)認為檢索效能及功能使用性是互動式資訊檢索評估關鍵的指標，前者是由檢索結果的相關性(精確率及回收率)所構成，後者則是以效能(effectiveness)為面向，包含了使用性、學習性、錯誤發生率、滿意程度等。其認為互動資訊檢索系統的評估必須要具備下列三個面向：

1. 評估互動資訊檢索系統的效能。
2. 評估介面好用性及其組織好用性。
3. 評估使用者與資訊檢索系統之間的互動過程。

而資訊檢索互動的評估則可以分為以下面向：

1. 系統支援的程度。
2. 系統效率。
3. 使用者滿意的程度。
4. 系統提供支援的方式。

將使用者納入資訊檢索評估中之後，傳統的系統效能指標並無法測量使用者對於系統效能的觀點，因此學者亦從使用者之角度進行研究，諸如介面好用度、檢索結果滿意度等指標均是為了測量使用者之觀點。而互動式資訊檢索評估必須要包含兩者之面向才能夠準確的評估系統之優劣，並進一步改善檢索之結果。

第二節 查詢擴展

一、定義

在資訊檢索的過程中，檢索詞對於檢索的結果有著最直接的影響。如何協助使用者表達他們的資訊需求是此一領域重要的議題(Belkin et al., 1982)。在自動索引的環境中，文件索引之詳盡性相當高，然而一般使用者的檢索詞卻非常短，約2-3 個詞，換言之，使用者檢索問題的詳盡性卻沒有相對的提高。簡短的檢索詞彙通常會造成龐大的搜尋結果且缺乏精確性，而此一現象可透過查詢擴展(Query expansion)的方式來建議並引導使用者提出較多的檢索詞彙以獲得改善。

相關回饋的技術即是根據讀者與系統的互動提供詞彙建議，此一技術已在過去已被證明可以改善檢索結果(Harman, 1988, 1992)。近來也有研究者嘗試利用像是閱讀時間、捲軸移動以及文件取得等隱含性相關回饋(implicit relevance feedback)之指標來作為改善檢索詞品質的依據(White, Ruthven, & J. M. Jose, 2005)。另一種提供檢索詞建議的方法是從文件代理(surrogate)或詮釋資料(metadata)中抽取出和原始查詢共同出現的詞彙作為詞彙建議來源。(Hearst et al., 2002; Joho & Joemon M. Jose, 2006)。

查詢擴展的定義為使用者在資訊檢索的過程中將初始查詢透過增加或刪除檢索詞彙以加強檢索效能。其建立於初始查詢的建立並無法呈現使用者確切的資訊需求的假設上。而目前應用索引典來擴展及重組查詢最近蔚為熱門。

檢索詞修改可由自動化的方式自動完成，亦可透過使用者操作完成。自動化的方式又稱為相關回饋(relevance feedback)，一般是透過點選「相關文獻」等超連結回傳資料給系統所完成。使用者操作則是可根據檢索結果對原本的檢索詞進行篩選，亦可利用檢索工具所提供的機制進行。

二、形式

為了減緩資訊檢索的使用者需求及文獻兩者間的不對稱呈現，目前之研究多以自動或互動的方式來擴展使用者的查詢。其中一個方法是以廣泛運用的詮釋資料輔助檢索，透過動態擷取初始檢索結果集之詮釋資料來加強使用者之查詢(Hearst et al., 2002; Lin, 1999; Pollitt, 1998)。此外，在醫藥科學文獻檢索中，亦由研究透過控制詞彙(Doms & Schroeder, 2005)及文摘摘要的全文(Goetz & von der Lieth, 2005; Perez-Iratxeta, BORK, & Andrade, 2001; Plikus, Zhang, & Chuong, 2006)中抽取索引詞彙來發展詞彙的共現詞彙表以提供詞彙建議。

Shiri 和 Revie (2006)認為查詢擴展的實作有以下三種形式：

1. 人工系統(manual system)。
2. 自動系統(automatic system)。
3. 互動系統(interactive system)。

目前在人工及互動系統間的分野上有所爭議，因為兩者都牽涉到人的行為。兩者間真正的差異是在於人工系統並未包含與文件集的交互諮詢，但互動系統的查詢修改是經由一連串的回饋而來。上述的兩種方式均可利用第三方資源，包含字典及索引典等詞彙工具來完成詞彙建議之功能。

Shiri 和 Revie (2006)將利用索引典建構的資訊檢索系統分為：

1. 標準人工建構索引典：標準索引典中階層、等同、相關的關係已廣泛應用在詞彙選擇及查詢擴展的目的上。多數的研究著重於控制詞彙與自然語彙的效率

及效能。這些索引典也以應用於一些專家系統的雛型中以幫助使用者選擇檢索詞及擴展查詢。

2. 搜尋索引典(search thesaurus): 搜尋索引典又被認為是終端使用者索引典，包含大量的同義詞、類同義詞、變形詞的進入詞(entry term)用以輔助使用者找尋替代詞彙加入查詢中。

3. 自動化建構索引典：基於詞彙共現關係所製作之索引典，又稱為關聯式索引典(association thesaurus)，詞彙之關係建立於詞彙與詞彙間同時出現的次數作為彼此關係密切與否之判斷。

Kristensen 和 Jarvelin (1990)研究小型的搜尋索引典在回收率及精確率的效能，使用者僅參與提供搜尋及回饋的過程。其結論是搜尋索引典可以提升全文資料庫的檢索效能。其亦調查了5種檢索模式:基本檢索、同義字檢索、狹義詞檢索、相關詞檢索、綜合檢索中回收率及精確率的效率。研究顯示搜尋索引典是加強回收率的工具，而使用者的參與可以提升精確率。

Spink (1995)認為自動化與半自動化的查詢擴展已有所發展。自動化查詢擴展技術利用使用者的查詢文字及由使用者輸入檢索之相關的文件產生一系列的檢索詞來檢索出更多相關的文件。而在互動查詢擴展中，使用者則可自行選擇由系統建議的候選詞。而其研究檢視了不同詞彙選擇的來源檢索相關文獻的效能，包含使用者問題陳述、使用者互動、中介者、索引典、從相關文件抽取。其中索引典為第三有效的工具，次於問題陳述及中介互動。結果顯示詞彙相關回饋的檢索詞較專業建議檢索詞及選自索引典的檢索詞更能有效的檢索到相關的物件。

多數的研究顯示使用索引典可以幫助使用者選擇替代詞彙來擴展查詢，同時能夠改善檢索結果。但研究多關注於檢索結果而非檢索過程。必須要將使用者與索引典的互動納入研究，包含了使用者態度、檢索過程的認知狀態、使用者選擇

詞彙的機制等。另外也要研究在多重主題下索引典輔助的查詢擴展之效能。

互動查詢擴展研究是為了調查使用者問題的呈現及其在詞彙選擇的過程中所使用的工具或經驗來修改查詢。此法強調了行為決策的重要性及認知特質對於查詢的影響。

本文所提出的詞彙建議工具是將由使用者初次檢索所得到的檢索結果，擷取其中的 MeSH 詞彙並且將它們呈現給使用者，做為使用者增加或改善檢索詞之依據。基於在文件中 MeSH 主題詞的共現性(co-occurrences)，這些擷取出的詞彙可被視為與原始檢索詞彙具有概念上的相關性。

第三節 使用者詞彙選擇行為

Bates (1979)研究檢索策略時，認為其包含了檢索組成策略及詞彙策略。前者是建構設計詞彙組的過程及分析資訊需求和查詢元素的方法，包含提高查詢的精確率，使檢索結果更廣泛或專注。詞彙策略則是在搜尋組成的情境下中特定詞彙的選擇或修改，關注於在階層架構中上位詞、下位詞或相關詞的詞彙選擇。

Fidel (1984)指出檢索詞彙的數量與操作的數量的關係在線上檢索行為是相當重要的模式。步驟越多的檢索會較步驟少的使用者使用更多的詞彙。

Saracevic 與 Kantor (1988a, 1988b)及 Saracevic, Kantor, Chamis 與 Trivison (1988)的研究檢驗不同使用者對相同問題的檢索詞彙選擇，透過問卷、訪談、log 檔、錄影資料來蒐集資料。其將檢索詞彙來源分成 1.使用者口頭說明問題但未參考紙本問題說明；2.從紀錄的口頭聲明及使用者遞交的紙本問題中抽取；3.僅使用紙本問題的詞彙；4.使用紙本問題詞彙及適合的索引典。

Spink 與 Saracevic (1997)的研究指出檢索詞彙選擇的研究分為兩種，分別為自動演算法及人工法。前者是發展及評估不同演算法在詞彙重組及擴展上的選

擇、權重、排序來加強資訊檢索。而後者則是研究評估使用者在檢索時重組、擴展及修改其檢索所選擇的詞彙。其關注於認知及行為模型和影響詞彙選擇的因素。其重要的變數包含資訊需求、使用者意圖、個人特質及不同的資訊尋求個人檔，並分析其與詞彙選擇的關係。研究結果指出五種詞彙來源，分別為 1.問題說明；2.使用者與中介者之互動；3.索引典；4.詞彙相關回饋。其認為索引典詞彙在與使用者詞彙結合於檢索說明時最為有效。

Vakkari (2000, 2001)調查了 11 個學生在撰寫碩士論文時所使用的檢索詞彙及策略的改變。其工具為問卷、訪談、大聲思考法及 log 檔。其檢驗了檢索詞彙的種類及數量和使用的策略及與學生檢索前之知識架構的關係。他將新增的詞彙分成上位詞、下位詞、相關詞及同義詞。研究結果顯示對於任務的理解多寡會影響使用者選擇下位、同義詞和使用的策略。他認為對初學者而言，結構化的術語支援不僅能透過使用下位詞改善檢索結果，同時亦能幫助使用者解構主題。

Efthimiadis (2000)認為 IR 領域迫切需要真實而非模擬的互動搜尋的研究來調查真實使用者所進行的相關權重行為。其研究透過 INSPEC 資料庫調查電機學院的教職員生所進行的 25 次真實資訊需求的開放式檢索，透過問卷、log 檔及相關評估等方法蒐集資料。其問卷分為四個部分：前檢索問卷蒐集受測者背景資料；查詢擴展問卷蒐集受測者所選擇之詞彙及其與原始查詢之詞彙關係；檢索後問卷則蒐集受測者之滿意度、整體檢索印象、查詢與結果之評價、預期檢得相關筆數、以及紙本之檢索結果相關評估；最終問卷包含整體評價及使用者評論。

該研究詞彙選擇的分析主要是紀錄系統所產生的建議詞彙數量及受測者選擇其作為檢索詞的數量。受測者必須指出清單上所有適合使用於檢索的詞彙及其選擇的原因是基於 1.變形或同義字；2.相關詞；3.妥協詞(no better term)；4.新觀念的產生。另一部分則是要求受測者說明所選擇最適合檢索的五個詞彙與其原始查詢詞彙的關係，如索引典中的階層架構。研究結果顯示多數詞彙對初始查詢而

言多為同義字或相關詞，而部分詞彙則是新的概念。而選擇詞彙與初始詞彙的關係為：34%無明顯關係，66%的選擇詞彙與初始詞彙有關聯。這些關係之比例為下位詞 46%、上位詞 3%、相關詞 17%。

檢索效益的評估則是比較初始查詢及擴展後的查詢之間的效益，主要是基於受測者對於檢索結果的相關判斷後進行比較。研究結果顯示查詢擴展提升了檢索的精確性且提供較多的相關文獻。

其研究提供了對於特定使用族群之詞彙選擇行為的描繪，說明索引典的階層關係在提供詞彙建議的查詢擴展系統上對使用者有極大的助益。此外，除了傳統階層關係外，研究者認為使用者選擇新觀念詞彙的判斷需要再進一步調查。該研究之限制在於其字彙關係的判定是由受測者自行判斷而非由索引典的內在架構進行辨識，此外亦無明確討論其建議詞彙產生的機制，僅說明是由 INSPEC 資料庫所產生，無法作為後續研究的參考。

Shiri, Revie 與 Chowdhury (2001)將檢索詞彙策略及詞彙選擇行為的研究依照研究對象之不同分成 1.專業檢索者；2.專業檢索者及終端使用者的中介環境；3.專業及新手的檢索行為比較；4.只有終端使用者。

Shiri 和 Revie (2006)調查於索引典輔助的系統中使用者查詢擴展互動的行為，其研究對象為獸醫學系的學術教職員工及大學生。資料收集自 30 人操作的 90 次透過 OVID 介面檢索 CAB 摘要資料庫的檢索。資料蒐集工具為問卷、螢幕錄影、訪談。其研究結果發現不同背景之受測者所選擇之詞彙有顯著之不同。同時搜尋任務的複雜性也會影響使用者與索引典系統的互動，較複雜的任務需要更多的查詢擴展及詞彙選擇。另外使用者是否曾經檢索過類似問題的經驗亦會影響其選擇行為及對系統的評價。

該研究的研究方法主要分成以下兩個部分，分別為認知行為與資訊處理任務

及影響檢索的因素：1.使用者的看法及態度：諸如使用者滿意度、影響、對於系統的態度等；2.與系統好用度相關的議題：檢索過程、檢索、檢索時間、易使性、易學性等。該研究所記錄的變數為使用者背景、搜尋主題、搜尋過程、搜尋詞、索引典介面、使用滿意度等。此外亦檢驗了以下假說：1.複雜主題的搜尋所用的搜尋詞會較簡單者來得多；2.複雜搜尋會需要更多的查詢重組；3.主題搜尋經驗會影響使用者對於索引典提供詞彙的實用性的判斷；4.詞彙選擇的數量與詞彙的實用性有關；5. 查詢重組的數量會與索引典瀏覽及導航的難易度有關。

其研究分析所有受測者所瀏覽、選擇、重組查詢的詞彙及其映射到索引典的詞彙和該行為所發生的檢索階段。此外亦調查了受測者在增加其初始搜尋的過程中，所修改或增加的詞彙以及其與索引典的關係。多數人選擇查詢擴展的原因可能是因為在索引典的輔助下可以接觸到較多與初始查詢語意相關的詞彙進而選擇之。在詞彙關係部分則將建議詞彙與初始查詢的關係分成7個類別，除了傳統索引典的階層關係外尚有統計共現及字母順關係類別。統計共現詞彙是缺乏索引典關係但被系統基於共現性判斷其與初始查詢詞彙潛在相關，其與初始查詢無任何階層關係。字母順詞彙則是因為 OVID 提供的階層關係是由字母所排序的。此外其分析受測者未擴展其查詢之原因為：1.檢索詞與索引詞一致；2.單一詞彙缺乏索引典關係。

該研究中受測者選擇之詞彙與初始詞彙之關係可以分為下位詞(47%)、上位詞(22%)及相關詞(22%)三種。這表示終端使用者知道索引典內的關係並將其應用在搜尋上。受測者操作行為的數量依賴於工作的複雜度，越複雜的工作使用者需要使用越多的系統功能，查詢重組的數量也隨著複雜度的提升而增加。此外，主題搜尋的經驗對於使用者對於詞彙選擇的評價亦有所影響，沒有經驗的使用者較有經驗的使用者認為索引詞彙較為有用，且能忠實呈現其原始資訊需求。研究中的使用者有強烈的應用索引典階層的傾向，40%查詢擴展是選自下位/上位詞。

這意味著索引典具有提供使用者全新或替代觀念的潛力，研究者認為查詢擴展詞彙的語意關係與索引典所處理及涵蓋階層及關聯有關。INSPEC 資料庫具有 10,200 相關詞及 5,995 個階層關係，而 CAB 資料庫則有 15,617 個相關詞，44,398 個下位詞，121,56 上位詞，後者可以提供較多的下位詞。研究結果顯示 50% 查詢擴充之使用者會找到檢索前未知的概念，且大部分查詢擴充的使用者都可檢索到相關的資訊。此一結果說明了索引典具有提供使用者未曾思考過或替代觀點的能力。

網路的興起帶來的大量終端使用者檢索行為，為了提供更好的系統，必須要探索使用者的行為及態度。使用者行為包含了認知與行為的層面，如詞彙如何與查詢相關、或是特定的介面設計。另外商業資訊檢索系統近來也大量的應用索引典輔助使用者進行檢索，隨著更具互動性的檢索功能的出現(White, Kules, & Drucker, 2006)，新的線上檢索行為也應運而生。對於線上檢索行為的研究將有助於系統設計者了解使用者的偏好以及對系統演變的心智模型(mental model)，並用以設計更為人性化且擁有良好使用性的檢索介面。

這些發展使得更多的研究投入於系統的實用性及使用性研究、並探討其如何影響使用者選擇詞彙及查詢擴展。使用者研究可以延伸研究者對於使用者行為的了解，並從而提供系統設計的準則。

第三章 研究方法與步驟

本研究將以實證的方式來測試不同介面的對於使用者利用詞彙建議系統的影響。由於傳統以精確率和回收率為評估標準(performance criteria)的方式並不足以回答本研究所欲回答的問題。根據互動性(interactive)與認知(cognitive)資訊檢索理論，資訊檢索並非單一且單向的過程，而是一個動態的回饋過程，使用者的認知架構及資訊尋求的情境往往會影響到檢索工具的選擇與檢索的有效性。學者從線上檢索行為的研究中發現使用者的需求狀況(problematic situations)和使用情境(use contexts)往往會影響資訊檢索過程之眾多面向，例如資訊來源的選擇、檢索詞彙、檢索技術的使用和相關性的判斷等(Belkin, Chang, Downs, Saracevic, & Zhao, 1990; Saracevic et al., 1988; Vakkari, 2001; Vakkari & Hakala, 2000; Vakkari, Pennanen, & Serola, 2003)。

在檢索工具以及其使用情境日趨多元的今日，應該就兩者間的互動影響給予更多的關注，亦即在不同的使用情境下，必須使用相對應的檢索工具，方能得到最大的效果。在檢索工具的評估方面，被測試的工具和測試的情境任務(task)之相容性的則必須加以關注。在評估一新穎的搜尋介面時，找出它能夠發揮最大功效的使用情境，並進而發展一套最適合的評估方法與標準是很重要的，若無法找出工具最適當的使用情境，其效用將無法彰顯出來。事實上，過去並不乏針對某些類型的檢索任務設計出最合適的評估標準的例子，譬如在 TREC 實驗中的問答式(question-answering)專題(track)和新穎性(novelty)專題，其資訊系統即是依個別情境所特別設計的評估標準來進行評鑑的。

本研究關注於實驗介面能支援與解決何種類型需求情境的關心具體呈現在研究的設計上。在研究設計中，如何描述並且分類如使用者之需求情境此類抽象概念則成為一相當重要的課題。目前以使用者為導向的互動式檢索評估中，研究

者都事先精心設計一套模擬的搜尋情境，然後以實驗要求受測者針對這些情境進行檢索。

本研究除了遵守此一模擬情境模式之外，更重要的也將探索其他的可能，亦即讓受測者檢索他們自己真正(genuine)的問題，而非如一般檢索實驗中由實驗者指定(assigned)檢索任務(task)的方式。

在互動式檢索系統評估中，研究者常面臨的一個兩難是如何讓研究設計在反映真實搜尋情境的同時兼顧評估結果的客觀性。本研究對此兩難的解決之道在於採用重複量測的方式進行實驗，並且以受測者的真實資訊需求做為控制變數，這樣的設計方能反應受測者真實的檢索情況。因此，在研究設計中，受測者將在螢幕錄影軟體之記錄下進行實驗，使用其真實的資訊需求，搜尋所需要的資訊，另將透過實驗前問卷調查方式由受測者描述其檢索問題的特性。採用上述方法的優勢在於使用真實的需求情境；而透過問卷的方式來了解使用者的檢索情境，也避免以敘事(narrative)方式操作化檢索問題特徵的困難。

第一節 研究方法與設計

本研究旨在瞭解使用者在不同情境及介面下的查詢擴展行為，以實驗法分別比較不同介面及不同熟悉度的查詢擴展行為，以無干擾觀察法完整記錄受測者利用介面之行為，以內容分析法探究受測者查詢擴展詞彙之類別，並以問卷法記錄受測者對於建議詞彙系統之滿意程度。

一、實驗法

本研究欲以 PubMed 介面/MAP 詞彙建議介面做為處理，受測者之真實資訊需求為自變數，並以受測者利用該介面時所進行的詞彙選擇行為及受測者對於檢索結果之滿意度等為依變數，進行實驗。

(一)介面設計

本研究為測試大型資料庫提供詞彙建議之檢索效能，因此設計 MAP(Multi-Faceted Access to PubMed) 介面，作為實驗之工具，圖 3-1 為該介面之呈現。



圖 3-1MAP 系統頁面

MAP 介面中建議詞彙主要是由資料庫紀錄中後設資料的 MeSH 主題詞來完成，並依層面來組織詞彙以協助使用者瀏覽，其輔助使用者探索資訊空間的方法與 Hearst et al.(2002)所提出的層面式瀏覽大致相同。但是以大型資料庫 PubMed 之特性進行修改。

該介面能夠傳遞查詢至 PubMed，並產生醫學索引典 MeSH 之敘述詞讓使用者能夠精鍊其查詢。大型資料庫如 PubMed 若採取與檢索同時擷取後設資料可能會產生運算時間過長的問題，因此為了避免即時產生詞彙之共現矩陣，因此事先建立詞彙出現之資料庫。資料庫之共現詞表可以定期更新來改善呈現已出版文獻彼此之間的概念關係。此一先建資料庫包含了從 2006~2007 PubMed 所有的書目紀錄中所抽取的 MeSH 敘述詞、作者以及刊名。

當使用者透過 MAP 介面遞交其查詢時，系統會透過先建資料庫辨別出前兩百個高頻與查詢詞共現的詞彙，同時顯示於介面上提供使用者瀏覽。而當使用者

遞交非 MeSH 敘述詞作為其查詢時，使用者之查詢必須在先建資料庫中對映 (mapping) 到合適的 MeSH 敘述詞，此功能透過利用 PubMed 的自動翻譯表(該功能依序檢查檢索詞是否為 MeSH 敘述詞、進入詞(entry term)、子款目等，比對到相關款目時即將檢索詞彙轉換成控制詞彙。)完成。因此，當使用者遞交其查詢時，為了辨識每個詞彙的共現關係，透過 PubMed 提供的自動翻譯表所產生合適的 MeSH 敘述詞也將一併在本地資料庫進行查詢。在沒有合適 MeSH 敘述詞的情況下，將會啟動從 MEDLINE 檢索結果中擷取前兩百個敘述詞的機制作為建議詞彙的來源，如圖 3-2 所顯示。

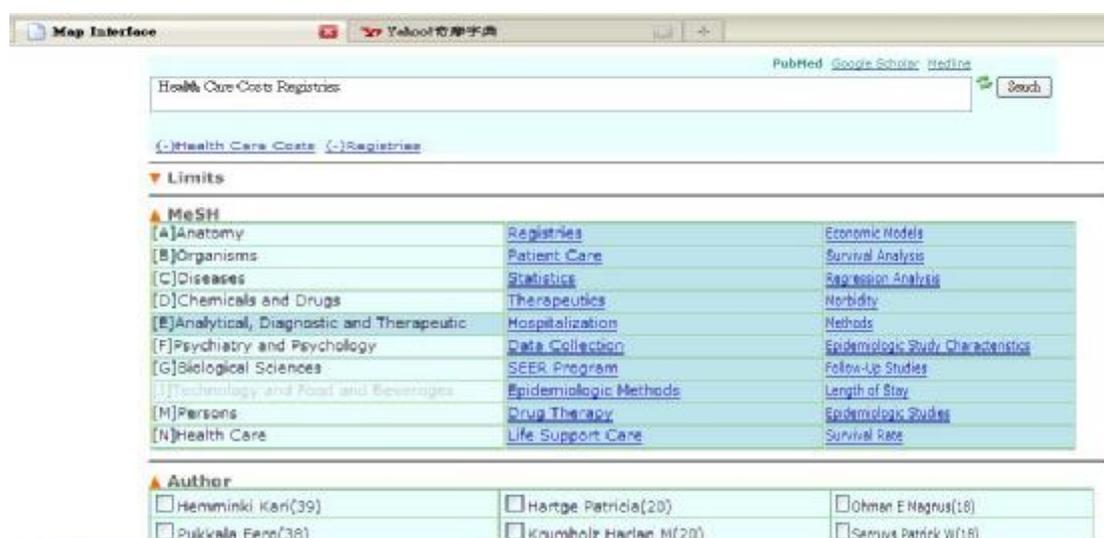


圖 3-2 MAP 介面之建議詞彙呈現

MAP 介面之建議詞彙清單的顯示方式有兩種，一個是簡單排序的清單形式且未經過任何形式的分類，而另一種則是以 MeSH 組織進行層面分類，其排序的方式是由其與查詢關鍵字之共現詞彙頻率與倒置詞彙頻率相乘後以降冪排序。

透過 MAP 的關鍵字推薦功能，使用者可輕易於「MeSH」欄位中獲取與輸入關鍵字相關的詞彙，提供他們不同面向與深度的詞彙選擇。而「Author」與「Journal」兩個欄位亦將即時回饋使用者與該關鍵字主題相關的期刊名與作者名，

使用者可再以期刊或作者的角度進行進一步檢索。此外，本介面不僅幫助使用者檢索 PubMed 與 Medline 資料庫，亦可運用於 Google Scholar 系統上；使用者可依個別需求調整檢索對象。

(二)實驗設計

本研究以重複量測(repeated measure)的方式設計實驗，受測者被要求以兩個介面(PubMed 及 MAP)檢索問題，即每一檢索問題都要求受測者交互使用兩個介面進行檢索。重複量測可以控制因為檢索問題(request)不同所產生的變異量(variance)並加強統計檢定的力量。但此一設計最明顯的風險在於次序效應會影響實驗結果，即前一介面之檢索結果會影響後一介面之檢索結果。為控制可能發生的次序效應，受測者的檢索需求會被隨機指派到替代的處理順序，而使得每一個檢索需求都有均等的機會以任何介面進行第一次檢索。受測者將被隨機指派到四者其中之一進行檢索，然後再轉移到另一介面。為降低次序效應，同一需求中的處理並非立即接續，每一處理之間均有一段中斷的時間，讓受測者能夠以較不受前次處理影響的情境來進行介面的操作。

雖然透過隨機指派替代的處理順序，次序效應仍存在於前一介面所提取的詞彙給與當前之介面過多的優勢。尤其在以 MAP 介面作為第一次介面時，受測者將可瀏覽較多的詞彙，而在受測者使用 PubMed 標準介面進行第二次檢索時，可能會大量利用這些詞彙，因此必須要對受測者說明使用相同的查詢進行檢索，且不要使用在前一介面所得到詞彙。

實驗之分組如表 3-1 所示，共有四組，以介面之差異與主題熟悉度作為控制，受測者將隨機指派至四個類別中，先以單一介面檢索不熟悉主題及熟悉主題，而後以另一介面進行不熟悉主題及熟悉主題之檢索，其目的在於比較不同主題熟悉度及不同介面對於使用者的檢索行為及查詢擴展行為所產生的差異。

表 3-1 實驗分組表

類型	PubMed		MAP	
1	熟悉主題	不熟悉主題	熟悉主題	不熟悉主題
2	不熟悉主題	熟悉主題	不熟悉主題	熟悉主題
	MAP		PubMed	
3	熟悉主題	不熟悉主題	熟悉主題	不熟悉主題
4	不熟悉主題	熟悉主題	不熟悉主題	熟悉主題

(三)工作任務

本研究將探索以使用者的真實需求情境而非研究者所指定的檢索任務為基本的分析單位(unit analysis)的可行性。如此設計的原因有二：一為如果採用指定檢索任務，在描述任務時將難免給與受測者檢索詞彙的提示，如此一來將與研究介面所欲提供的詞彙建議功能相抵消，受測者或許會過於仰賴任務描述中有限的詞彙，失去對瀏覽式介面的需求。即使已有研究顯示設計良好的語意開放的任務描述即能夠讓參與者融入假想的需求情境中，並足以獲得與真實情境相當的評估結果(Borlund & Ingwersen, 1997, 2000)，然而在對於醫學如此具備高度專業知識的領域而言，設計具有開放語意的任務描述是有困難的。

第二個採取真實檢索問題的原因在於本研究欲找出檢索問題與工具間的對應性。在本研究中研究者將有機會觀察、紀錄不同檢索工具在不同搜尋情況下的效用。若採取使用研究者所假想的檢索任務模式，則必須將不同的搜尋情境用敘事的方式操作化，否則將喪失研究兩者相容性的機會，因此本研究將以受測者的真實資訊需求做為檢索任務。

二、無干擾觀察法

為了能夠真實記錄受測者操作之行為，利用螢幕影響擷取軟體 Morae®記錄

自受測者開啟瀏覽器進行實驗起，包含頁面之點選、相關紀錄之點選、關鍵字之刪改、系統功能之操作等活動，完整錄製受測者利用介面之過程，其範圍為受測者提供之兩個檢索問題所進行的四次操作過程。

三、內容分析法

分析以無干擾觀察法紀錄之受測者螢幕畫面，將受測者所操作之關鍵字之新增、刪改之動作進行編碼，並分析其修改之關鍵字與其初始查詢之間的語意關係。

四、問卷法

問卷包含背景資料問卷、檢索前問卷及檢索後問卷以及實驗後問卷。背景資料問卷詢問受測者的背景資料，如職業、性別、專業領域、使用 PubMed 資料庫及 Google Scholar 之頻率、對於 MeSH 之熟悉度、使用進階檢索指令的頻率等。檢索前問卷則是詢問受測者欲檢索之問題以及對該議題之熟悉度等。檢索後問卷則詢問受測者對於檢索結果、介面的滿意度等。實驗後問卷則是詢問介面對於受測者的影響，包含最佳檢索詞彙、熟悉程度等。前測之問卷詳如附錄一，經過前測後刪去不適宜之問項後形成正式問卷，詳如附錄二。

第二節 研究對象與研究工具

一、研究對象

由於 PubMed 資料庫屬於醫學資料庫，使用者多是具有醫學相關的專業知識且基於特定專業的資訊需求而檢索該資料庫，故本研究之研究對象為生命科學及醫學相關系所的專業人士而非一般的普通使用者。而研究對象之取得則是於 BBS 及校內訊息公佈欄網頁張貼告示，公開徵求生命科學及醫學相關系所之研究生，並以新臺幣 500 元做為參與實驗之報酬。

二、研究工具

資料蒐集工具為：檢索前問卷、檢索後問卷、實驗後問卷、螢幕影像紀錄軟體。

分析工具則使用 SPSS 進行量化數值的比較。

螢幕影像紀錄軟體 Morae®則為 Techsmith 公司出品的可用性測試軟體，它可以記錄使用者實際操作中的行為，及對於特定應用程式和電腦的詳細操作數據。本研究應用此軟體記錄使用者在使用 MAP 及 PubMed 介面時的操作行為，自第一次檢索至第四次檢索中完整的影像紀錄，同時以高亮度標示使用者的滑鼠位置以指示使用者的注意位置。

第三節 資料蒐集與分析

本研究之資料來源分別為實驗前後要求受測者所填寫之問卷以及實驗過程之螢幕影像紀錄。據此，欲採取以下分析方法：

一、系統功能之分析

1. 分析初始查詢與最終查詢之相似性，比較不同介面重組(reformulation)使用者初始查詢的能力。
2. 分析 Map 介面之檢索結果在 PubMed 介面之檢索結果中的位置，瞭解 MAP 介面是否能夠有效挖掘原本埋於 PubMed 檢索結果深處之資料。
3. 記錄受測者對詞彙建議系統之評價，並進行分析。

二、受測者詞彙選擇行為之分析

Efthimiadis (2000)的研究將受測者選擇建議之詞彙之原因分成1.相關詞彙；2.同義詞；3.妥協詞(no better term)；4.新概念詞，而其與初始查詢的語意關係分成 1.無關係；2.下位詞；3.上位詞；4.相關詞。而 Shiri & Revie (2006)除了分析受

測者所選擇之詞彙與初始詞彙之關係外，亦透過計算初始查詢之詞彙數及布林邏輯運算元數量，做為任務複雜度之類別。據此本研究對於受測者詞彙選擇行為欲進行以下分析：

1. 就螢幕錄影資料進行編碼，記錄受測者所選擇及刪改的詞彙。
2. 計算受測者初始查詢詞彙、新增詞彙、刪除詞彙之數量，並比較不同介面之詞彙選擇行為之間的差異。
3. 針對受測者之檢索進行內容分析，目的在於分辨其檢索詞彙之類別，並就不同介面及不同熟悉度比較詞彙選擇行為是否有明顯差異。

第四節 研究實施與步驟

一、實驗步驟

當受測者抵達時，會先徵得其同意實驗，同時要求其填寫一問卷收集其專業知識、教育狀態及 PubMed/Medline 使用經驗。此外提供一影片教學來解釋如何操作介面以讓使用者熟悉介面。

在每個檢索開始前，要求使用者寫下搜尋前問卷說明其欲檢索的任務。同時寫下其所認為最適合的查詢。尺度資料可以取得受測者對於檢索問題的熟悉度及檢索結果的詳盡性的量化指標。

每一受測者會用兩個介面交互檢索兩個查詢，因此會有四個階段。當每一檢索階段開始時，受測者輸入其在搜尋前問卷所寫下的查詢，並要求其用指派的介面在該階段檢索 10 筆有用的紀錄。在最初輸入之後，受測者可以透過介面的回饋及檢索結果來修改其查詢。受測者被告知每一任務均有 15 分鐘去完成，但可以提前完成。

每一階段結束後，受測者會寫下這次其所認為最適合的查詢詞。並且以六點尺度要求其評估檢索前及檢索後的理想查詢詞的適用性(goodness)。另外亦要求受測者評估對於檢索結果的滿意度及介面的好用程度。所有受測者與介面的互動均以螢幕錄製軟體所記錄。本研究特別關注於重複查詢的次數、選擇關鍵字的數量、檢視的紀錄數，此有助於理解使用者可能如何與不同的系統互動。

在同一查詢在兩個介面都完成後，受測者被要求用六點尺度挑出在先前所找到的20筆紀錄的相關程度(每一介面10筆)。如此則可比較兩個介面的檢索品質。其它的指標包含使用者對於查詢所表達的檢索之評估，即對於檢索結果的滿意度。同時要求受測者對詞彙建議功能進行評論同時說明在何種蒐尋情況下該功能會有所幫助。

具體實驗流程如下：

1. 於檢索開始使先請受測者填寫檢索前問卷，紀錄其背景資料及所欲檢索的問題及使用之關鍵字。
2. 檢索開始後使用螢幕記錄軟體錄製受測者的檢索過程，每一檢索均耗時15分鐘。
3. 檢索完畢後請受測者填寫檢索後問卷，記錄其認為最適合的關鍵字，同時提出介面評價。
4. 每一受測者均有兩個檢索問題，而每一檢索問題均透過MAP及PubMed介面進行兩次檢索。
5. 檢索順序的安排將確保受測者不會馬上重覆檢索同一問題，以期降低實驗的次序效應。
6. 結束實驗。

7. 填寫實驗後問卷，記錄受測者對於詞彙建議系統之評價。

二、研究步驟

本研究之研究步驟如圖 3-3 所示：

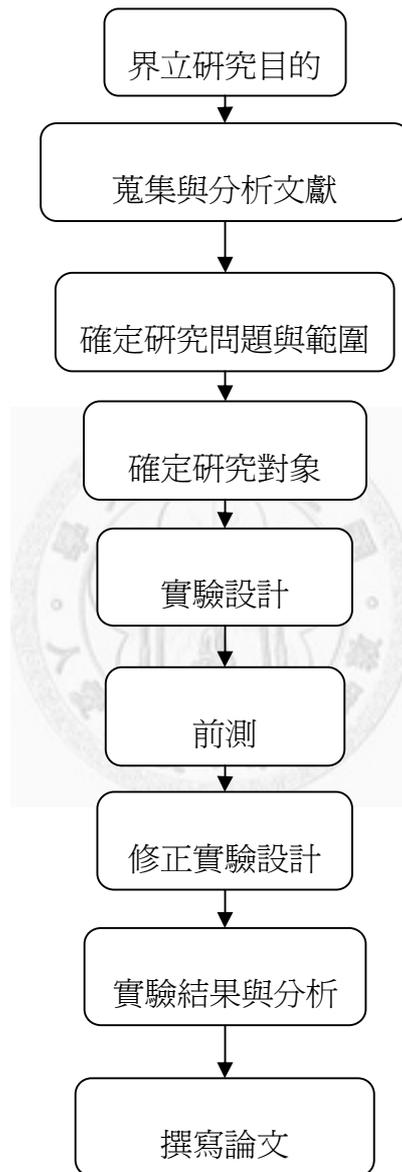


圖 3-3 研究步驟圖

第四章 研究結果與發現

本研究旨在了解使用者利用詞彙建議系統之查詢擴展行為與利用傳統介面查詢之間的差異，並分析詞彙建議系統在何種情境下對使用者較有幫助。本章共分七節，第一節說明使用者基本資料，第二節分析使用者在不同介面及不同主題熟悉度下之檢索特性，第三節分析不同介面及不同主題熟悉度下使用者查詢擴展之行為，第四節說明 MAP 介面輔助使用者檢索排序較低的相關文獻之能力，第五節解釋因實驗設計所產生的次序效應，第六節分析使用者評價以說明 MAP 介面在何種情境下能夠發揮較大的功用。

第一節 使用者基本資料

本實驗之受測者經由台灣大學電子佈告欄系統(Bulletin Board System,BBS)批踢踢實業坊公開徵求相關生命科學及醫學相關系所之研究生參與，共募集 44 名受測者，其就讀學院及性別等資料如表 4-1 所示，受測者除 1 人為研究助理外其餘均為研究生。系所之學院分類參考國立臺灣大學教學單位之分類¹，如表 4-2。

為符合實驗之目的，即測試 PubMed 資料庫之一般使用者，因此將受測者之背景限制於相關生醫理工系所，其中之例外分別為電機所及特殊教育研究所，故分於其他類之中。

¹ 參考自 <http://www.ntu.edu.tw/chinese2010/academics/academics.html>

表 4-1 受測者基本資料表

學院	性別		總和
	男	女	
醫學院	3	6	9
公衛學院	4	2	6
生命科學學院	3	4	7
理學院	2	2	4
工學院	4	3	7
農學院	6	3	9
其他	1	1	2
總和	23	21	44

表 4-2 學院分類表

學院別	系所
醫學院	生理學研究所
	微生物與生化學研究所
	醫學檢驗暨生物技術研究所
	獸醫所
	藥學研究所
	公共衛生
公衛學院	公共衛生學研究所
	流行病學研究所
	衛生福利研究所
	環境衛生研究所
生命科學學院	分子細胞所
	生化科技所
	生命科學研究所
	植物科學研究所
理學院	心理所
	認知與神經科學研究所
工學院	化學工程學研究所
	醫學工程學研究所
農學院	生物工程所

	生物資訊研究所
	動物學研究所
	動物學研究所
	農業化學研究所
	營養學研究所
其他	特殊教育研究所
	電機研究所

除受測者就讀系所及性別資料外，實驗前問卷並蒐集受測者利用 PubMed 及 Google Scholar 的頻率與對於 MeSH 醫學索引典的熟悉程度和檢索時利用布林邏輯進行進階檢索的頻率，要求受測者以李克特 6 點尺度進行選擇，0 表示程度最低，5 表示程度最高，其結果如表 4-3 及圖 4-1。結果顯示受測者在利用 PubMed 及 Google Scholar 的頻率大致相近，而對於 MeSH 醫學索引典及布林邏輯檢索則較不熟悉。由是觀之，可以發現一般使用者並不熟悉檢索系統所提供之進階檢索功能，對於相關之資訊組織工具 MeSH 也缺乏理解，在進行檢索時仍只會以關鍵字組合進行檢索。而透過 MAP 介面進行檢索，系統能夠提供使用者相關的 MeSH 敘述詞，進一步的呈現領域內的知識體系，使用者即便不熟悉 MeSH 之詞彙亦可用點選的方式利用 MeSH 中豐富的詞彙進行檢索。

表 4-3 受測者檢索背景資料

	平均數	標準差	個數
使用PubMed資料庫之頻率	3.68	0.86	44
使用Google Scholar之頻率	3.39	1.59	44
MeSH醫學索引典之熟悉程度	1.07	1.50	44
使用布林邏輯等進階檢索功能之頻率	2.52	1.58	44

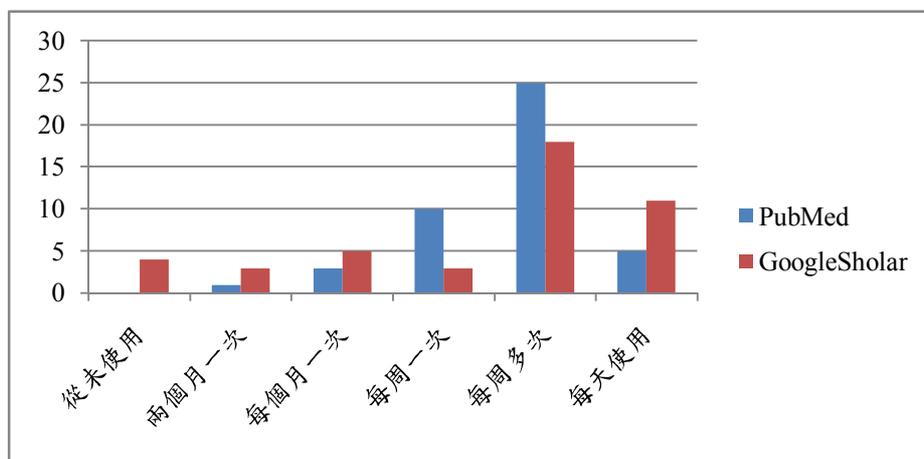


圖 4-1 受試者利用 PubMed 及 Google Scholar 頻率圖

本實驗採用使用者之真實資訊需求做為檢索任務，每一使用者提供熟悉與不熟悉各1個檢索任務，共有88個檢索任務，並在PubMed及MAP各進行一次檢索，總計176次檢索。在88個檢索任務中受測者未曾檢索之主題為28個，實驗前已檢索過之主題為60個。受試者之檢索任務詳如附錄三。

本研究針對主題熟悉度此一變數採用兩種方式測量，其一以六點尺度由受試者評估其對於檢索主題之熟悉度。其二為類別變數，以受試者是否檢索過該檢索問題做為分類之基準。尺度資料用以分析整體使用者對於檢索主題的熟悉程度並與其他變數進行相關檢定，而類別變項則是以不同介面與不同主題熟悉度針對受試者之檢索行為進行anova變異數分析，並用於比較不同類別變數對於查詢擴展行為的影響。

在檢索前問卷的部分則要求受測者填寫欲檢索之問題以及查詢關鍵字，而後詢問其受試前是否檢索過該主題，並要求受測者以李克特六點尺度，0表示程度最低，5表示程度最高，說明其對檢索主題之熟悉程度、對於檢索結果完整度的要求以及由受測者評估其問題可以由他人代為查詢的取代程度。由於缺乏一能夠客觀評估查詢(Query)之指標，因此設計查詢「適用性」(goodness)指標，由受測

者於檢索後問卷中以李克特六點尺度，0表示最低，5表示最高，主觀評估其初始查詢以及最終查詢能夠表達其資訊需求的程度，其結果如表4-4，其相關程度如表4-5。

結果顯示主題熟悉度與檢索結果之完整度呈現正相關.550($p=.000$)。受測者對檢索主題愈為熟悉，由於對檢索系統中的相關文獻已有所認知，因此對檢索結果之完整性也要求更高。而在受測者在未曾檢索主題的情況下，由於對主題相對不熟悉，因此不知檢索結果之完整程度，所以較不要求檢索結果之完整性。

此外主題熟悉度與初始查詢適用性亦呈正相關.250($p=.019$)，受測者對於檢索主題越熟悉，其初始查詢愈能夠完整陳述其資訊需求。而初始查詢適用性與最終查詢適用性亦成正相關.256($p=.016$)，顯示最終查詢適用性會受到初始查詢適用性的影響，因此在評估介面時必須控制檢索主題之變異。而主題熟悉度未與最終查詢適用性產生相關，其原因可能在於受試者經過介面檢索後，降低了主題熟悉度對於查詢適用性的影響。

在受測者評估他人取代查詢方面，受測者在面對不熟悉主題時，對於檢索結果之相關判斷會較為寬鬆，因此對於他人代為檢索之結果其相關判斷亦趨於簡單，所以受測者會較容易認為他人之檢索結果與己身之需求相關，是故會認為他人能夠取代己身之檢索。但在已檢索過主題的情境下，受測者對於檢索結果之相關文獻已有進一步認知，其相關判斷亦趨嚴苛，故他人代為檢索之結果反而無法滿足其資訊需求。Pennanen 與 Vakkari (2003)的研究亦指出對於主題較不熟悉且的使用者能較熟悉主題者發現更多的參考文獻，其原因在於不熟悉主題者無法分辨相關與不相關文獻之間的差異，因此會較容易接受文件為相關。

表4-4 檢索主題特性

	檢索與否	平均數	標準差	N
檢索主題熟悉度	N	1.50	1.20	28
	Y	4.50	1.55	60
檢索結果完整度	N	3.07	1.76	28
	Y	4.55	1.06	60
代為檢索之取代程度	N	3.04	1.32	28
	Y	2.80	1.44	60
初始查詢適用性	N	3.61	1.52	28
	Y	3.95	1.20	60

表4-5 檢索主題特性相關程度

Spearman's rho	熟悉度	完整度	初始適用	最終適用
熟悉度				
完整度	.550(**)			
初始適用	.250(*)	.098		
最終適用	.070	.011	.256(*)	

*p<.05 ** p<.01

第二節 使用者檢索特性分析

受測者之檢索特性是指受測者之查詢各個性質，包含查詢相似度、最終查詢、新增詞彙、移除詞彙之字數以及送出查詢之次數和瀏覽文件數，經由統計分析以了解在初次檢索及重複檢索和不同介面上其檢索特性是否有明顯差異。

查詢相似度是指使用者最終查詢與初始查詢之相似程度，以 Jaccard 係數 (Jaccard's Coefficient) 進行運算，查詢之相似度為 $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$ ， $A \cap B$ 為初始查詢字數減去移除字數， $A \cup B$ 為初始查詢字數以及擴充查詢字數減去 $A \cap B$ ，比較兩種介面之查詢相似度能夠說明 MAP 介面是否較傳統 PubMed 介面較能改變受測者之查詢。最終查詢、新增詞彙、移除詞彙之字數則是受測者在與檢索系統

互動後所產生之明顯變化，其可說明受測者經過實驗後所受到的影響程度。送出查詢之次數則表示受測者改變其檢索之次數，瀏覽文件數則是指受測者在檢索結果中所點閱的頁面之數量。

一、 重複檢索與初次檢索之檢索特性比較

重複檢索與初次檢索之檢索特性如表 4-6，經過獨立樣本平均數檢定之後，結果顯示重複檢索與否在查詢相似度有明顯差異($t=3.82, p=.000$)，重複檢索之查詢相似度($m=.53$)明顯高於初次檢索($m=.35$)，此一現象說明了使用者在熟悉主題的情境下，由於認知結構相對穩定，其查詢之關鍵字較在與資訊檢索系統互動後較不易改變。反之，在不熟悉主題的情境下，由於使用者對於檢索主題缺乏完整的認知，因此在與資訊檢索系統互動後較會改變其查詢關鍵字而進一步重構其資訊需求表述。

在移除詞彙數的部分，初次檢索之移除詞彙數($m=.96$)明顯高於重複檢索($m=.59$) ($t=2.323, p=.021$)，此一結果說明了在不熟悉主題的情境下，使用者的初始查詢是較為不固定的，因此在與資訊檢索系統互動後會刪除初始查詢中較不符合其資訊需求之詞彙。

而重複檢索與否雖然對於新增詞彙數($t=1.858, p=.065$)並未產生明顯影響，但是在不熟悉的情境下，由於使用者資訊需求較不明確，因此在與資訊檢索系統互動後其新增詞彙數($m=2.45$)會高於熟悉情境下資訊需求較為穩定者($m=1.89$)。

至於在最終查詢字數方面($t=-.235, p=.815$)，重複檢索與否兩者之間並無明顯之差異，此即是使用者之最終查詢之字數並不會受到主題熟悉與否所影響。而查詢次數($t=.151, p=.880$)則說明了不論對於主題熟悉與否，使用者仍會經過一定次數測試後，才會決定其認為最適合的查詢詞彙。而瀏覽文件數($t=-.155, p=.877$)則顯示不論對於主題熟悉與否，使用者仍會瀏覽一定數量之檢索結果頁面才會結束

其查詢。

表 4-6 新舊查詢之檢索特性比較

	重複查詢與否	平均數	標準差	個數
相似度 ***	新	0.35	0.26	56
	舊	0.53	0.30	120
新增詞彙數	新	2.45	1.86	56
	舊	1.89	1.84	120
刪除詞彙數 *	新	0.96	1.19	56
	舊	0.59	0.88	120
最終查詢詞彙數	新	3.96	2.12	56
	舊	4.04	1.99	120
瀏覽文件數	新	90.09	51.42	56
	舊	91.63	65.98	120
查詢次數	新	4.55	2.98	56
	舊	4.48	2.82	120

*p<.05

***p<.000

二、不同介面之檢索特性比較

不同介面之檢索特性如表 4-7，經過獨立樣本平均數檢定之後，結果顯示不同介面之查詢相似度平均數有明顯差異($t=3.313, p=.001$)，PubMed 介面($m=.54$)之相似度高於 MAP 介面($m=.40$)，此結果說明了使用者利用 PubMed 介面時，由於僅能透過與檢索結果的互動來產生新的檢索詞彙，因此其最終查詢與初始查詢的相似度會較高，而在 MAP 介面時，由於介面提供了大量的相關建議詞彙，因此使用者能從中獲得較多的檢索詞彙並透過檢索驗證其相關性而置入最終查詢詞彙中，因此其查詢相似度會較低。

此外，不同介面之新增詞彙數亦有明顯差異($t=-3.705, p=.000$)，MAP 介面($m=2.58$)明顯高於 PubMed 介面($m=1.57$)。由於使用者在利用 PubMed 介面檢索時，其查詢詞彙僅能從與檢索結果互動中所產生，反之 MAP 介面能夠提供大量的相關 MeSH 詞彙供使用者點選利用，因此兩者之新增查詢詞彙數會有明顯之

不同。

不同介面在最終查詢字數方面亦有明顯差異($t=-2.370, p=.019$)，MAP 介面之最終查詢字數($m=4.38$)明顯高於 PubMed 介面($m=3.66$)。這亦可說明透過 MAP 介面所提供的大量相關詞彙，使用者可以不受己身對於檢索主題理解的限制，從中取用合適的查詢詞，因此其最終查詢詞彙字數會較 PubMed 介面來的高。

而不同介面在查詢次數具有明顯差異($t=-3.683, p=.000$)，MAP 介面之查詢次數($m=5.27$)明顯高於 PubMed 介面($m=3.74$)。MAP 介面雖然能夠提供使用者大量的相關詞彙，然而此一相關判斷是透過系統所產生，其對於主題之相關性仍須經由受測者檢索後方能決定。MAP 介面之查詢次數高於 PubMed 介面的原因在於使用者必須透過個別的查詢來驗證 MAP 介面所提供的建議詞彙是否能夠滿足其資訊需求。

在移除詞彙數($t=-1.431, p=.154$)方面，不同介面間並無明顯差異。使用者移除之詞彙是以初始查詢作為比較的對象，而使用者移除詞彙主要是受到檢索結果的影響。透過關鍵字查詢，使用者可從檢索結果中判斷該關鍵字是否與其主題相關，並決定其是否適合查詢。因此在 2 個介面之中，使用者均是透過檢索結果來判斷查詢詞彙之適用與否，因此兩者間的數量並未產生明顯差異。

於瀏覽文件數($t=-.326, p=.745$)方面，不同介面間並無明顯差異。此結果顯示使用者無論是利用何種介面進行檢索，其仍須瀏覽一定數量的檢索結果，方認定其資訊需求得到滿足，而後結束其檢索。

表 4-7 不同介面之檢索特性比較

	介面	平均數	標準差	個數
相似度	PubMed	.54	.31	88
***	Map	.40	.27	88
新增詞彙數	PubMed	1.57	1.54	88
*	Map	2.58	2.01	88

刪除詞彙數	PubMed	.60	.90	88
	Map	.82	1.09	88
最終查詢詞彙數	PubMed	3.66	1.91	88
	Map	4.38	2.09	88
查詢次數	PubMed	3.74	2.74	88
	Map	5.27	2.78	88
瀏覽文件數	PubMed	89.63	60.10	88
	Map	92.66	63.32	88

*** p<.001
*p<.05

三、不同介面與重複搜尋與否對檢索特性之交互作用

經雙因子變異數分析，檢驗重複檢索與否和不同介面對於各個檢索特性之交互作用，結果如表 4-8，其平均數及標準差如表 4-9。

結果顯示在查詢相似度($F(1,172)=.61, p=.43$)，其平均數圖如圖 4-2，不同介面及重複檢索與否雖未呈現明顯交互作用，但可發現在 PubMed 中的重複檢索其查詢相似度最高($m=.59$)，MAP 介面中的全新檢索最低($m=.25$)，此一現象可說明在 MAP 介面在使用者於不熟悉的情境下，由於提供了大量的相關詞彙，因此能夠大幅度的重構使用者之資訊需求，所以其查詢相似度較低。反之使用者在熟悉情境下，操作 PubMed 介面時由於接受之刺激有限，因此其查詢相似度呈現最高。

在新增詞彙數方面($F(1,172)=.39, p=.533$)，其平均數圖如圖 4-3，兩者雖無明顯交互作用，但可發現在 PubMed 中的重複檢索其新增詞彙數最少($m=1.45$)，MAP 介面中的全新檢索最多($m=3.07$)，此一結果可以說明在不熟悉情境下，使用者會透過 MAP 介面提供之大量的相關建議詞彙進行檢索，重構其資訊需求。反之，在熟悉主題之情境下，使用者之查詢詞彙受到己身既有知識架構影響，且與僅能與檢索結果產生有限的互動，因此其新增查詢詞彙數會較低。

在刪除詞彙數($F(1,172)=.42, p=.52$)中，其平均數圖如圖 4-4，兩者並無明顯交互作用，但可發現 PubMed 中的重複檢索其移除詞彙數最少($m=.52$)，MAP 介面

中的全新檢索最多($m=1.14$)。在熟悉的情境下，使用者資訊需求較為穩定，且在 PubMed 介面中僅能與查詢結果進行互動，因此其刪除詞彙數會較低。而在不熟悉的情境下，使用者初始查詢關鍵字之相關性必須透過檢索後才能判斷，且受到 MAP 大量的建議詞彙所刺激，因此會大量修改初始查詢之詞彙，因此其移除詞彙數則較高。

而在最終查詢詞彙數($F(1,172)=.10, p=.751$)，其平均數圖如圖 4-5。兩者雖未見明顯交互作用，但可發現在 PubMed 中的全新檢索其最終查詢詞彙數最少($m=3.54$)，MAP 介面中的重複檢索最多($m=4.42$)。在於不熟悉之情境下，使用者缺乏足夠的領域內知識去構成其資訊需求，且在 PubMed 介面中只有檢索結果能夠給予使用者刺激，因此其查詢之新增關鍵字極其有限，是故其最終查詢字數亦有限。而在熟悉的情境下，透過 MAP 介面提供的大量建議詞彙，使用者能從中判斷初始之查詢詞彙是否合適，且透過點選檢索即可知查詢詞彙之適用性，因此其最終查詢詞彙數會高於其他情況。

在瀏覽文件數($F(1,172)=.48, p=.49$)方面，其平均數圖如圖 4-6。在熟悉情境下，受測者之瀏覽文件數並無差異，但在不熟悉情境下使用者在 PubMed 介面中其所瀏覽之文件數($m=83.82$)明顯低於 MAP 介面($m=96.36$)。因此推論不論在不熟悉的情境下，使用者均會檢驗數量相當的檢索結果以判斷是否滿足其資訊需求，因此其瀏覽文件數並無明顯差異。反之在熟悉的情境下，使用者在 PubMed 介面中由於對於主題內相關資訊已有所認知，因此其瀏覽檢索結果之過程會較為簡略，而在 MAP 介面中，由於建議詞彙已改變檢索結果之結構，因此使用者會瀏覽較多文件以判斷其資訊需求是否得到滿足。

至於查詢次數($F(1,172)=.49, p=.48$)，其平均數圖如圖 4-7，兩者亦無明顯交互作用，但可發現在 PubMed 介面中全新檢索其查詢次數最低($m=3.57$)，MAP 介面中全新檢索最高($m=5.54$)。在不熟悉的情境下，使用者在 PubMed 介面時僅能與

檢索結果互動，其查詢之重構亦有限，因此不須透過多次的檢索來驗證查詢詞彙之適合性，反之在 MAP 介面中，使用者由於能夠透過介面提供的大量檢索詞彙來重構其資訊需求，因此必須透過多次的檢索來驗證其查詢詞彙之適合性，因此其查詢次數會較高。

表 4-8 重複檢索與介面對查詢行為各變數之二因子變異數分析摘要

變數	變異來源	自由度	平方和	平均平方和	F 檢定
相似度	介面	1	0.96	0.96	12.19**
	重複檢索	1	1.22	1.22	15.45***
	介面 * 重複檢索	1	0.05	0.05	0.61
	誤差	172	13.58	0.08	
	總和	176	54.52		
新增詞彙數	介面	1	43.56	43.56	13.69***
	重複檢索	1	11.60	11.60	3.65
	介面 * 重複檢索	1	1.24	1.24	0.39
	誤差	171	544.03	3.18	
	總和	175	1350.00		
移除詞彙數	介面	1	2.46	2.46	2.51
	重複檢索	1	5.30	5.30	5.41*
	介面 * 重複檢索	1	0.41	0.41	0.42
	誤差	172	168.46	0.98	
	總和	176	265.00		
最終查詢字數	介面	1	21.68	21.68	5.34*
	重複檢索	1	0.23	0.23	0.06
	介面 * 重複檢索	1	0.41	0.41	0.10
	誤差	172	697.76	4.06	
	總和	176	3561.00		
瀏覽文件數	介面	1	1183.68	1183.68	0.31
	重複檢索	1	91.03	91.03	0.02
	介面 * 重複檢索	1	1853.77	1853.77	0.48
	誤差	172	661167.60	3844.00	
	總和	176	2125527.00		
檢索次數	介面	1	103.80	103.80	13.48***
	重複檢索	1	0.19	0.19	0.02

介面 * 重複檢索	1	3.80	3.80	0.49
誤差	172	1324.45	7.70	
總和	176	5005.00		

*P<.05 **P<.01 ***P<.000

表 4-9 重複檢索與否及不同介面交互作用之查詢行為

	介面	重複檢索	平均數	標準差	個數
相似度	PubMed	全新檢索	0.44	0.29	28
		重複檢索	0.59	0.31	60
	Map	全新檢索	0.25	0.19	28
		重複檢索	0.46	0.28	59
新增詞彙數	PubMed	全新檢索	1.82	1.70	28
		重複檢索	1.45	1.45	60
	Map	全新檢索	3.07	1.82	28
		重複檢索	2.34	2.08	59
刪除詞彙數	PubMed	全新檢索	0.79	1.03	28
		重複檢索	0.52	0.83	60
	Map	全新檢索	1.14	1.33	28
		重複檢索	0.66	0.94	59
最終查詢詞彙數	PubMed	全新檢索	3.54	2.03	28
		重複檢索	3.72	1.87	60
	Map	全新檢索	4.39	2.17	28
		重複檢索	4.42	2.04	59
瀏覽文件數	PubMed	全新檢索	83.82	46.11	28
		重複檢索	92.33	65.80	60
	Map	全新檢索	96.36	56.37	28
		重複檢索	91.44	67.16	59
查詢次數	PubMed	全新檢索	3.57	2.74	28
		重複檢索	3.82	2.76	60
	Map	全新檢索	5.54	2.92	28
		重複檢索	5.20	2.72	59

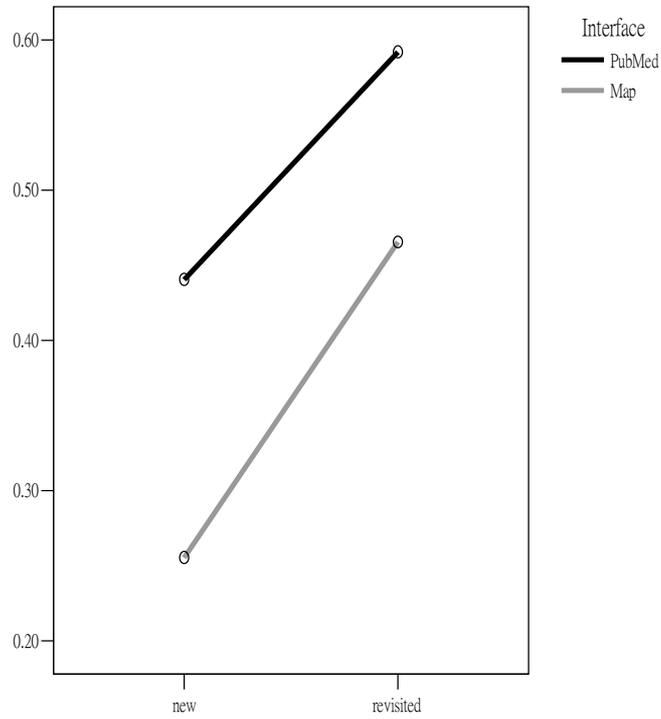


圖 4-2 不同介面及不同熟悉度之相似度平均數圖

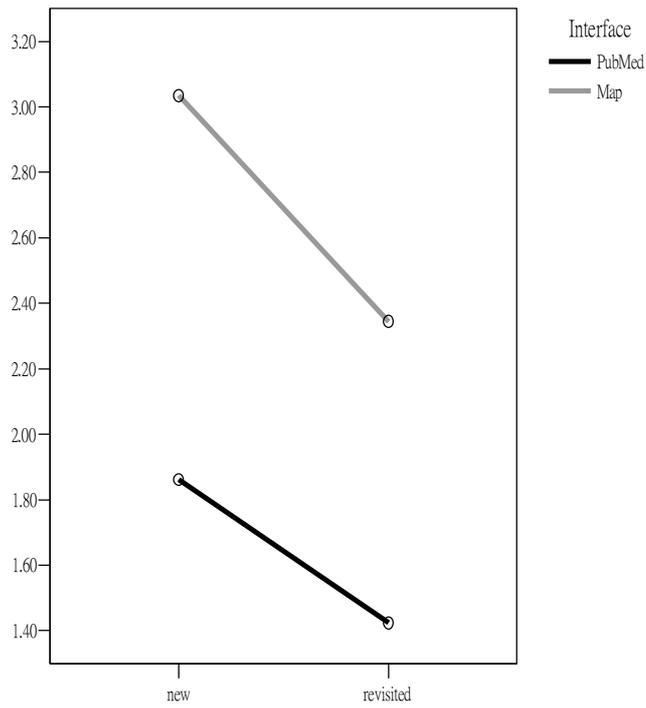


圖 4-3 不同介面及不同熟悉度之新增詞彙數平均數圖

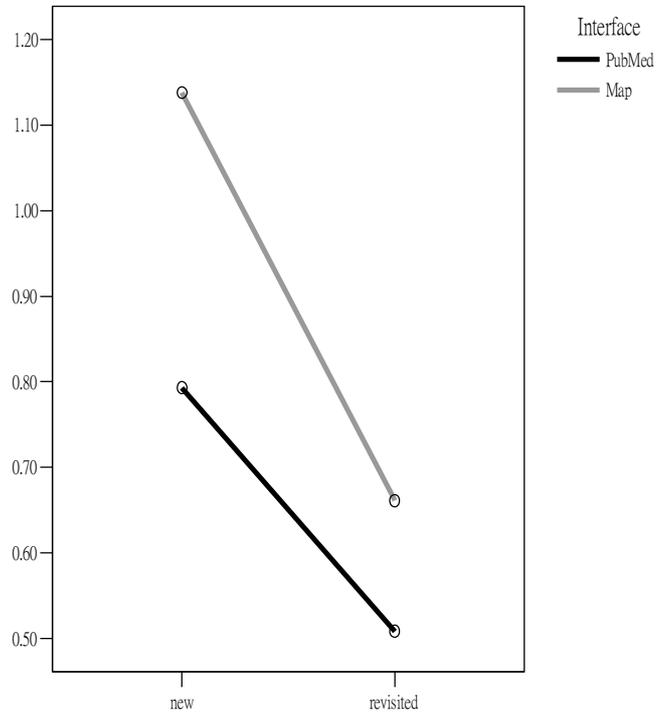


圖 4-4 不同介面及不同熟悉度之移除詞彙數平均數圖

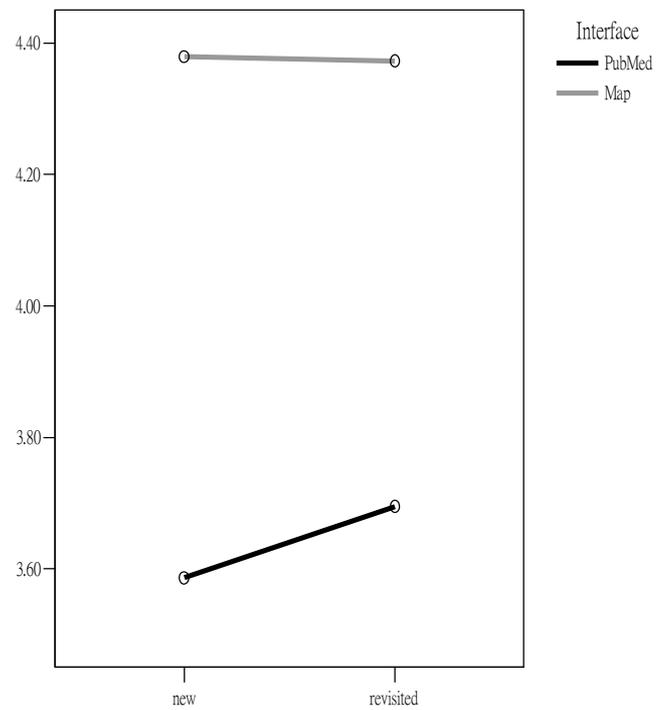


圖 4-5 不同介面及不同熟悉度之最終查詢字數平均數圖

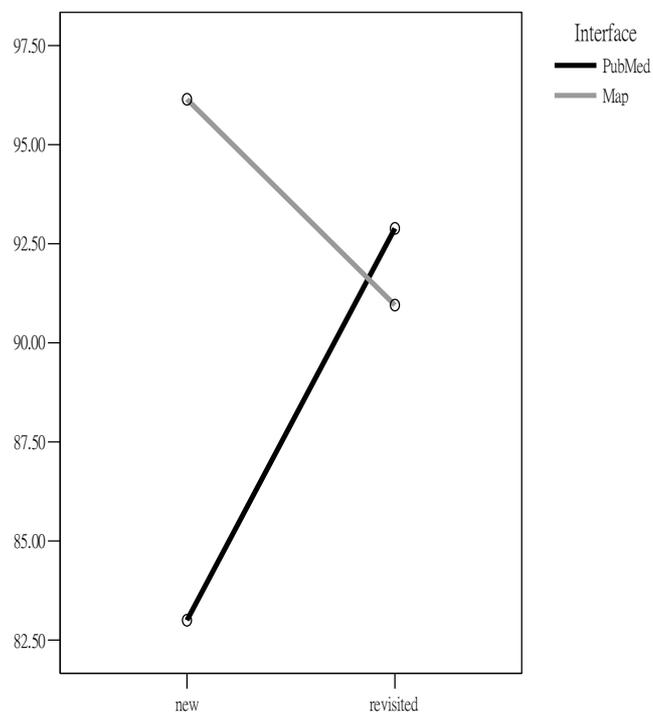


圖 4-6 不同介面及不同熟悉度之瀏覽文件數平均數圖

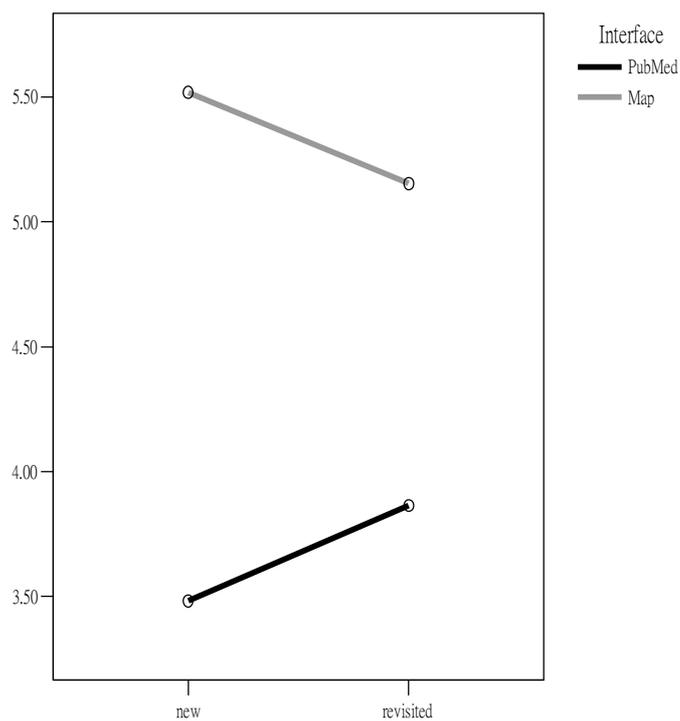


圖 4-7 不同介面及不同熟悉度之查詢次數平均數圖

第三節 使用者查詢擴展行為

本節關注於使用者操作介面時所使用之查詢詞彙，主要分成兩個部分，其一為分析使用者於操作 PubMed 時在檢索過程中輸入的詞彙以及利用 MAP 介面時所選用之建議詞彙。其二為在實驗後其所認知的最佳查詢詞彙。兩者均依據 MeSH 醫學主題詞表進行分類

本研究依據 MeSH 醫學主題詞表之結構，將受測者之查詢擴展詞彙分成 11 類，如表 4-10。其他類是為研究者在以 MeSH 層面進行詞彙分類時，無法斷定該詞彙之 MeSH 分層。如使用者之查詢擴展詞彙為 MeSH 醫學主題詞表中的限定詞(Qualifier)，無所斷定來自於某一特定的 MeSH 分層，或是使用者之查詢擴展詞彙並未包含於 MeSH 詞彙之中。此外，若使用者在實驗錄影的過程中並未選擇或輸入詞彙，而在檢索後問卷中的最終查詢詞彙填入，亦無法判斷該詞彙之層面。再者，使用者在利用 MAP 介面時以鍵盤輸入而非選擇建議詞彙的方式輸入之關鍵字，此種詞彙亦無法判斷其 MeSH 類別。

但 MAP 介面所提供建議詞彙類別並未完全包含 MeSH 既有之類別，其主要原因在於 1.該類別可透過 PubMed 之進階搜尋完成，如 Publication Characteristics、Geographicals、Named Groups 等類別。2.該類別與使用者主題較不相關，如 Information Science、Humanities、Technology Industry Agriculture 等類別，與本研究對象之生醫背景較無相關，在設計介面時就沒有考慮加入該層面。

表 4-10 詞彙分類表

類別	定義
Anatomy	解剖學
Organisms	有機組織
Diseases	疾病
Chemicals and Drugs	化學藥物

Analytical Diagnostic and Therapeutic Techniques and Equipment	診斷分析、治療技術及設備
Psychiatry and Psychology	精神病學和心理學
Phenomena and Processes	現象和過程
Disciplines and Occupations	學科與職業
Anthropology Education Sociology and Social Phenomena	教育社會學和社會現象
Health Care	健康保健
其他	無法歸類至 MeSH 層面之詞彙

一、不同主題熟悉度下的使用者查詢擴展行為

(一)、 瀏覽詞彙階段

瀏覽詞彙是指使用者在操作介面時，透過與系統的互動所新增的詞彙，在 PubMed 介面中指稱使用者透過與檢索結果互動所輸入替換的關鍵字，在 MAP 介面中則指稱使用者所點選之建議詞彙。

不同主題熟悉度瀏覽詞彙之類別如表 4-11，其橫條圖如圖 4-8。在全新檢索中，使用者最常瀏覽 Chemicals and Drugs、Diseases、Disciplines and Occupations 三種類別，而在重複檢索中使用者瀏覽之詞彙亦復如是，呈現相同的比例，因此可以推論不同主題熟悉度對於使用者瀏覽詞彙類別並無影響。

表 4-11 不同主題熟悉度瀏覽詞彙之類別

類別	全新檢索		重複檢索	
	次數	比例	次數	比例
[A]Anatomy	14	10%	20	9%
[B]Organisms	11	8%	10	4%
[C]Diseases	16	12%	23	10%
[D]Chemicals and Drugs	24	17%	43	18%
[E]Analytical Diagnostic and	15	11%	15	6%

Therapeutic Techniques and Equipment				
[F]Psychiatry and Psychology	3	2%	5	2%
[G]Phenomena and Processes	4	3%	5	2%
[H]Disciplines and Occupations	16	12%	33	14%
[I]Anthropology Education Sociology and Social Phenomena	0	0%	1	0%
[N]Health Care	7	5%	17	7%
其他	28	20%	62	26%
總計	138		234	

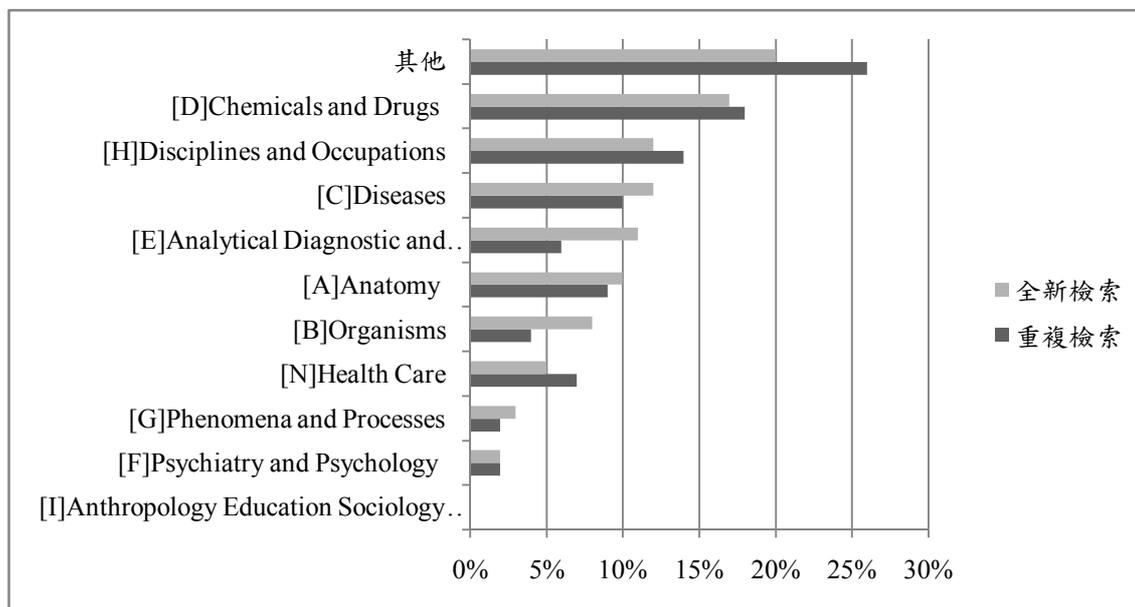


圖 4-8 不同主題熟悉度瀏覽之詞彙

(二)、 選用詞彙階段

選用詞彙階段則是指使用者在經過與檢索系統互動後，認定其輸入之關鍵字符合其資訊需求並能檢得其所需之資訊，於實驗後問卷中所填答之最終查詢詞彙。刪去與初始查詢重複之詞彙後，即為使用者之選用詞彙。

使用者選用之詞彙類別如表 4-12 及圖 4-9，結果顯示，在全新檢索中使用者最常選用 Chemicals and Drugs、Analytical Diagnostic and Therapeutic Techniques

and Equipment、Disciplines and Occupations 及 Diseases 4 類，而在重複檢索中除 Analytical Diagnostic and Therapeutic Techniques and Equipment 類，其較常選用之詞彙類別比例亦與重複檢索相同，因此可說明主題熟悉度對於選用詞彙類別並無明顯影響。

表 4-12 不同主題熟悉度選用之詞彙類別

類別	全新檢索		重複檢索	
	次數	比例	次數	比例
[A]Anatomy	5	6%	8	5%
[B]Organisms	6	7%	3	2%
[C]Diseases	7	9%	13	8%
[D]Chemicals and Drugs	11	14%	23	15%
[E]Analytical Diagnostic and Therapeutic Techniques and Equipment	8	10%	8	5%
[F]Psychiatry and Psychology	1	1%	3	2%
[G]Phenomena and Processes	2	2%	6	4%
[H]Disciplines and Occupations	7	9%	16	10%
[I]Anthropology Education Sociology and Social Phenomena	0	0%	1	1%
[N]Health Care	4	5%	8	5%
其他	30	37%	69	44%
總計	81		158	

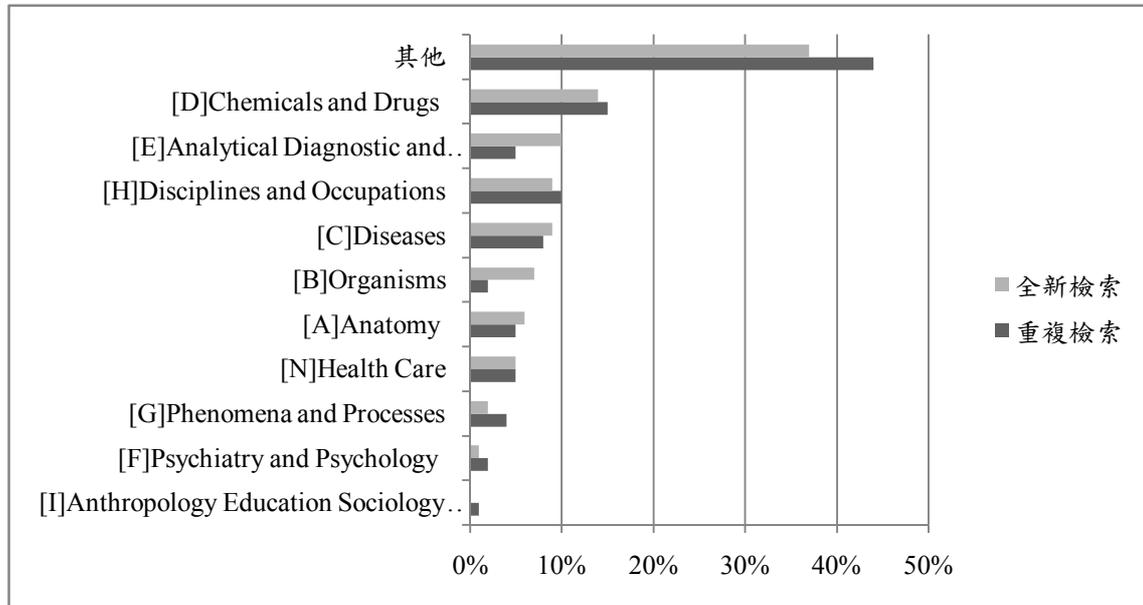


圖 4-9 不同主題熟悉度選用之詞彙類別

(三)、 不同主題熟悉度之瀏覽及選用詞彙

在全新檢索中使用者瀏覽及選用之詞彙類別如表 4-13 及圖 4-10。結果顯示，在瀏覽階段使用者最常選用 Chemicals and Drugs、Disciplines and Occupations 及 Diseases 3 類，而在選用階段除 Analytical Diagnostic and Therapeutic Techniques and Equipment 類亦占多數外，選用之詞彙類別亦與瀏覽詞彙之類別比例相同，而就數量而言則是呈微幅減少，因此可推論在不熟悉的情境下，使用者選用詞彙類別會與瀏覽詞彙類別比例相同，而透過多次檢索來驗證瀏覽詞彙之適用性，而略作刪減。

表 4-13 全新檢索之瀏覽及選用詞彙

類別	瀏覽階段		選用階段	
	次數	比例	次數	比例
[A]Anatomy	14	10%	5	6%
[B]Organisms	11	8%	6	7%

[C]Diseases	16	12%	7	9%
[D]Chemicals and Drugs	24	17%	11	14%
[E]Analytical Diagnostic and Therapeutic Techniques and Equipment	15	11%	8	10%
[F]Psychiatry and Psychology	3	2%	1	1%
[G]Phenomena and Processes	4	3%	2	2%
[H]Disciplines and Occupations	16	12%	7	9%
[I]Anthropology Education Sociology and Social Phenomena	0	0%	0	0%
[N]Health Care	7	5%	4	5%
其他	28	20%	30	37%
總計	138		81	

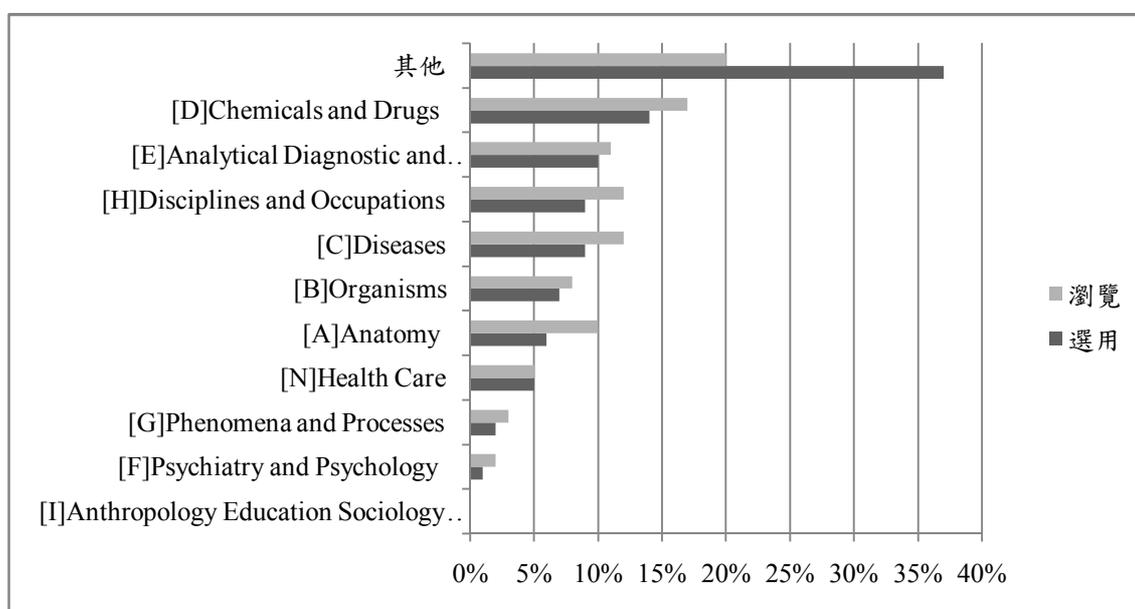


圖 4-10 全新檢索之瀏覽與選用詞彙

在重複檢索中使用者瀏覽及選用之詞彙類別如表 4-14 及圖 4-11。結果顯示，在瀏覽階段使用者最常選用 Chemicals and Drugs、Disciplines and Occupations 及 Diseases 3 類，而在選用階段使用者選用之詞彙類別亦與瀏覽詞彙之類別比例相同，而就數量而言則是呈減少的狀態，因此可推論在熟悉情境之下使用者選用詞彙與瀏覽詞彙之類別比例呈現相同的趨勢，並透過多次檢索來驗證瀏覽詞彙之適用性，而略作刪減。

表 4-14 重複檢索之瀏覽與選用詞彙

類別	瀏覽階段		選用階段	
	次數	比例	次數	比例
[A]Anatomy	20	9%	8	5%
[B]Organisms	10	4%	3	2%
[C]Diseases	23	10%	13	8%
[D]Chemicals and Drugs	43	18%	23	15%
[E]Analytical Diagnostic and Therapeutic Techniques and Equipment	15	6%	8	5%
[F]Psychiatry and Psychology	5	2%	3	2%
[G]Phenomena and Processes	5	2%	6	4%
[H]Disciplines and Occupations	33	14%	16	10%
[I]Anthropology Education Sociology and Social Phenomena	1	0%	1	1%
[N]Health Care	17	7%	8	5%
其他	62	26%	69	44%
總計	234		158	

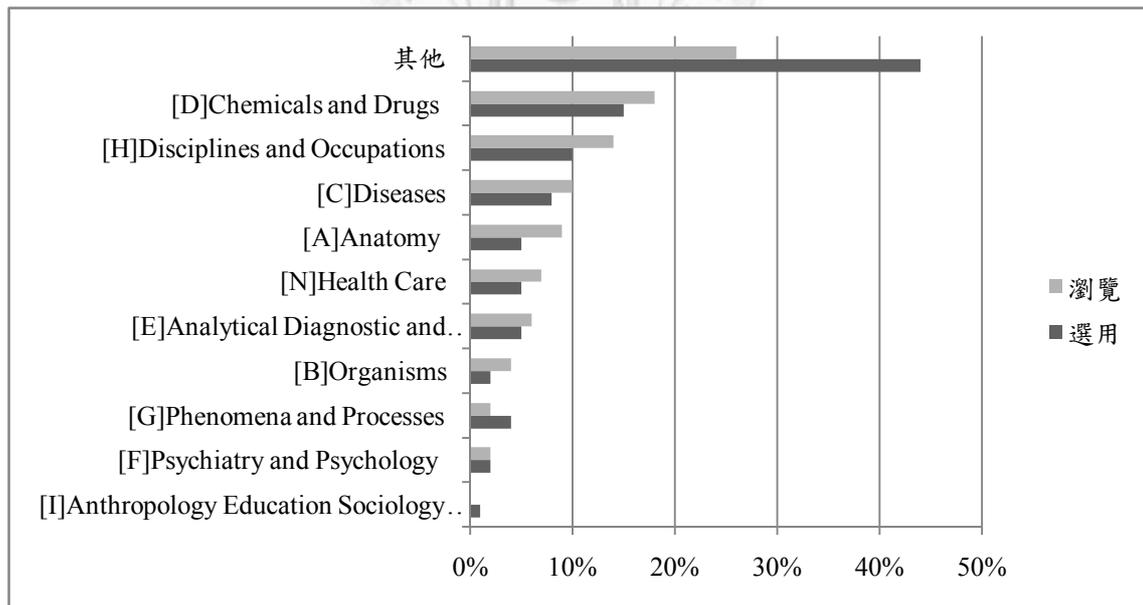


圖 4-11 重複檢索之瀏覽與選用詞彙

二、不同介面下的使用者查詢擴展行為

本研究利用 Morae®視窗錄影軟體完整記錄使用者檢索之過程，觀察發現使用者於 PubMed 介面中其詞彙利用行為可以分成三類，其一為利用檢索結果之特定子句，複製檢索結果中符合資訊需求之子句進行檢索。其二為利用檢索結果之文獻作者，針對特定作者之文獻進行檢索。其三是以檢索需求說明時所填寫的關鍵字進行排列組合交叉檢索。

MAP 介面中的使用者行為則是輸入初始查詢後，系統傳回相關的建議詞彙，而後使用者點選檢索之，並自行修改詞彙而非採用原始 MeSH 詞彙之結構。而在瀏覽檢索結果的過程中亦會刺激使用者輸入其認為相關的詞彙並進行檢索。

(一)、 瀏覽詞彙階段

使用者不同介面之瀏覽建議詞彙次數分配表如表 4-15 及圖 4-12。結果顯示在瀏覽階段，在 PubMed 介面中使用者最常輸入 Chemicals and Drugs、Diseases、Analytical Diagnostic and Therapeutic Techniques and Equipment 3 類之詞彙。在 MAP 中使用者則較常點選 Chemicals and Drugs、Disciplines and Occupations、Anatomy 3 類詞彙進行查詢。此一現象有別於使用者在不同熟悉度下所瀏覽之詞彙仍趨一致的結果，由此可見不同介面對於使用者之瀏覽詞彙有所影響。

表 4-15 瀏覽建議詞彙次數分配表

類別	PubMed		Map	
	次數	比例	次數	比例
[A]Anatomy	6	6%	28	10%
[B]Organisms	5	5%	16	6%
[C]Diseases	14	15%	25	9%
[D]Chemicals and Drugs	25	26%	42	15%
[E]Analytical Diagnostic and Therapeutic Techniques and Equipment	9	9%	21	8%

[F]Psychiatry and Psychology	0	0%	8	3%
[G]Phenomena and Processes	9	9%	N/A	N/A
[H]Disciplines and Occupations	2	2%	47	17%
[I]Anthropology Education Sociology and Social Phenomena	1	1%	N/A	N/A
[N]Health Care	1	1%	23	8%
其他	24	25%	66	24%
Total	96		276	

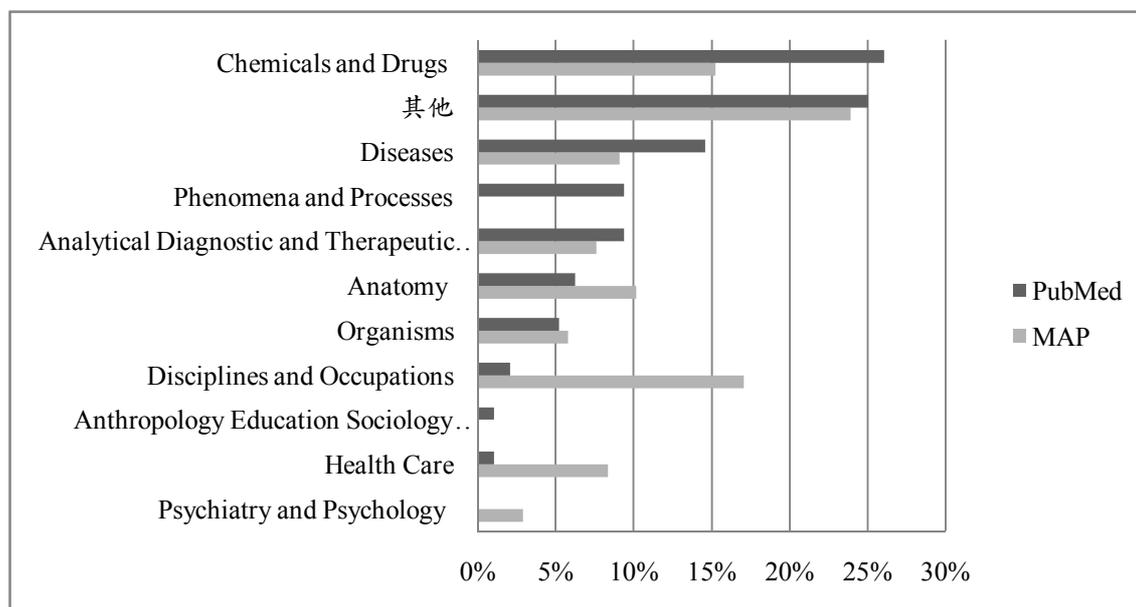


圖 4-12 使用者不同介面之瀏覽之建議詞彙次數長條圖

(二)、 選用詞彙階段

使用者在不同介面下所選用之詞彙類別如表 4-16 及圖 4-13。結果顯示在選用詞彙階段，於 PubMed 介面中使用者輸入之查詢詞彙類別最多者為 Chemicals and Drugs、Diseases、Analytical Diagnostic and Therapeutic Techniques and Equipment 3 類。在 MAP 中則是 Chemicals and Drugs、Disciplines and Occupations、Anatomy 及 Health Care 等類，不同介面對於使用者選用之詞彙確有影響。

表 4-16 不同介面選用詞彙之類別

類別	PubMed		Map	
	次數	比例	次數	比例
[A]Anatomy	4	4%	9	7%
[B]Organisms	4	4%	5	4%
[C]Diseases	14	13%	6	5%
[D]Chemicals and Drugs	21	19%	13	10%
[E]Analytical Diagnostic and Therapeutic Techniques and Equipment	13	12%	3	2%
[F]Psychiatry and Psychology	1	1%	3	2%
[G]Phenomena and Processes	8	7%	N/A	N/A
[H]Disciplines and Occupations	1	1%	22	17%
[I]Anthropology Education Sociology and Social Phenomena	1	1%	N/A	N/A
[N]Health Care	3	3%	9	7%
其他	39	35%	60	46%
總計	110		130	

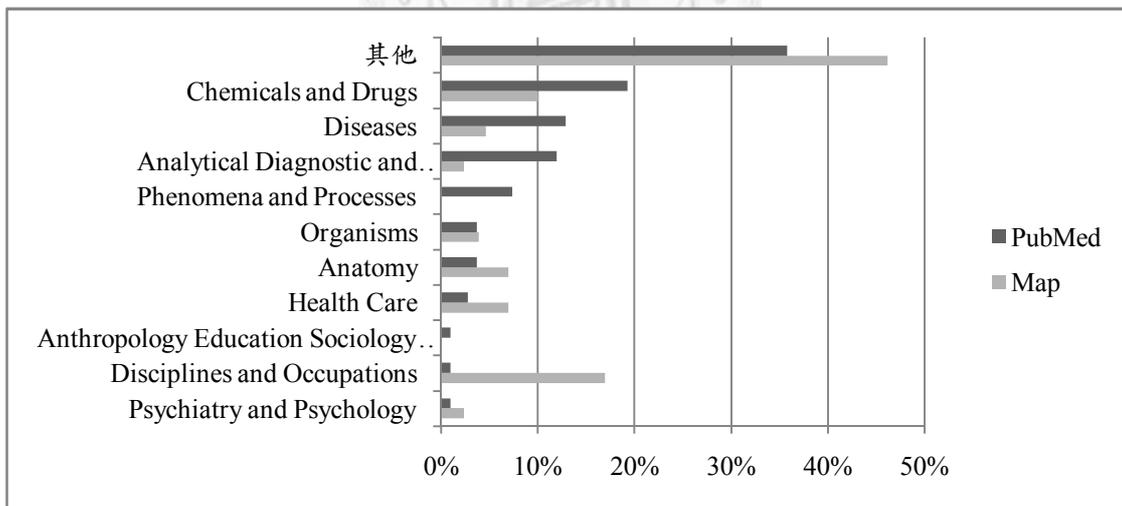


圖 4-13 使用者不同介面之最終選擇之建議詞彙次數長條圖

(三)、 不同介面瀏覽詞彙與選用詞彙

PubMed 使用者瀏覽及選用之詞彙類別如表 4-17 及圖 4-14。兩階段之詞彙類別之比例呈現一致的方向，而在數量上呈現微幅減少，其原因為，在 PubMed 介

面中，使用者受到之刺激多來自檢索結果之互動，而就檢索結果中的相關字再次檢索。其前提在於該關鍵字已然出現在使用者認為相關之文獻之中，而使用者以該相關文獻為基準，而以其中之關鍵字進行檢索。是故該關鍵字契合使用者資訊需求程度愈高，因此使用者所瀏覽輸入之詞彙亦成為使用者認為之最佳查詢詞彙。

表 4-17 PubMed 介面使用者瀏覽與最終查詢詞彙類別

類別	瀏覽階段		選用階段	
	次數	比例	次數	比例
[A]Anatomy	6	6%	4	4%
[B]Organisms	5	5%	4	4%
[C]Diseases	14	15%	14	13%
[D]Chemicals and Drugs	25	26%	21	19%
[E]Analytical Diagnostic and Therapeutic Techniques and Equipment	9	9%	13	12%
[F]Psychiatry and Psychology	0	0%	1	1%
[G]Phenomena and Processes	9	9%	8	7%
[H]Disciplines and Occupations	2	2%	1	1%
[I]Anthropology Education Sociology and Social Phenomena	1	1%	1	1%
[N]Health Care	1	1%	3	3%
其他	24	25%	39	36%
總計	96		109	

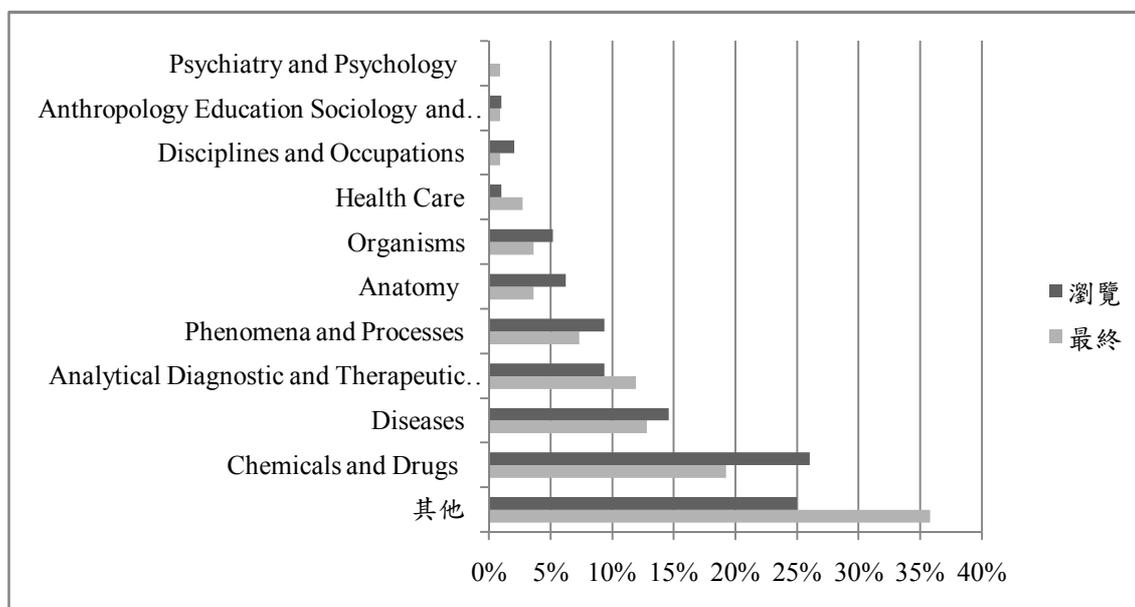


圖 4-14 PubMed 介面瀏覽及選用之建議詞彙橫條圖

MAP 介面瀏覽及選用詞彙類別如表 4-18 及圖 4-15。在 MAP 介面之中，使用者在瀏覽階段及選用階段之詞彙類別比例趨於一致，而選用查詢詞彙之數量明顯低於瀏覽詞彙，其原因在於 MAP 介面在瀏覽階段時提供使用者大量的相關建議詞彙，而使用者亦會點選大量的建議詞彙進行檢索。但建議詞彙與檢索主題之相關性是由系統而非使用者判斷。而選用詞彙則是使用者就多次檢索之結果判斷建議詞彙與其資訊需求相關與否後選用。因此會形成使用者選用大量的瀏覽詞彙以及經過篩選後少量的選用查詢詞彙。

表 4-18 MAP 介面使用者瀏覽與最終查詢詞彙類別

類別	瀏覽		選用	
	次數	比例	次數	比例
Anatomy	28	10%	9	7%
Organisms	16	6%	5	4%
Diseases	25	9%	6	5%
Chemicals and Drugs	42	15%	13	10%
Analytical Diagnostic and Therapeutic Techniques and Equipment	21	8%	3	2%

Psychiatry and Psychology	8	3%	3	2%
Disciplines and Occupations	47	17%	22	17%
Health Care	23	8%	9	7%
其他	66	24%	60	46%
總計	276		130	

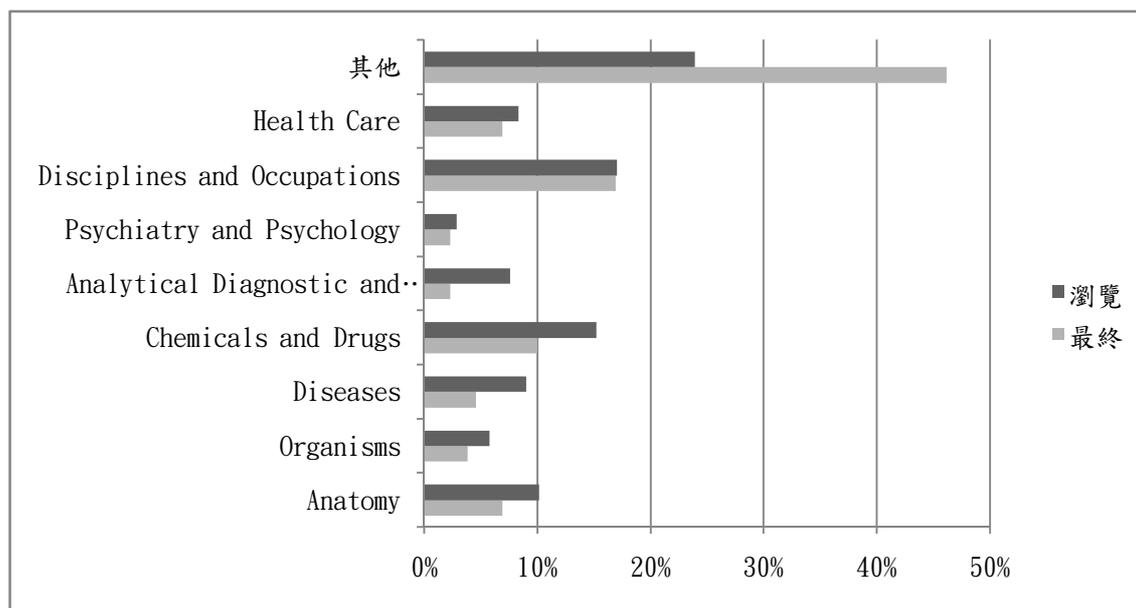


圖 4-15 MAP 介面瀏覽及最終選擇之建議詞彙橫條圖

三、重複檢索與否與不同介面對於查詢擴展之交互作用

(一)、 瀏覽詞彙階段

使用者不同介面瀏覽之詞彙類別如表 4-19、圖 4-16 及圖 4-17。結果顯示在瀏覽詞彙階段，PubMed 介面中不同主題熟悉度選擇之詞彙類別比例未見顯著差異，多集中於 Diseases 及 Chemicals and Drugs 2 類之中，說明不論主題熟悉與否，在 PubMed 介面中，其詞彙瀏覽行為並未見明顯差異。而在 MAP 介面中，不同主題熟悉度使用者之瀏覽詞彙均集中於 Chemicals and Drugs 及 Disciplines and Occupations 2 類，主題熟悉度對 MAP 介面中的瀏覽詞彙亦未有明顯影響。

表 4-19 不同介面及主題熟悉度之瀏覽詞彙

類別	PubMed		PubMed		MAP		MAP	
	全新檢索		重複檢索		全新檢索		重複檢索	
	次數	比例	次數	比例	次數	比例	次數	比例
[A]Anatomy	2	6%	4	6%	12	11%	16	9%
[B]Organisms	4	12%	1	2%	7	7%	9	5%
[C]Diseases	5	15%	9	14%	11	10%	14	8%
[D]Chemicals and Drugs	7	21%	18	29%	17	16%	25	15%
[E]Analytical Diagnostic and Therapeutic Techniques and Equipment	5	15%	4	6%	10	10%	11	6%
[F]Psychiatry and Psychology	0	0%	0	0%	3	3%	5	3%
[G]Phenomena and Processes	4	12%	5	8%	N/A	N/A	N/A	N/A
[H]Disciplines and Occupations	0	0%	2	3%	16	15%	31	18%
[I]Anthropology Education Sociology and Social Phenomena	0	0%	1	2%	N/A	N/A	N/A	N/A
[N]Health Care	0	0%	1	2%	7	7%	16	9%
其他	6	18%	18	29%	22	21%	44	26%
總計	33		63		105		171	

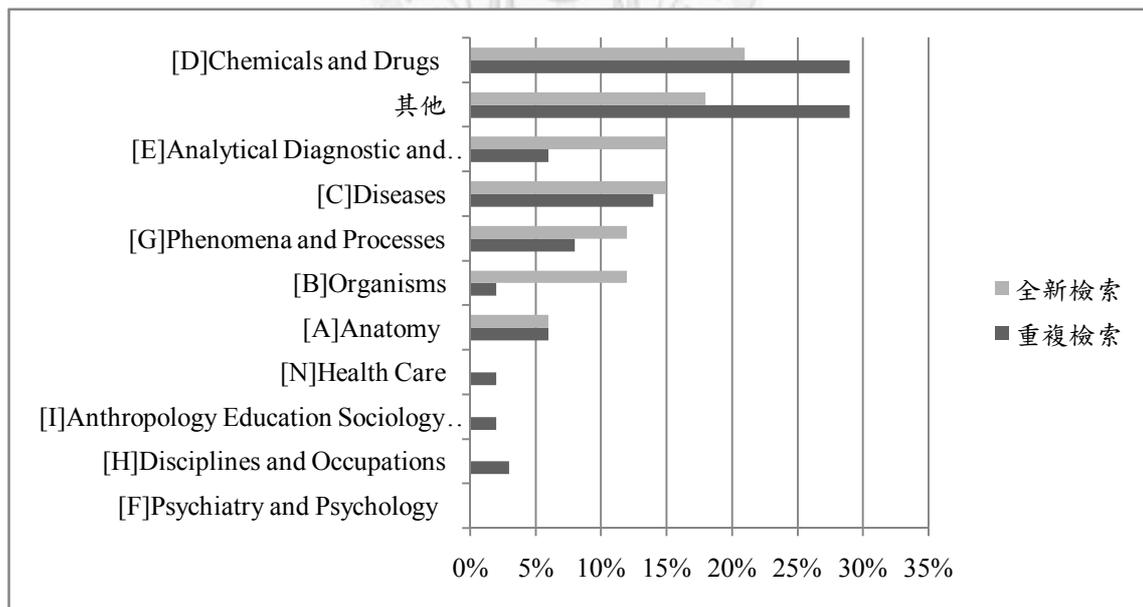


圖 4-16 Pubmed 介面及不同熟悉度之瀏覽詞彙

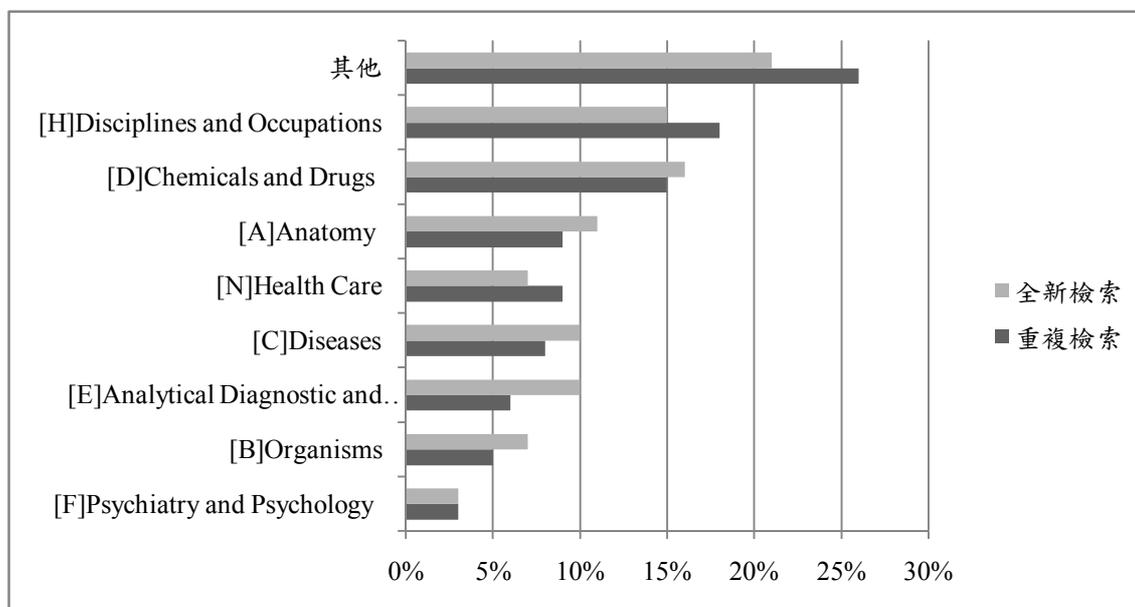


圖 4- 17 MAP 介面及不同主題熟悉度之瀏覽詞彙

(二)、 選用詞彙階段

使用者不同介面選用之詞彙類別如表 4-20、圖 4-18 及圖 4-19。結果顯示在選用詞彙階段，PubMed 介面中不同主題熟悉度選擇之詞彙類別比例未見顯著差異，多集中於 Diseases 及 Chemicals and Drugs 2 類之中，說明不論主題熟悉與否，在 PubMed 介面中，其詞彙瀏覽行為並未見明顯差異。而在 MAP 介面中，不同主題熟悉度使用者之選用詞彙均集中於 Chemicals and Drugs 及 Disciplines and Occupations 2 類，主題熟悉度對 MAP 介面中的瀏覽詞彙亦未有明顯影響。

表 4- 20 不同介面及主題熟悉度之選用詞彙

類別	PubMed 全新檢索		PubMed 重複檢索		MAP 全新檢索		MAP 重複檢索	
	次數	比例	次數	比例	次數	比例	次數	比例
[A]Anatomy	1	3%	3	4%	4	10%	5	6%
[B]Organisms	3	8%	1	1%	3	7%	2	2%
[C]Diseases	6	15%	8	11%	1	2%	5	6%
[D]Chemicals and Drugs	7	18%	14	20%	4	10%	9	10%
[E]Analytical Diagnostic and	7	18%	6	9%	1	2%	2	2%

Therapeutic Techniques and Equipment								
[F]Psychiatry and Psychology	0	0%	1	1%	1	2%	2	2%
[G]Phenomena and Processes	2	5%	6	9%	N/A	N/A	N/A	N/A
[H]Disciplines and Occupations	0	0%	1	1%	7	17%	15	17%
[I]Anthropology Education Sociology and Social Phenomena	0	0%	1	1%	N/A	N/A	N/A	N/A
[N]Health Care	1	3%	2	3%	3	7%	6	7%
其他	12	31%	27	39%	18	43%	42	48%
總計	39		70		42		88	

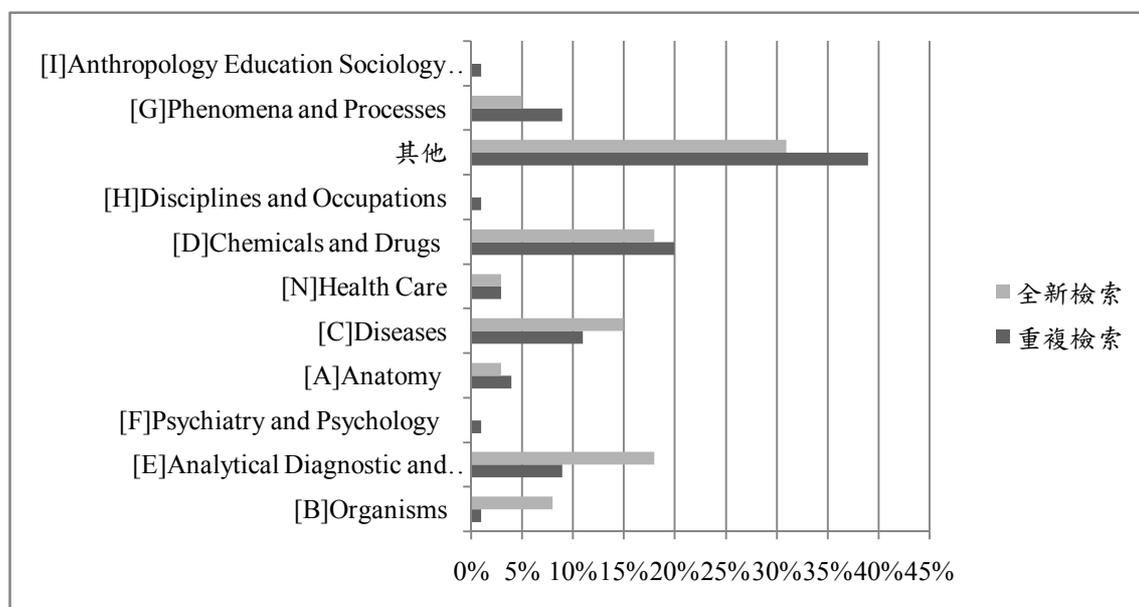


圖 4- 18 PubMed 介面及不同主題熟悉度之選用詞彙

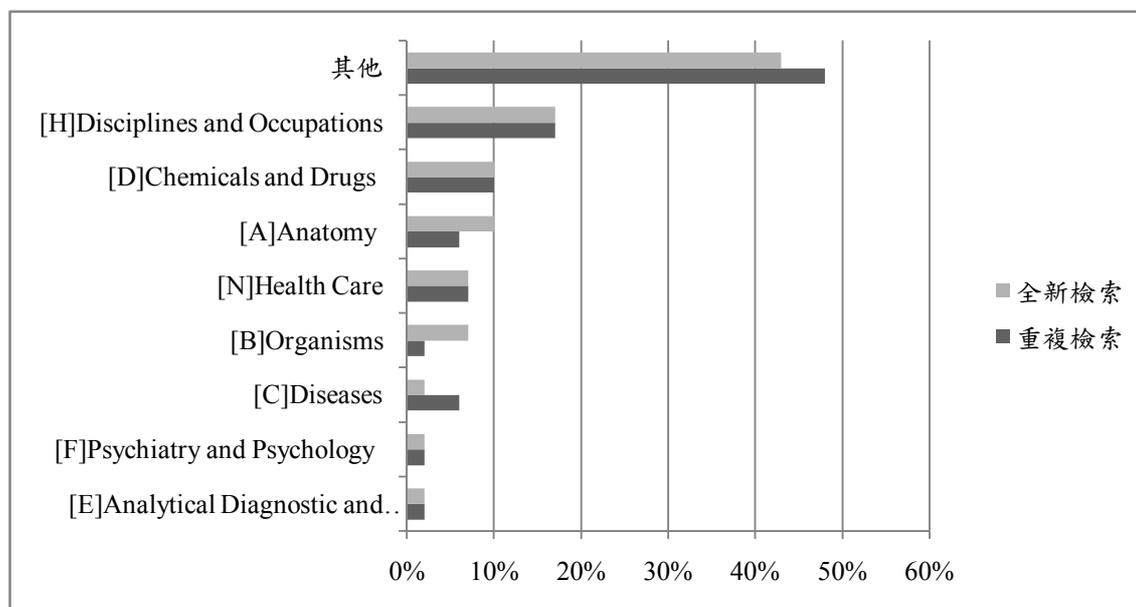


圖 4-19 MAP 介面及不同主題熟悉度之選用詞彙

四、小結

不同主題熟悉度高度利用查詢擴展詞彙類別如表 4-21。研究結果顯示在不同主題熟悉度下，使用者之瀏覽詞彙及選用詞彙在數量及類別上均無差異，說明無論在熟悉與否的情境下，使用者運用詞彙的過程並無二致。

表 4-21 不同主題熟悉度高度利用查詢擴展詞彙類別

不同主題熟悉度-瀏覽	常用詞彙類別
全新檢索	Chemicals and Drugs、Diseases、Disciplines and Occupations
重複檢索	Chemicals and Drugs、Diseases、Disciplines and Occupations
不同主題熟悉度-選用	
全新檢索	Chemicals and Drugs、Analytical Diagnostic and Therapeutic Techniques and Equipment、Disciplines and Occupations、Diseases
重複檢索	Chemicals and Drugs、Disciplines and Occupations、Diseases

不同介面高度利用查詢擴展詞彙類別如表 4-22。在不同介面下，使用者之瀏覽及選用詞彙在數量以及類別均有所不同。不同介面之瀏覽與選用詞彙類別之所以會產生差異的原因在於，在 PubMed 介面中，使用者的詞彙來自於記憶中對於

主題的認知。而在 MAP 介面中，系統能夠以 MeSH 結構呈現主題內的完整的知識結構，使用者的檢索詞彙除了出自於己身對於主題的理解外，同時可透過點選與主題相關的 MeSH 詞彙進行檢索，其查詢之關鍵字能夠涵括檢索主題的各個面向，因此在不同介面下使用者瀏覽以及選用詞彙類別會有所不同。就瀏覽及選用詞彙數量而言，在 PubMed 下使用者詞彙來源有限，因此數量並無明顯變化。而在 MAP 介面中，系統會大量提供與主題相關之建議詞彙讓使用者點選，故其瀏覽詞彙數會較多。而其選用之詞彙則是經由檢索結果進一步驗證其與主題之相關性進行篩選，因此其數量會較瀏覽詞彙少。

表 4-22 不同介面高度利用查詢擴展詞彙類別

不同介面-瀏覽	常用詞彙類別
PubMed	Chemicals and Drugs、Diseases
Map	Analytical Diagnostic and Therapeutic Techniques and Equipment Chemicals and Drugs、Disciplines and Occupations、Anatomy
不同介面-選用	
PubMed	Chemicals and Drugs、Diseases
MAP	Analytical Diagnostic and Therapeutic Techniques and Equipment Chemicals and Drugs、Disciplines and Occupations、Anatomy、 Health Care

不同介面及不同主題熟悉度高度利用查詢擴展詞彙類別如表 4-23。在不同介面下，不同主題熟悉度之使用者瀏覽及選用詞彙在數量及類別上並無差異。此一現象說明不同主題熟悉度與不同介面對於使用者之查詢擴展行為並無明顯交互作用。

表 4-23 不同介面及不同主題熟悉度高度利用查詢擴展詞彙類別

交互作用-瀏覽	常用詞彙類別
PubMed-全新檢索	Diseases、Chemicals and Drugs
PubMed-重複檢索	Diseases、Chemicals and Drugs
MAP-全新檢索	Chemicals and Drugs、Disciplines and Occupations
MAP-重複檢索	Chemicals and Drugs、Disciplines and Occupations
交互作用-選用	
PubMed-全新檢索	Diseases、Chemicals and Drugs

PubMed-重複檢索	Diseases、Chemicals and Drugs
MAP-全新檢索	Chemicals and Drugs、Disciplines and Occupations
MAP-重複檢索	Chemicals and Drugs、Disciplines and Occupations

本研究嘗試將使用者在不同主題熟悉度以及不同介面中的瀏覽及選用詞彙以 MeSH 之層面進行歸類，雖然 MeSH 唯一專業的醫學索引典，但仍無法包含使用者所有可能查詢之關鍵字，也因此會造成詞彙無法分類的情形。所以會形成在不同的檢索階段中，其他類的詞彙比例會較任一 MeSH 層面之詞彙來的高。儘管研究者無法分類使用者所有利用的檢索詞彙，但透過分析各個檢索階段使用者的利用詞彙，仍有助於理解使用者的查詢擴展行為。

第四節 檢索結果過濾分析

使用者的初始查詢多使用過於廣泛的關鍵字進行檢索，而後透過與檢索結果互動後更進一步的重構(reformulation)其查詢，本研究的研究目的之一即是比較使用者透過不同介面重構其查詢而過濾初始查詢中不相關的資訊，同時挖掘位於檢索結果底部的相關資訊的能力。

一般而言，使用者缺乏領域內的相關知識或僅有模糊的資訊需求時，從而以少量關鍵字進行檢索，其檢索結果則是以詞彙共現之程度進行排序的紀錄，在 PubMed 之中檢索結果則是以年代排序，造成在大量的檢索結果之中，排序較低的相關文獻則被忽略。而透過 MAP 提供的大量相關詞彙進行查詢重構，能夠進一步強化其資訊需求表述，其目的在於使能檢得在大量檢索結果之中排序較低但相關程度亦高的文獻。

若欲比較 PubMed 與 MAP 介面何者較能輔助使用者取得於在大量檢索結果中排序較低的相關文獻，其前提為在實驗過程中，使用者並未刪除其初始查詢中的任一關鍵字，以確保最終查詢是透過新增詞彙過濾初始查詢結果中的不相關文獻。由於實驗設計之故，並非所有的檢索均能符合此一條件。在 88 組檢索之中，

共有 10 組檢索符合此一條件(11.36%)。相關文獻之中數、最大值、最小值、全距如表 4-24。

使用者在進行檢索時，均須選取 10 筆相關文獻，此即為比較不同介面輔助取得排序較低的相關文獻之基準，案例之比較是以文件在初始查詢結果中的位置進行。

在 10 組符合檢索初始查詢結果的案例中可以發現，PubMed 及 MAP 介面其相關文獻位置之最小值未有明顯的差距，但在最大值的部分則 MAP 介面多遠高於 PubMed 介面，顯示 MAP 介面確能較 PubMed 介面輔助使用者取得在原始查詢中排序較低的相關文獻。

表 4-24 不同介面之相關文獻於初始檢索結果之位置

使用者	介面	中位數	平均數	標準差	最小值	最大值	全距	個數
2	PUBMED	134.5	131.17	54.25	39	197	158	6
	MAP	503	516.17	259.39	219	829	610	6
24	PUBMED	52	54.80	38.17	11	139	128	10
	MAP	78.5	99.50	50.36	38	167	129	10
25	PUBMED	440	487.83	182.87	345	820	475	6
	MAP	1393.5	5078.83	6513.36	366	13812	13446	6
27	PUBMED	38.5	44.90	34.60	2	103	101	10
	MAP	42	54.10	44.73	7	128	121	10
29	PUBMED	133	115.10	42.80	52	171	119	10
	MAP	90	114.60	58.50	58	254	196	10
30	PUBMED	60	69.50	30.05	39	121	82	10
	MAP	476	483.00	207.92	194	867	673	10
31	PUBMED	695.5	687.63	38.74	629	745	116	8
	MAP	1371.5	1459.13	793.21	616	2549	1933	8
32	PUBMED	107	103.29	38.90	33	160	127	7
	MAP	151	162.71	97.45	33	297	264	7
44	PUBMED	88	84.90	32.79	37	125	88	10
	MAP	71	141.00	198.15	26	676	650	10
46	PUBMED	583.5	873.70	803.87	90	2420	2330	10

以使用者 2 號為例，其不同介面之文獻排序如圖 4-20，在 PubMed 介面中使用者選取之相關文獻集中於檢索結果中的前 197 筆，而在 MAP 介面下，在相同的檢索結果之中，使用者反而能夠取得檢索結果中排序末端的第 829 筆記錄。

而使用者 30 號之不同介面之文獻排序如圖 4-21，在 PubMed 介面中使用者選取之相關文獻則集中於檢索結果中的前 121 筆，而在 MAP 介面下使用者反而能夠取得第 867 筆記錄。

這些案例反映了 MAP 介面能夠較 PubMed 介面輔助使用者能進一步取得於大量檢索結果底層之相關文獻。

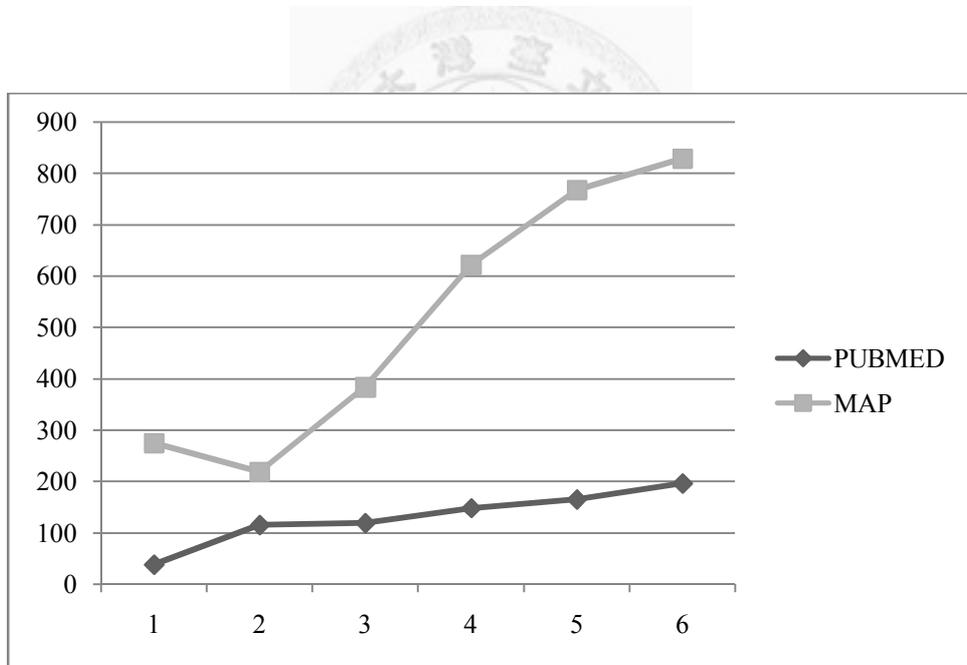


圖 4- 20 使用者 2 號之不同介面文獻排序

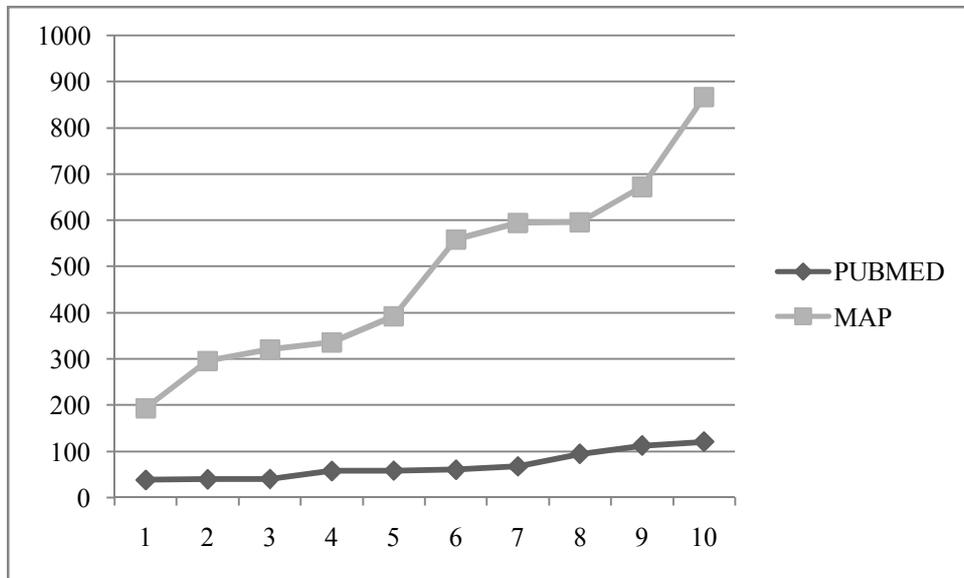


圖 4-21 使用者 30 號之不同介面文獻排序

第五節 實驗之次序效應

本研究關注於使用者在利用不同檢索介面時，其查詢關鍵字改變的過程。由於使用者之內在知識會受到檢索過程的影響而有所改變，因此介面相互操作的過程極易影響使用者對於關鍵字的認知。而實驗流程為使用者以 MAP 介面進行檢索後，以相同關鍵字以 PubMed 介面再次進行檢索或反之操作。在實驗流程中在兩個介面交互操作時已設定使用者中斷時間，同時要求使用者在兩個介面時均使用相同之詞彙作為初始查詢，目的均是為了降低實驗之次序效應。本研究判定使用者產生次序效應之基準在於，若受試者在進行第二個介面檢索時，運用了前次介面檢索所獲得的詞彙或是如前次檢索刪除檢索詞彙時，即視為次序效應之產生。

研究結果顯示在 88 次的檢索中總共出現了 40 次的次序效應(45.45%)，如表 4-25。其中 PubMed 先 MAP 後之實驗出現了 16 次，MAP 先 PubMed 後之實驗

則出現了 24 次。

表 4-25 實驗產生之次序效應

實驗順序	主題熟悉程度	N
PubMed → MAP	重複檢索	9
	全新檢索	7
MAP → PubMed	重複檢索	19
	全新檢索	5

次序效應的產生，說明了實驗中使用者的確會受到前一介面操作過程的影響，進而將前次操作經驗代入第二種介面之檢索。先進行 MAP 介面而後 PubMed 介面之實驗設計產生較多次的次序效應，原因在於使用者利用 MAP 介面檢索時，系統會主動提供基於 MeSH 的相關領域內詞彙，其數量之多，遠大於使用者利用 PubMed 介面時與檢索結果互動後所能激發的概念。

表 4-26 為先 PubMed 後 MAP 之次序效應案例，舉例而言，2 號使用者在 PubMed 介面中最後選擇 microbiology 作為最佳的檢索詞彙之一，而後在 MAP 介面中同樣也選擇 microbiology 作為最佳的檢索詞彙之一。22 號使用者則是在 PubMed 介面中的最終查詢詞彙刪去 proteomic 及 HUVEC 二字，而後在 MAP 介面的最終查詢詞彙中亦刪去同一詞彙。

表 4-27 則為先 MAP 後 PubMed 之次序效應案例，其中 7 號使用者在 MAP 介面中輸入 subtype 作為最終查詢詞彙，並在 PubMed 做出相同的行為。而 12 號使用者則是在 MAP 介面中在最終查詢詞彙刪去 social behavior，而在 PubMed 介面亦復如是。以上案例足證介面利用之先後對於使用者詞彙之選擇有所影響。

表 4-26 先 PubMed 介面後 MAP 介面之次序效應案例

ID	REVISIT	介面	初始查詢	最終查詢	ADDED	Removed	Overlap
2	Y	PubMed	respiratory, airborne	respiratory AND airborne microbiology	microbiology		microbiology
		MAP		Respiratory airborne microbiology, Cough, Sneezing	microbiology, cough, sneezing		
22	Y	PubMed	proteomic, Sal B, HUVEC	PA1-l; Sal B; endothelial cell	endothelial cell	proteomic; HUVEC	proteomic; HUVEC
		MAP		Sal B; PA1-l; Card; vascular system; hypertension	Card; vascular system; hypertension	proteomic; HUVEC	

表 4-27 先 MAP 介面後 PubMed 介面之次序效應案例

ID	REVISIT	介面	初始查詢	最終查詢	ADDED	Removed	Overlap
7	N	MAP	hemagglutinin, binding affinity; influenza	hemagglutinin, binding affinity, influenza, subtype, H1N1;H3N2; H5N1	subtypes, H1N1;H3N2; H5N1		subtypes
		PubMed		hamagglutinin binding affinity; ingluenza,subtypes	subtypes		
12	Y	MAP	TBI & social behavior & ToM & empathy	TBI & ToM & empathy		social behavior	social behavior
		PubMed		TBI & ToM & empathy		social behavior	

第六節 使用者評論

本研究在實驗過後要求使用者填寫問卷，要求使用者說明 MAP 介面在何時較有幫助，並給予相關的改進建議。經過分析後，使用者認為 MAP 介面的主要的功能在於資料過濾、關鍵詞聯想等兩個面向。舉例而言，使用者 6、9、18、

23、26、31、34 均提出了資料過濾的概念：

如果研究主題屬於大範圍，詞彙建議會有幫助。(6)

當搜尋範圍過大以致資料難以消化(數百、數千筆資料)則可以此減少資料彙整。(9)

對牽涉到較多領域的主題能夠幫助集中焦點。對於比較細部的研究能提供精確的資料，幫助節省過濾資料的時間。(18)

縮小 search 的 range。(23)

在分類下，有系統的更多相關 TERM 更容易符合較細的檢索目標。(26)

在做研究時，往往需要擬定方向，才能夠縮小範圍，在詞彙建議下，可協助使用者找到相對應的資源。(34)

而在幫助聯想的部分，使用者 1、3、11、23、27 則認為 MAP 介面有助於聯想相關的檢索詞彙：

提供更精確的字眼或同義字幫助聯想。(1)

適時的提供 KEYWORD、新的想法。(3)

可提供研究者當初無想到之關鍵詞彙。(11)

找到更有興趣的主題。(23)

提供新的關鍵字或新的訊息，也許原本不知道(不熟)的字。也許原本知道，但仍可快速的分類資料或提醒提供更精確的字眼或同義字幫助聯想。(27)

使用者認為在熟悉主題與不熟悉主題的情況下，MAP 的功用互有不同。在不熟悉主題的情境下，MAP 介面可以提供 MeSH 中的語意架構，輔助使用者了解領域內的知識架構並取得相關的關鍵詞進行檢索，37 名使用者均提出了相似的概念，以使用者 2、4、16、20、30、33 為例：

題目不熟時；不知下關鍵字時；有助 Help 與知道下何種 KEYWORDS。(2)

在對該主題不熟悉時，可觸類旁通，知道某些 KEYWORD 後，MESH 可提供我其他可能的 WORD 是我之前所不知道的，剛進入此領域時有幫助。(4)

我認為在一開始完全進到新的、完全不懂的領域下，MAP 的 MESH 功能有很多分類可以了解這疾病如何在人類產生，及其機制，這可以幫助剛進來的生手。(16)

當對查找主題不熟悉時，MESH 提供了許多對該主題的建議，幫助使用者更快取得對主題的概觀。(20)

當使用者想要收集某一較 GENERAL 的主題時，MAP 是很好的幫手，使用者可以得到很多新的靈感也不受關鍵字限制，可得到更多資料。(30)

在僅知大範圍的關鍵字時，詞彙建議可以幫助提供原先沒有想到的 term 或 narrow down 搜尋範圍，更容易找到目標文獻。(33)

而在熟悉主題的情境下，使用者認為 MAP 介面的輔助效益不大，但仍有助於相關概念的聯想及指引，以使用者 14、20、21、22、30、36、37、43 為例：

對於熟悉的主題，差異則不大，但會使搜尋更精確。(14)

而在針對熟悉領域時，MESH 也扮演了一個提醒者的角色，可確保任一關鍵字不被遺忘。(20)

在平常已有確定目標的搜尋中，pubmed/map 的效果差不多。(21)

較熟悉的研究主題，能提供更多不同層面的知識。(22)

當使用者已有一非常 SPECIFIC 的題目時，MAP 幫助不大，甚至會使資料搜尋速度下降。(30)

再者，如果對所蒐集詞彙了解十分透徹，可藉詞彙建議了解其他研究面向。(35)

在已熟悉的前提下，MAP 仍有助於縮小尋找範圍，並有助於條列式搜尋。(36)

當對主題很瞭解時，詞彙建議能限制出想要的主題並開啓新的思維。(37)

但是在對於要檢索的文獻已相當熟悉的情況下，有時甚至毋需使用到 MESH 即可關成檢索的工作，但 MESH 仍能提供新的方向，發現自己熟悉題目的新走向，有助於開展視野。(41)

在熟悉的題目上，感覺詞彙建議的功能就比較貧弱，因此其提供的關鍵字，對於熟悉的題目而言，實在不夠具體。並且知識上較清楚，相關的重要程度。因此在策略上寧可運用既有知識去搜尋。(44)

在實驗過後，要求使用者以五點尺度填答對於建議詞彙之滿意度以及以七點尺度填答，題目為若 MAP 介面有所幫助則是在何種面向較有幫助，其面向分別為 1.其輔助產生新的概念進行下一步蒐尋 2.幫助釐清檢索問題 3.能夠顯示資料庫文獻的架構 4.管理大量的檢索結果。

研究結果如表 4-28，雖然新舊檢索對於建議詞彙之滿意度雖無顯著差異，但新的檢索對於建議詞彙的滿意度(m=4.28)略高於重複檢索(m=4.08)。而在 MAP 的輔助面向部分，則只有在「產生新的概念」此一面向新舊檢索有明顯差異

($p=.036$)。分析結果與使用者意見相符，顯示 MAP 介面在使用者面臨不熟悉主題時能夠提供相關主題詞彙供使用者利用檢索之。

表 4- 28 使用者對於 MAP 輔助功能之評估

		平均數	標準差	個數
詞彙滿意程度	全新檢索	4.39	1.10	28
	重複檢索	4.03	1.68	60
產生新的概念*	全新檢索	4.57	1.10	28
	重複檢索	3.95	1.35	60
釐清檢索問題	全新檢索	4.18	1.09	28
	重複檢索	4.03	1.57	60
顯示文獻結構	全新檢索	4.18	1.28	28
	重複檢索	4.50	1.10	60
管理檢索結果	全新檢索	4.29	0.90	28
	重複檢索	4.42	1.16	59

*: $p<.05$

第五章 結論與建議

為調查使用者查詢擴展行為在不同介面上之差異及詞彙建議系統輔助之功效，本研究利用實驗法及問卷法了解使用者在不同介面的詞彙選擇行為以及實驗過後的差異。茲將本研究所得之結論歸納如下，同時依據研究發現提出建議及未來研究方向，以供未來從事相關研究者之參考。

第一節 結論

本研究旨在了解不同介面下的使用者查詢擴展行為，以及 MAP 介面輔助使用之成效，期望能夠提供資訊檢索系統設計者及使用者行為研究者參考。以下綜合本研究分析使用者詞彙選擇之結果以及相關特殊案例之說明，針對使用者查詢擴展行為提出以下結論。

一、 主題熟悉度會影響使用者查詢行為

研究結果顯示，使用者在熟悉主題以及不熟悉主題的查詢行為上有明顯不同。在不熟悉的情境下，使用者會大量的嘗試並修正檢索詞彙並進行多次檢索，此一現象在 2 個檢索介面均有出現，而在 MAP 介面中尤為明顯。從查詢相似度來看，熟悉主題之查詢相似度會較不熟悉主題之相似度來的高，意即在熟悉主題的情況下，使用者之最終查詢與初始查詢間並不會產生明顯的變化。這乃是因為在熟悉主題的情境下，使用者對於其所需的資訊有一定程度的了解，對於資訊需求之表述亦能有效表達之，因此其初始及最終查詢彼此之間並不會有太大的差距。

而在不熟悉主題的情境下，使用者缺乏對於該主題的知識以及對於資訊需求的理解，因此只能粗略的進行資訊需求表述，而後與資訊檢索系統互動後而進一步改善其資訊需求，以檢索其所需的資訊。透過與系統互動的過程，重構使用者初始查詢，因此其查詢相似度會明顯低於熟悉主題之查詢。

而在瀏覽文件數及查詢次數，不同主題熟悉度並無明顯影響，此說明了無論在何種情境下，使用者仍須瀏覽一定數量的檢索結果來確認資訊需求以得到滿足，並經過一定次數的檢索過程才能驗證檢索詞彙之適用性。

二、不同介面會影響使用者查詢行為

查詢相似度是以使用者初始查詢及最終查詢互相比較，用以說明不同介面重構使用者之查詢的能力，其代表著使用者對於初始查詢之新增及移除詞彙之間的關係。

研究結果顯示，不同介面之查詢行為的確有所不同。MAP 介面之查詢相似度低於 PubMed 介面，其原因在於 MAP 系統能夠提供使用者大量的建議相關詞彙，因此使用者可從中檢驗初始查詢關鍵字之適用性，而在 PubMed 介面中，使用者僅能就檢索結果來驗證之，因此缺乏其他驗證查詢關鍵字之工具，其查詢詞彙之刺激有限，所以其最終查詢是故高度與初始查詢相似。在同樣的原因之下，PubMed 之而新增詞彙數及移除詞彙數亦較 MAP 介面來的少。

在瀏覽文件數及查詢次數不同介面並無明顯差異，此即是使用者在判斷檢索結果之適用性及檢索詞彙之適用性時，仍須經過瀏覽一定的檢索結果以及多次的查詢來驗證之，因此使用不同介面並不會造成影響。

三、不同主題熟悉度與不同介面對使用者查詢行為產生之交互作用

研究結果中主題熟悉度與不同介面對於查詢行為之交互作用雖然不顯著，然而根據使用者之評論以及相關變數之測量，結果顯示 MAP 介面在使用者較不熟悉檢索主題時可以發揮較大的作用，其能提供與檢索主題相關之建議詞彙，輔助使用者重構其資訊需求，釐清模糊的初始查詢。而在熟悉的情境下，使用者因為對於該主題有所認知，並掌握了一定數量的關鍵詞彙，因此認為 MAP 系統的輔

助作用不大，無助於改善資訊檢索結果。

四、 不同主題熟悉度未影響使用者查詢擴展行為

使用者在不同主題熟悉度下，其瀏覽及最終選用詞彙類別之比例並未呈現明顯之變化。在全新檢索及重複檢索情境之下，兩者之瀏覽及選用詞彙類別之比例均集中於相同之類別。而相同情境下，其瀏覽及選用之類別亦成一致的方向，僅在字數上有所變化。此一結果說明了主題之熟悉度對於使用者之查詢擴展行為並未有所影響。

五、 不同介面會影響使用者查詢擴展行為

在不同介面下，使用者瀏覽詞彙之類別及選用詞彙之類別的確有所不同。在瀏覽詞彙階段，PubMed 介面之詞彙類別比例確實不同於 MAP 介面。而在選用詞彙階段，不同介面下的選用詞彙類別比例也不盡相同，是故推論不同介面會影響使用者查詢擴展行為。

不同介面之瀏覽與選用詞彙類別之所以會產生差異的原因在於，在 PubMed 介面中，使用者的詞彙來自於己身的知識架構，因此其構成查詢的詞彙會受限於其對於主題的認知而有所偏重。而在 MAP 介面中，系統能夠以 MeSH 結構呈現主題內的完整的知識結構，使用者的檢索詞彙除了出自於己身對於主題的理解外，同時可透過點選與主題相關的 MeSH 詞彙進行檢索，其查詢之關鍵字能夠涵括檢索主題的各個面向，因此在不同介面下使用者瀏覽以及選用詞彙類別會有所不同。

六、 不同主題熟悉度與不同介面對使用者查詢擴展行為並未產生交互作用

在 PubMed 介面中，無論是全新檢索或重複檢索，其瀏覽以及選用之詞彙其

類別均集中於少數類別之中。是故，在 PubMed 介面下，主題熟悉度對於使用者查詢擴展行為並未有所影響。此外，使用者之瀏覽及選用詞彙類別呈現一致的方向，而其瀏覽字數亦與選用詞彙字數相近，此一現象是肇因於在 PubMed 介面中，使用者之相關詞彙來源僅有己身之知識架構以及檢索結果中的相關詞彙，因此使用者在輸入關鍵字的同時，其已認定該關鍵字與主題之相關性，因此其瀏覽之詞彙即為選用之詞彙。

而在 MAP 介面中，亦呈現相同的結果，全新檢索及重複檢索之瀏覽及選用之詞彙均集中於少數類別之中。因此推論主題熟悉度亦不會影響使用者在 MAP 介面中的查詢擴展行為。使用者之瀏覽及選用詞彙類別雖亦呈現一致的方向，然在字數上則是瀏覽詞彙數會高於選用詞彙數，此一現象乃是因為 MAP 提供之相關建議詞彙，是系統基於詞彙之共現性所提供的 MeSH 詞彙，其完整揭示了該主題領域內的相關詞彙。使用者雖從系統中取得了大量的相關詞彙，但該詞彙與使用者之需求的匹配程度仍須經過進一步的檢索才能驗證詞彙與主題之相關性。因此使用者會瀏覽大量的建議詞彙，而透過檢索結果來刪去不適合之詞彙，是故瀏覽之詞彙數多，選用之詞彙數少。

七、 MAP 介面能夠輔助使用者取得在大量檢索結果中排序較低的 相關文件

PubMed 在處理使用者之查詢時，是將使用者查詢之關鍵字以布林邏輯 AND 連接，然後傳回含有該關鍵詞之文獻，並以出版年份排序。然而在使用者輸入語意較為廣泛之詞彙時，則檢索結果也趨於龐大，而使用者又僅會瀏覽點選前數頁之檢索結果，使得使用者所需的相關文獻埋於大量檢索結果之底部，且無從取得。

在符合初始查詢及最終查詢均檢索同一資料集之案例之中，研究結果發現透

過 MAP 提供之建議詞彙，使用者能夠在相同檢索結果中，利用增加關鍵字的方式來縮小檢索結果，逐步逼近位於大量檢索結果底部之相關文獻，並從而取得其所需的資訊。

第二節 建議

綜合文獻分析、使用者查詢查詢擴展行為分析及使用者意見調查，研究者提出下列建議供資訊檢索研究者參考。

一、發展能夠支援不熟悉情境之檢索工具

資訊檢索學界傳統上認為資訊檢索系統只要存有使用者所需的相關資訊，而使用者有能力檢索之，則該系統必對使用者有所幫助。然而不同的使用者即使擁有相同的檢索主題，其對於檢索結果的要求以及對檢索主題的理解也不盡相同，因此認為檢索系統能夠一體適用(one fits all)的觀念，忽略了使用者在不同搜尋情境下如對主題的熟悉程度所產生的資訊需求。

White et al. (2006)提出探索式檢索(Exploratory Search)，認為在使用者缺乏領域內的專業知識時，亦無適當的關鍵字表述其資訊需求，於此檢索系統必須要提供瀏覽及探索的功能，以輔助使用者能夠進一步理解此一不熟悉領域內之知識架構，從而形塑其查詢以進行資訊檢索。

本研究之結果亦發現，詞彙建議系統在使用者於不熟悉情境之下能夠發會較大的效用，使用者可以藉由相關的建議詞彙理解其不熟悉的領域，並從中取得相關之關鍵字，而能檢得其所需的資訊。

由此可見，使用者對於檢索主題之熟悉度以及其對於檢索結果之要求均會影響其利用資訊檢索系統，因此未來之研究可進一步區分使用者之檢索情境，而設計與其需求相符之資訊檢索系統進行研究，以求提供使用者最大的支援(Kules &

Capra, 2008)。

二、採用使用者之真實資訊需求進行研究

傳統資訊檢索評估，主要是以研究者先建的檢索主題進行，並以事先決定的相關判斷做為比較的基準，而後以使用者的檢索結果進行比較。此種設計的優點在於結果較為客觀且容易計算，且不同系統間能互相比較，而缺點在於研究之生態效度較低，無法真實呈現使用者在真實情境下的資訊檢索，同時忽略了使用者在操作過程中與系統所產生的互動。而採用使用者真實資訊需求作為資訊檢索主題進行檢索，其目的即是在於提高研究之生態效度，期能夠進一步貼近使用者的真實檢索情境。

然而，採用使用者之真實資訊需求所產生之問題在於不同主題間缺乏比較之基準點。在研究者先建的主題下，可以比較在不同處理下所產生的檢索結果，從而比較資訊檢索系統之優劣。但採用使用者真實資訊需求做為檢索主題時，因為使用者行為會受到檢索主題的強烈影響，是故檢索結果亦會劇烈變化，則使得系統間難以進行比較。如何在控制與真實之間取得平衡，是未來研究的一大課題。

三、避免實驗產生次序效應

本研究以使用者之真實需求做為控制變項，研究之目的在於比較不同介面間使用者重構查詢之行為，所關注之焦點在於使用者所選擇的關鍵詞並透過重複量測的方式進行。然而由於實驗設計之故，使用者之檢索過程為 2 介面交互操作進行，因此初次實驗之介面詞彙選用經驗會強烈影響另一介面詞彙選用行為。研究結果顯示次序效應相當明顯。

若欲避免實驗產生次序效應，可透過延長實驗之時間降低次序效應。本研究產生次序效應之原因在於要求使用者於短時間內完成不同介面之檢索，因此前次

檢索之經驗會強烈影響後一檢索之過程。若將實驗時間拉長為數日間隔，並以更多的檢索需求要求使用者交錯檢索，則檢索經驗對於檢索過程之影響會明顯減少，使用者之反應將會來自於其原始的刺激。

第三節 進一步研究之建議

本研究囿於時間與人力之限制未能趨於完善，無法針對相關議題做全面性的探討。茲將進一步的研究建議陳述如下，以做參考。

一、 針對其他學科領域進行研究

檢索主題對於檢索過程及結果影響甚為劇烈，本研究僅以生物醫學等相關系所研究生為對象，針對其詞彙擴展行為進行分析，但未能擴及其他學科領域。建議未來之研究能夠擴及其他學科領域，進行使用者查詢擴展行為研究，比較學科間的異同，以對此一議題獲得全面性的理解。

二、 利用現有的檢索系統進行比較

本研究以自行設計之檢索介面以及 PubMed 為主要比較的對象，進行使用者查詢擴展行為之分析，並未擴及其他大型資料庫，其結果無法推論至其他資訊檢索系統。建議未來可關注市場上較為成熟之產品，進行相關檢索功能之比較及實驗，以期能將研究之結果應用於實際之系統之上。

三、 利用自動化的記錄工具

本研究透過 MORAE 軟體錄製使用者查詢擴展之行為，並於實驗後進行編碼，然並未使用自動化記錄工具如 log 檔紀錄等，因此使用者之查詢關鍵字偶有

錯落。建議未來針對使用者利用檢索系統行為之研究可以同時採用視窗錄影軟體以及自動化記錄軟體。

四、關注使用者利用詞彙彼此的關係

本研究以 MeSH 之架構分析使用者查詢擴展之詞彙，其關係判斷係以 MeSH 索引典之架構為主。但此一分析並無法理解使用者是在何種判斷下利用該詞彙。建議未來除以既有之資訊組織工具進行語意關係之判斷，亦可針對使用者利用詞彙之方式進行分析。



參考文獻

一. 中文部分

黃慕萱(民 85)。檢索系統評估。在資訊檢索(第十章, 頁 269-275)。台北市: 台灣學生。

二. 西文部分

Baeza-Yates, R. (1999). *Modern Information Retrieval*. New York: ACM Press.

Bates, M. J. (1979). Information search tactics. *Journal of the American Society for information Science*, 30(4), 205-214.

Belkin, N. J. (1993). Interaction with texts: Information retrieval as information-seeking behavior. *Information retrieval*, 93, 55–66.

Belkin, N. J., Chang, S., Downs, T., Saracevic, T., & Zhao, S. (1990). Taking account of user tasks, goals and behavior for the design of online public access catalogs. In *Proceedings of the 53rd ASIS annual meeting* (pp. 69–79). Medford, NJ: Learned Information.

Belkin, N. J., KELLY, D., Kim, G., Kim, J. Y., Lee, H. J., Muresan, G., Tang, M. C., et al. (2003). Query length in interactive information retrieval. In *Proceedings*

of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval (pp. 205-212). ACM New York, NY, USA.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982). ASK for information retrieval: Part I. Background and theory. *Journal of documentation*, 38(2), 61-71.

Borlund, P., & Ingwersen, P. (1997). The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation*, 53(3), 225-250.

Borlund, P., & Ingwersen, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. *Journal of Documentation*, 56(1), 71-90.

Cleverdon, C. W. (1967). The Cranfield tests on index language devices. In *Aslib proceedings*, 19 (Vol. 6, pp. 173–194).

Doms, A., & Schroeder, M. (2005). GoPubMed: exploring PubMed with the gene ontology. *Nucleic Acids Research*, 33(Web Server Issue), 783-786.

Dumais, S. T., & Belkin, N. J. (2005). The TREC interactive tracks: putting the user into search. In E. M. Voorhees & D. Harman (Eds.), *TREC: Experiment and*

evaluation in information retrieval (pp. 123-153). Cambridge, MA: MIT Press.

Efthimiadis, E. N. (1996). Query Expansion. In M. E. Williams (Ed.), *Annual Review of Information Systems and Technology (ARIST)* (Vol. 31, pp. 121-187).

Retrieved from

<http://faculty.washington.edu/efthimis/pubs/Pubs/qe-arist/QE-arist.html>

Efthimiadis, E. N. (2000). Interactive query expansion: a user-based evaluation in a relevance feedback environment. *Journal of the American Society for Information Science*, 51(11), 989-1003.

Fidel, R. (1984). Online searching styles: A case-study-based model of searching behavior. *Journal of the American Society for Information Science*, 35(4), 211-221.

Goetz, T., & von der Lieth, C. W. (2005). PubFinder: a tool for improving retrieval rate of relevant PubMed abstracts. *Nucleic acids research*, 33(Web Server Issue), 774-778.

Harman, D. (1988). Towards interactive query expansion. In *Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 321-331). ACM New York, NY, USA.

Harman, D. (1992). Relevance feedback revisited. In *Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 1-10). ACM New York, NY, USA.

Hearst, M., Elliott, A., English, J., Sinha, R., Swearingen, K., & Yee, K. P. (2002). Finding the flow in web site search. *Communications of the ACM*, 45(9), 42-49.

Hersh, W., Pentecost, J., & Hickam, D. (1996). A task-oriented approach to information retrieval evaluation. *Journal of the American Society for Information Science*, 47(1), 50-56.

Hersh, W. R., & Hickam, D. H. (1998). How well do physicians use electronic information retrieval systems? A framework for investigation and systematic review. *JAMA*, 280(15), 1347-1352.

Ingwersen, P., & Järvelin, K. (2005). *The turn: integration of information seeking and retrieval in context*. Dordrecht, Netherlands: Springer.

Jansen, B. J., Spink, A., Bateman, J., & Saracevic, T. (1998). Real life information retrieval: a study of user queries on the Web. In *ACM SIGIR Forum* (Vol. 32, pp. 5-17).

Joho, H., & Jose, J. M. (2006). Slicing and dicing the information space using local contexts. In *Proceedings of the 1st international conference on Information interaction in context* (pp. 66-74). Copenhagen, Denmark: ACM.
doi:10.1145/1164820.1164836

Kristensen, J., & Järvelin, K. (1990). The effectiveness of a searching thesaurus in free text searching of a full-text database. *International Classification*, 17(2), 77–84.

Kules, B., & Capra, R. (2008). Creating exploratory tasks for a faceted search interface. In *Second Workshop on Human-Computer Interaction (HCIR 2008)*.

Lin, X. (1999). Visual MeSH. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 317-318). ACM New York, NY, USA.

Manning, C. D. (2008). *Introduction to Information Retrieval*. New York: Cambridge University Press.

Pennanen, M., & Vakkari, P. (2003). Students' conceptual structure, search process, and outcome while preparing a research proposal: a longitudinal case study. *Journal of the American Society for Information Science and Technology*, 54(8), 759–770.

- Perez-Iratxeta, C., BORK, P., & Andrade, M. A. (2001). XplorMed: a tool for exploring MEDLINE abstracts. *Trends in biochemical sciences*, 26(9), 573-575.
- Plikus, M. V., Zhang, Z., & Chuong, C. M. (2006). PubFocus: semantic MEDLINE/PubMed citations analytics through integration of controlled biomedical dictionaries and ranking algorithm. *BMC bioinformatics*, 7(1), 424.
- Pollitt, A. S. (1998). The key role of classification and indexing in view-based searching. *International cataloguing and bibliographic control*, 27(2), 37-40.
- Robertson, S. E., & Hancock-Beaulieu, M. M. (1992). On the evaluation of IR systems. *Information Processing and Management*, 28(4), 457-466.
- Ruthven, I. (2008). Interactive information retrieval. In B. Cronin (Ed.), *Annual Review of Information Science & Technology* (Vol. 42, pp. 43-91). Medford, NJ: Information Today.
- Saracevic, T. (1997). The stratified model of information retrieval interaction: Extension and applications. In *Proceedings of the American Society for Information Science* (Vol. 34, pp. 313–327).
- Saracevic, T., & Kantor, P. (1988a). A study of information seeking and retrieving. II.

Users, questions, and effectiveness. *Journal of the American Society for Information Science*, 39(3), 177–196.

Saracevic, T., & Kantor, P. (1988b). A study of information seeking and retrieving. III. Searchers, searches, and overlap. *Journal of the American Society for information Science*, 39(3), 197–216.

Saracevic, T., Kantor, P., Chamis, A. Y., & Trivison, D. (1988). A study of information seeking and retrieving: Background and methodology. *Journal of the American Society for information Science*, 39(3), 161-176.

Shiri, A., & Revie, C. (2006). Query expansion behavior within a thesaurus-enhanced search environment: A user-centered evaluation. *Journal of the American Society for Information Science and Technology*, 57(4).

Shiri, A., Revie, C., & Chowdhury, G. (2001). Thesaurus-assisted search term selection and query expansion: a review of user-centred studies. *Knowledge Organization*, 29(1), 1–19.

Spink, A. (1995). Term relevance feedback and mediated database searching: implications for information retrieval practice and systems design. *Information Processing and Management*, 31(2), 161–171.

Spink, A., & Saracevic, T. (1997). Interaction in information retrieval: selection and effectiveness of search terms. *Journal of the American Society for Information Science*, 48(8), 741–761.

Su, L. T. (1992). Evaluation measures for interactive information retrieval. *Information Processing and Management*, 28(4), 503–516.

Su, L. T. (1994). The relevance of recall and precision in user evaluation. *Journal of the American Society for Information Science*, 45(3), 207–217.

Su, L. T. (2003). A comprehensive and systematic model of user evaluation of web search engines: Ii. an evaluation by undergraduates. *Journal of the American Society for Information Science and Technology*, 54(13), 1193–1223.

Vakkari, P. (1999). Task complexity, problem structure and information actions integrating studies on information seeking and retrieval. *Information Processing and Management*, 35(6), 819-837.

Vakkari, P. (2000). Cognition and changes of search terms and tactics during task performance: A longitudinal case study. In *Proceedings of the RIAO 2000 Conference* (pp. 894–907).

Vakkari, P. (2001). A theory of the task-based information retrieval process: a

summary and generalisation of a longitudinal study. *Journal of Documentation*, 57(1), 44–60.

Vakkari, P., & Hakala, N. (2000). Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, 56(5), 540–562.

Vakkari, P., Pennanen, M., & Serola, S. (2003). Changes of search terms and tactics while writing a research proposal A longitudinal case study. *Information Processing and Management*, 39(3), 445–463.

White, R. W., Kules, B., & Drucker, S. M. (2006). Supporting Exploratory Search, Introduction, Special Issue, Communications of the ACM. *Communications of the ACM*, 49(4), 36-39.

White, R. W., Ruthven, I., & Jose, J. M. (2005). A study of factors affecting the utility of implicit relevance feedback. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 35-42). ACM New York, NY, USA.

Xie, I. (2008). *Interactive Information Retrieval in Digital Environments*. Hershey, PA: IGI Pub.

附錄一 前測問卷

感謝您同意參與本研究，本研究試圖了解使用者參與「醫學資料庫詞彙建議研究」的相關資訊行為，藉以評估該介面系統的使用效能。在介面測試中，您將參與以下程序，請您仔細閱讀，若有問題也可以提出：

介面測試前：

1. 您將會閱讀及簽署本同意書，您可以在這個階段提出任何有關的問題。您同樣能夠保留此同意書做為未來參考。
2. 您將會填寫一份簡短的問題卷，藉以了解您先前的背景資料和搜尋經驗。

介面測試開始時：

1. 本測試於您登入本資料庫之時開始。
2. 本測試的輔助程式將會紀錄您在使用介面時的螢幕畫面，在測驗的期間您的使用過程將會被記錄下來。

測試的任務：

您將與我們所提供的介面互動並且搜尋您自己的檢索問題。您的主要任務有二：一為找出八筆（兩個介面各四筆）對您最有用的文獻；二為在與介面互動之後，寫下您認為最能夠表達您的資訊需求的檢索詞彙，您一共有 30 分鐘的時間來完成這兩個任務，若提早完成可舉手告知研究者。

介面測試進行中：

每一階段操作結束，您將針對該階段操作經驗填答一份後測問卷，並且為前五十筆的檢索結果評分。

本研究立易於了解使用者的資訊使用行為及該資料庫介面的使用性，測驗的過程以及結果並非測驗您的搜尋能力，而在於測試介面的使用度，請您放鬆心情作答。您的個人資料在研究報告結果中將會以匿名的方式處理，所有的資料將只用在本研究。

本研究乃自願性質，為時約一個半小時，結束後您將獲得新台幣三百八十元的報酬。您若中途放棄測試將不會得到任何酬勞。

我，_____已經確實閱讀以及了解這份同意書，並且同意參本研究的測驗。

_____（使用者簽名） 日期： 年 月 日

_____（施測者簽名） 日期： 年 月 日

背景資料

1 · 請問您的職業是？

- 大學生
- 研究生
- 研究人員
- 教師
- 醫技／醫療護理人員
- 其他

2 · 性別

- 男 女

3 · 請問您的研究／工作專業領域為何？如物理治療，生物化學，皮膚科等

4 · 請問您使用 PubMed 或 Medline 的頻率為何？

從未使用	兩個月一次	每個月一次	兩周一次	每周一次	每天使用

5.請問您使用 Google 學術搜尋(Google Scholar)的頻率為何？

從未使用	兩個月一次	每個月一次	兩周一次	每周一次	每天使用

6.請問您對醫學標題表(MeSH)的熟悉程度為何？直接圈選下排數字作答。

完全不熟悉							非常熟悉
0	1	2	3	4	5	6	

7.請問當您使用 PubMed 或 Medline 檢索問題時，您運用進階檢索技巧(如布林邏輯 AND / OR / NOT ，查詢擴展，限制條件等)的頻率為何？

從未使用							每次都
0	1	2	3	4	5	6	

前測問卷

1.請敘述您想要檢索的問題。

2.假設下面的空格代表了 檢索欄位，你會輸入哪些檢索詞來檢索這個問題呢？

--

3.請問您對此檢索問題涉及的主題熟悉程度為何？

完全不熟悉							非常熟悉
0	1	2	3	4	5	6	

4.請問您對檢索結果完整度的要求有多高？

有一篇代表性文獻即可							完整的文獻回顧
0	1	2	3	4	5	6	

5.請評估您目前的檢索詞彙能夠完整表達您的資訊需求的程度。

完全不能							完全能
0	1	2	3	4	5	6	

6.如果今天您無法親自檢索這個問題，必須請他人代為檢索，您認為他人的檢索結果能夠滿足您的資訊需求的程度是？

完全不能							完全能
0	1	2	3	4	5	6	

後測問卷

請跟據您上一個階段的檢索經驗回答下列問題，直接圈選下排數字以作答：

1. 請問您認為該介面容易使用的程度為何？

完全不易			部分容易		非常容易
1	2	3	4	5	6

2. 請問您認為此介面對檢索的有用程度為何？

完全沒用			部分有用		非常有用
1	2	3	4	5	6

3. 請問您對使用此介面檢索的滿意程度為何？

非常不滿意			部分滿意		非常滿意
1	2	3	4	5	6

4. 請問您對檢索出的資料滿意程度為何？

非常不滿意			部分滿意		非常滿意
1	2	3	4	5	6

5. 如果此介面對您而言有幫助，請問是在哪些方面有助於檢索呢？請指出以下敘述和您的經驗的符合程度。

a. 介面提供的關鍵字給予我新的想法，使我可能因此發展出新的研究方向。

完全不符合			部分符合		非常符合
1	2	3	4	5	6

b. 介面提供的關鍵字使我的檢索詞更加明確地表達我原本的問題。

完全不符合			部分符合		非常符合
1	2	3	4	5	6

c. 此介面呈現了資料庫的相關文獻的範圍和結構

完全不符合			部分符合		非常符合
1	2	3	4	5	6

d. 此介面幫助我處理龐大的檢索結果

完全不符合			部分符合		非常符合
1	2	3	4	5	6

End of the search questionnaire

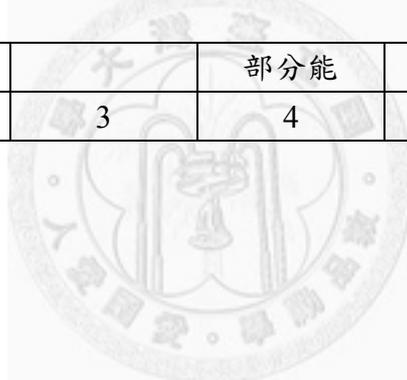
1.經過方才與介面的互動之後，您或許將修改一開始的檢索詞。請在下方空格中寫下現在您認為最能理想地表達您原本的檢索問題的檢索詞彙。

2 · 請問您現在對這個檢索問題涉及的主題熟悉程度為何？

完全不熟悉			部分熟悉		非常熟悉
1	2	3	4	5	6

3 · 請您評估您目前的檢索詞彙能夠完整表達您的資訊需求的程度。

完全不能			部分能		完全能
1	2	3	4	5	6



附錄二 正式實驗問卷

感謝您同意參與本研究，本研究試圖了解使用者參與「醫學資料庫詞彙建議研究」的相關資訊行為，藉以評估該介面系統的使用效能。本研究立意於了解使用者的資訊尋求行為及該資料庫介面的使用性，測驗的過程以及結果並分測研您的搜尋能力，請您放鬆心情作答。您的個人資料在研究報告結果終將會以匿名的方式處理，所有的資料將只用在本文中。

在介面測試時，您將參與以下程序，請您仔細閱讀，若有問題也可以提出：

介面測試前：

1. 您將會閱讀及簽署本同意書，您可以在這個階段提出任何有關的問題。
2. 您將會填寫一份簡短的問題卷，藉以了解您先前的背景資料和搜尋經驗。

介面測試開始時：

1. 本測試的輔助程式將會紀錄您在使用介面時的螢幕畫面，在測驗的期間您的使用過程將會被記錄下來。
2. 在檢索過程中請嘗試解釋選取檢索詞的理由，您的對話將被錄音。

測試的任務：

您將與我們所提供的介面互動並且搜尋您兩個自己的檢索問題。請盡可能至少選擇一個您比較不熟悉的主題。

1. 您有三分鐘的時間來完成每個問題，在檢索每一個問題時，您各有十五分鐘的時間分別使用兩種檢索模式，一為一般PubMed介面(無詞彙建議功能)，另一有詞彙建議功能，若提早完成可舉手告知研究者。
2. 在搜尋過程中，您的主要任務有二：一為找出並存取二十筆對您最有用的文獻(每一檢索模式各十筆； $10 \times 2 = 20$)；二為在與介面互動之後，寫下您認為最能夠表達您的資訊需求的檢索詞彙。

介面測試後：

1. 每一問題搜尋結束後，您將針對該檢索經驗填答一份後測問卷，並且找到的每筆資料評分。

本研究乃自願性質，為時約一個半小時，結束後您將獲得新台幣三百八十元的報酬。您若中途放棄測試將不會得到任何酬勞。

我，_____已經確實閱讀以及了解這份同意書，並且同意參本研究的測驗。

_____ (使用者簽名) 日期： 年 月 日

_____ (施測者簽名) 日期： 年 月 日

背景資料

1 · 請問您的主要身分是？

- 大學生
- 研究生
- 研究人員
- 教師
- 醫技／醫療護理人員
- 其他

2 · 性別

- 男 女

3 · 請問您的研究／工作專業領域為何？如物理治療，生物化學，皮膚科等

4 · 請問您使用 PubMed 或 Medline 的頻率為何？

從未使用	兩個月一次	每個月一次	兩周一次	每周一次	每天使用

5.請問您使用 Google 學術搜尋(Google Scholar)的頻率為何？

從未使用	兩個月一次	每個月一次	兩周一次	每周一次	每天使用

6.請問您對醫學標題表(MeSH)的熟悉程度為何？直接圈選下排數字作答。

完全不熟悉						非常熟悉
0	1	2	3	4	5	6

7.請問當您使用 PubMed 或 Medline 檢索問題時，您運用進階檢索技巧(如布林邏輯 AND / OR / NOT ，查詢擴展，限制條件等)的頻率為何？

從未使用						每次都用
0	1	2	3	4	5	6

前測問卷

1.請敘述您想要檢索的問題。

2.假設下面的空格代表了 檢索欄位，你會輸入哪些檢索詞來檢索這個問題呢？

--

3.是否曾經檢索過這個問題？

是 否

4.請問您對此檢索問題涉及的主題熟悉程度為何？

完全不熟悉				非常熟悉		
0	1	2	3	4	5	6

4.請問您對檢索結果完整度的要求有多高？

有一篇代表性文獻即可				完整的文獻回顧		
0	1	2	3	4	5	6

5.如果今天您無法親自檢索這個問題，必須請他人代為檢索，您認為他人的檢索結果能夠滿足您的資訊需求的程度是？

完全不能				完全能		
0	1	2	3	4	5	6

後測問卷

請跟據您剛剛利用介面的經驗回答下列問題，直接圈選下排數字以作答：

1.經過方才與介面的互動之後，您或許將修改一開始的檢索詞。請在下方空格中寫下現在您認為最能理想地表達您原本的檢索問題的檢索詞彙(可加入布林邏輯，如：AND,OR,NOT)。

--

1.請問您認為該介面容易使用的程度為何？

完全不易			部分容易			非常容易
0	1	2	3	4	5	6

2.請問您認為此介面對檢索的有用程度為何？

完全沒用			部分有用			非常有用
0	1	2	3	4	5	6

3.整體而言，您對使用關鍵詞推薦介面所搜尋到的結果滿意程度為？(請針對剛剛的檢索經驗作答)

完全不滿意				部分滿意		非常滿意
0	1	2	3	4	5	6

4.請問您對推薦詞彙的滿意程度為何？

完全不滿意				部分滿意		非常滿意
0	1	2	3	4	5	6

5.相較於傳統介面，若關鍵詞推薦介面對您有所幫助，請問是在哪些方面有助於檢索呢？請指出以下敘述和您的經驗的符合程度。

a.介面提供的關鍵字給予我新的想法，使我可能因此發展出新的研究方向。

完全不符合			部分符合			非常符合
0	1	2	3	4	5	6

example

b. 介面提供的關鍵字使我的檢索詞更加明確地表達我原本的問題。

完全不符合			部分符合			非常符合
0	1	2	3	4	5	6

Example

c. 此介面呈現了資料庫的相關文獻的範圍和結構

完全不符合			部分符合			非常符合
0	1	2	3	4	5	6

Example

d. 此介面幫助我處理龐大的檢索結果

完全不符合			部分符合			非常符合
0	1	2	3	4	5	6

請判斷您所檢索初的每一篇文件與您的檢索主題相符程度(relevance)與提供您之前不知道(novelty)的程度。由0到6，0代表完全不6代表非常。

Topic /Interface

#	Seen Before?(Y/N)	Relevance	Novelty
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

End of the search questionnaire

1.請您評估您在檢索之前所填的理想檢索詞彙能夠完整表達資訊需求的程度

完全不能						完全能
0	1	2	3	4	5	6

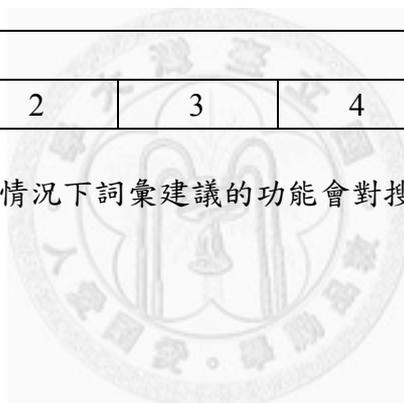
2.請您評估在與 PubMed 互動後所填的理想檢索詞彙能夠完整表達您的資訊需求的程度。

完全不能						完全能
0	1	2	3	4	5	6

3.請您評估在與 MAP 互動後所填的檢索詞彙能夠完整表達您的資訊需求的程度。

完全不能						完全能
0	1	2	3	4	5	6

4.請您談談您認為在什麼情況下詞彙建議的功能會對搜尋有幫助



附錄三 受試者檢索問題

USER REVISIT	REQUEST
2	全新檢索 Legionella sp. 與氣象因子的關係
2	重複檢索 呼吸系統與空氣中微生物的關係
6	重複檢索 ER 的 CPG island 分布
6	重複檢索 HBV 影響的 signaling pathway
7	全新檢索 不同亞型流感病毒血球凝集素的鍵結親和力
7	重複檢索 近年來流感病毒神經胺酸酶的變異性及其在病毒複製過程中所扮演的角色
8	全新檢索 the effect of TMV-encoded p126 protein in suppression of VIGS
8	重複檢索 the role of Grb7 regulation of cellular functions and tumor formation
9	全新檢索 果膠酯酶抑制劑
9	重複檢索 大豆異黃酮及其相關水解酵素
10	全新檢索 以導電高分子製作奈米及尺寸表面，觀察對心肌細胞行為之影響
10	重複檢索 利用表面地形的改變觀察肝細胞型態及功能之影響關聯性
11	全新檢索 心理衛生相關醫療資源
11	重複檢索 肺癌病患接受外科手術後之醫療品質
12	重複檢索 traumatic brain injury(TBI) & dementia & Alzheimer's disease(AD)
12	重複檢索 traumatic brain injury(TBI) & social behavior & theory of mind(ToM)
13	重複檢索 總額預算制度對於抑制醫療費用成長之成效
13	重複檢索 paruresis 之危險因子
14	全新檢索 如何分離犬隻的 T-reg Cells
14	重複檢索 TGF-B 如何影響 canine dendritice cell 的 mafuration
15	重複檢索 光動力療法治療與奈米粒子包覆 PPIX
15	重複檢索 癌症幹細胞的抗藥性
16	重複檢索 asthma 的體內免疫反應形況
16	重複檢索 類風濕性關節炎的動物模式
17	全新檢索 biochip 用於檢測癌症方面
17	重複檢索 preparing nanoparticle, drug delivery 如何製備 nanoparticle, 包附化療藥物 dox
18	全新檢索 奈米粒子在 EPR 效應下，是否有效果上的提升
18	重複檢索 了解 hydrogel 在藥物釋放方面所具有的特性及效果
19	全新檢索 嗎啡計量的耐受性和相關的研究數值及作用途徑
19	重複檢索 rumination depression
20	重複檢索 protein docking prediction

USER REVISIT	REQUEST
20	重複檢索 protein function prediction using network property
21	全新檢索 drug resistance in mTOR signaling
21	重複檢索 Rheb
22	重複檢索 探討 GSK3-B 在肺炎細胞中的調控
22	重複檢索 應用蛋白質體學探討丹參酚酸 B 調控血管內皮細胞
23	全新檢索 吞嚥障礙患者之電刺激療法
23	重複檢索 研究使用於失語症患者的健康生活品質
24	全新檢索 PBDE 的毒性資料
24	重複檢索 針對 fFuran 分析方法
25	全新檢索 bioreactor 生物反應器
25	重複檢索 幾丁聚醣相關文章
26	全新檢索 low back pain, LBC, prevalence
26	重複檢索 芳香烴受體功能
27	重複檢索 斑馬魚的肌肉發育過成為何
27	重複檢索 lactoferricin 的功能為何
28	重複檢索 the repair of hypoxanthine in DNA
28	重複檢索 the repair of uracil in double stranded DNA
29	重複檢索 database
29	重複檢索 text-mining
30	重複檢索 腦內神經傳導物質對睡眠的影響
30	重複檢索 兒茶素的抗癌機制
31	重複檢索 cell cycle 的機制
31	重複檢索 VP28 的影響及預防
32	重複檢索 醫院臨床癌症病患的營養狀況
32	重複檢索 放射線治療癌症患者營養介入成效
33	重複檢索 關於神經細胞於胚胎發育及 larva stage 發育的關聯機制
33	重複檢索 apoptosis 的過程中 engulfment 的機制
34	全新檢索 scan zebrafish image by using ultrasound
34	重複檢索 the design of high-temperature superconducting coil for magnetic resonance imaging
35	重複檢索 the asymmetrical development of human embryo
35	重複檢索 the relation between the genetic variants of the COMT gene and schizotypal personality
36	重複檢索 了解 probiotics 的定義及內涵，以及在水產養殖上用於抵抗病毒病之功用或實力
36	重複檢索 NNN(nervous necrosis virus): 了解時空背景，近年新發現等。

USER REVISIT	REQUEST
37	全新檢索 the effect of Brevetoxin on development, heart rate & c-fos expression
37	重複檢索 effect of Brevetoxin on Bovine Adrenal Chromaffin cell
38	全新檢索 the study of arabidopsis promote of ubiquitin genes
38	重複檢索 the functional study of Toc33 in Arabidopsis
39	全新檢索 micelle
39	重複檢索 study of cancer stem cell
40	全新檢索 骨質疏鬆症與運動習慣的相關性
40	全新檢索 營養與基因的交互作用對運動表現之影響
41	全新檢索 gsps
41	重複檢索 glutaredoxin
42	全新檢索 flower color examination o Oncidium
42	重複檢索 Function of MADs-box genes in sweet potato
43	全新檢索 using MRI techniques to evaluate the leakage pattern of blood brain barrier in ischemic stroke model
43	重複檢索 source memory recollection ERP
44	全新檢索 motion perception
44	重複檢索 consciousness implicit-implicit processing
45	全新檢索 narcolepsy sleepdisorder orexin
45	重複檢索 app flatillin,lipiraft
46	全新檢索 aflatoxin
46	重複檢索 hypo cholesterolemic drug
47	全新檢索 關於 HIV 之治療
47	重複檢索 SVD 之單源抗體應用
48	全新檢索 目前組織工程在心肌梗塞治療上的幫助與應用
48	重複檢索 cartilage repair and tissue engineering

附錄四 次序效應案例

先 PubMed 介面後 MAP 介面之次序效應案例

ID	REVISIT	interface	ORIGINAL	FINALIZED	ADDED	Removed	Overlap
2	Y	PubMed	respiratory, airborne	respiratory AND airborne microbiology	microbiology		microbiology
		MAP		Respiratory airborne microbiology ,Cough, Sneezing	microbiology, cough, sneezing		
17	N	PubMed	biochip AND cancer	biochip AND diagnosis AND carcinoma	diagnosis, carcinoma	cancer	diagnosis
		MAP		biochip AND cancer AND diagnosis AND Breast	diagnosis, breast		
18	Y	PubMed	hydrogel AND drug release	hydrogel AND delivery OR characterization	delivery, characterization	drug release	drug release
		MAP		hydrogel AND delivery AND drug carriers	drug carriers	drug release	
22	Y	PubMed	proteomic, Sal B, HUVEC	PA1-1; Sal B; endothelial cell	endothelial cell	proteomic; HUVEC	proteomic; HUVEC
		MAP		Sal B; PA1- 1; Card; vascular system; hypertension	Card; vascular system; hypertension	proteomic; HUVEC	
23	N	PubMed	dysphasia, electro stimulation	Dysphasia; laryngeal	laryngeal	electro stimulation	electro stimulation
		MAP		electric stimulation therapy	electric stimulation therapy	dysphgia, electro stimulation	
23	Y	PubMed	health-related quality of life(HRQOL) , stroke or aphasia	SAQOL; Aphasia	SAQOL	health-related quality of life (HRQOL), stroke	SAQOL,HRQOL,sroke
		MAP		SAQOL	SAQOL	health-related quality of life (HRQOL), stroke, aphasia	

25	N	PubMed	bioreactor, cell	bioreactor, cell culture, rotating bioreactor	culture, rotating bioreactor		culture
		MAP		bioreactor, protein, cell culture	protein, culture		
25	Y	PubMed	chitin, chiesan	chitesan particle drug delivery	particle drug delivery	chitin	chitin
		MAP		chitesan, drug therapy, blood, biomaterial	drug therapy, blood, Biomaterial	chitin	
26	N	PubMed	low back pain, prevalence	(low back pain) and (prevalence) and (occupational)	occupational		occupational,
		MAP		(low back pain) and (prevalence) and (occupational)	occupational		
26	Y	PubMed	AhR Polymorphism review	(aryl hydrocarbon receptor repressor) AND (polymorphism)	hydrocarbon receptor repressor	review, AhR	review, repressor
		MAP		AhR repressor	repressor	review	
30	Y	PubMed	cholinergic	cholinergic system and sleep	system, sleep		system, sleep
		MAP		cholinergic system AND sleep stages	system, sleep stages		
34	Y	PubMed	magnetic resonance image AND high-temperature superconductor	magnetic resonance image AND high temperature superconductor		diffusion tensor imaging	diffusion tensor imaging
		MAP	re superconductor AND Diffusion Tensor Imaging	magnetic resonance image AND high temperature superconductor		diffusion tensor imaging	
37	N	PubMed	Brevetoxin, c-fos. expression	brevetoxin& development, brevetoxon & c-fos, brevetoxin & heart rate	heart rate, development	expression	heart rate, development, expression
		MAP		brevetoxin, c-fos, development, heart rate	heart rate, development	expression	
39	N	PubMed	micelle, preparation,	micelle		preparation, polymer	preparation

		MAP	polymer	micelle, drug carrier, polymer	drug carrier	preparation	
46	Y	PubMed	cholesterol and drug	hypercholesterolemia	hypercholesterolemia	cholesterol	hypercholesterolemia, cholesterol
		MAP		hypercholesterolemia, LDL cholesterol	hypercholesterolemia, LDL cholesterol	cholesterol	
47	N	PubMed	HIV	AIDS, therapy	AIDS, THERAPY	HIV	AIDS, THERAPY, HIV
		MAP		AIDS, THERAPY, DRUG	AIDS, THERAPY, DRUG	HIV	

先MAP介面後PubMed介面之次序效應案例

ID	REVISIT	interface	ORIGINAL	FINALIZED	ADDED	Removed	Overlap
6	Y	MAP	HBV GSK JNK ERK Src	HBV AND signaling	signaling	GSK JNK ERK Src	signaling ,GSK JNK ERK Src
		PubMed		HBV AND signaling & HCC	signaling, HCC	GSK JNK ERK Src	
7	N	MAP	hemagglutinin, binding affinity; influenza	hemagglutinin, binding affinity, influenza, subtype, H1N1;H3N2; H5N1	subtypes, H1N1;H3N2; H5N1		subtypes
		PubMed		hamagglutinin binding affinity; ingluenza,subtypes	subtypes		
8	Y	MAP	growth receptor bound protein 7, cell migration AND FAK AND Grb7	Grb7 AND integrin AND FAK AND cell migration	cell integrin	growth receptor bound protein 7	growth receptor bound protein 7
		PubMed		Grb7 AND cell migration AND adhesion	adhesion	growth receptor bound protein 7, FAK	
9	N	MAP	pectinase, inhibitor, pectin	pectinase		inhibitor	inhibitor
		PubMed		methylesterase	methylesterase	pectinase, inhibitor, pectin	
10	N	MAP	conductive polymer,	polypyrrole, cardiomyocytes	polypyrrole	conductive polymer	conductive polymer

		PubMed	cardiomyocyte	polyethylene, cardiomyocytes, glycol	polyethylene, glycol	conductive polymer	
10	Y	MAP	nanotextured, hepatocyte	topography, hepatocytes	topography	nanotextured	nanotextured
		PubMed		hepatocytes, cytoskeleton	cytoskeleton	nanotextured	
12	Y	MAP	TBI & dementia & AD	TBI & dementia & cognitive impairment	cognitive impairment	AD	AD
		PubMed		TBI & dementia		AD	
12	Y	MAP	TBI & social behavior &	TBI & ToM & empathy		social behavior	social behavior
		PubMed	ToM & empathy	TBI & ToM & empathy		social behavior	
13	Y	MAP	global budget, expenditure target, expenditure capping. Medical	global budget system AND health expenditure Or capitation fee OR PPS NOT adjustment	health, capitation, fee, PPS, adjustment, system	target, capping, medical, MCPI	system, health, capitation, target, capping, medical, MCPI
		PubMed	expenditure, MCPI	health expenditure growth AND capitation OR prospective payment system	health, growth, capitation, system	target, capping, medical, MCPI	
13	Y	MAP	paruresis, avoidant paruresis, psychogenic urinary retention, bathroom	psychogenic urinary retention AND causes NOT treatment	causes, treatment	avoidant paruresis, bathroom phobia, risk factors, predicting factors	treatment, avoidant paruresis, bathroom phobia, risk factors, predicting factors
		PubMed	phobia, risk factors, predicting factors	parursis OR psychogenic urinary retention NOT treatment	treatment	avoidant paruresis, bathroom phobia, risk factors, predicting factors	
15	Y	MAP	photo dynamic therapy and nanoparticle and PPIX	photodynamic therapy and nanoparticle and cancer or PPIX	cancer		cancer

		PubMed		photodynamic therapy and nanoparticle and cancer	cancer	PPIX	
15	Y	MAP	cancer stem cell and anti-drug resistance	cancer stem cell and drug resistance or MDR	MDR	anti-drug	anti-drug
		PubMed		cancer stem cell and drug resistance		anti-drug	
16	Y	MAP	asthma, immune, immunology	asthma, immune, system, cytokine	cytokine, system	immunology	cytokine
		PubMed		asthma, immune, immunology, cytokine	cytokine		
20	Y	MAP	function prediction, network, system biology	computation protein function prediction, network, gene ontology	computation protein, gene ontology	systems biology	systems biology
		PubMed		protein function network topology graph, interaction	topology graph, interaction	systems biology	
21	Y	MAP	Rheb cell growth OR proliferation	Rheb		cell growth proliferation	cell growth proliferation
		PubMed		Rheb		cell growth proliferation	
32	Y	MAP	Radiotherapy, nutrition intervention, malnutrition	radiotherapy, nutrition status, malnutrition, nutrition intervention, cancer	nutrition status, cancer		cancer
		PubMed		radiotherapy, malnutrition, nutrition intervention, cancer	cancer		
33	Y	MAP	neuron development AND asymmetrical cell division	cell division, neuron development		asymmetrical	asymmetrical
		PubMed		neuroblasts, cell division, development	neuroblasts	neuron, asymmetrical	
33	Y	MAP	apoptosis, programmed cell death, engulfment, recognition	apoptosis, engulfment, recognition		programmed cell death	programmed cell death
		PubMed		apoptosis, engulfment, recognition		programmed cell death	

35	Y	MAP	asymmetry, embryo	asymmetry, laterality, embryo, PITX2	laterality, PITX2		PITX2
		PubMed		NODAL, PIXT2, embryo, asymmetry	NODAL, PIXT2		
36	Y	MAP	NNN, piscine	betanodavirus, fish, RNA-virus infection, biological control, diagnosis	betanodaviruses, fish, RNA-virus infection, biological control, diagnosis	NNN, piscine	betanodaviruses, fish, diagnosis, control, NNN, piscine
		PubMed		betanodavirus, fish, diagnosis, control	betanodaviruses, fish, diagnosis, control	NNN, piscine	
42	N	MAP	flower AND Oncidium AND color, anthocyanin AND Oncidium, orchid AND flower And color	orchid And color, pigmentation AND flower, anthocyanin, carotenoid	pigmentation, carotenoid	oncidium	pigmentation
		PubMed		anthocyanin flower color, oncidium flower color, pigmentation	pigmentation		
43	N	MAP	MRI and ischemic stroke	ischemic stroke, blood-brain barrier, reperfusion, MCAO	blood-brain barrier, reperfusion, MCAO	MRI	blood-brain barrier, reperfusion, MCAO, MRI
		PubMed		ischemic stroke, blood brain barrier, reperfusion, MCAO	blood-brain barrier, reperfusion, MCAO	MRI	
44	Y	MAP	implicit processing, subliminal, unawareness, unconscious	unconscious, unawareness, binocular rivalry, attention, implicit processes	binocular rivalry	subliminal	subliminal
		PubMed	attention	implicit, unawareness, attention, unconscious		subliminal, processing	

48	Y	MAP	cartilage repair and tissue engineering	articular cartilage and biomaterial	articular cartilage, biomaterials	cartilage repair, tissue engineering	biomaterials, cartilage repair, tissue engineering
		PubMed		osteoarthritis and biomaterial	osteoarthritis, biomaterials	cartilage repair, tissue engineering	

