

視覺輔助相關回饋生物醫學資訊檢索雛型系統之建置

邱子恒 Tzu-Heng Chiu

臺北醫學大學通識教育中心副教授兼圖書館副館長
Associate Professor, Center of General Education,
Associate Direct of the Library,
Taipei Medical University

蔣以仁 I-Jen Chiang

臺北醫學大學醫學資訊研究所副教授
Associate Professor, Graduate Institute of Medical Informatics,
Taipei Medical University

【摘要】

美國國家醫學圖書館的生醫書目資料庫 PubMed，收錄全世界 1960 年代以來至今的核心生物醫學論文書目超過 2 千萬，且每天都增加 2000-4000 筆的書目。該資料庫呈現檢索結果的方式，是將使用者檢索到之文獻書目以出版年代之新穎性來排序，但最新的文獻不見得是品質最高、最符合使用者需求的文獻，因此使用者面對數量龐大的搜尋結果時，很需要依「主題」來篩選相關資料。本研究運用文本探勘中資料自動分群技術，當使用者進行檢索的同時，直接從美國國家醫學圖書館之 PubMed 資料庫擷取書目、摘要與主題標目資料，即時進行歸納分群，以視覺化的方式呈現知識架構，並讓使用者依自己的相關判斷點選知識架構中的關鍵字，以調整檢索結果。本文首先簡述國外自動分群與視覺化檢索引擎之特色，之後介紹研究團隊開發出來的雛型系統與第一次小型使用者評估的結果。開發出來的檢索系統，將是一個「視覺輔助相關回饋」之生物醫學資訊檢索系統，預定命名為 **Visual Medline**，並期望未來進一步產學合作，發展成為商品化的系統。

關鍵字：資訊檢索、視覺化、知識架構(Ontology)、生物醫學資訊、Medline

壹、前言

臺灣的醫學圖書館大致可分為研究型、學術型和應用型等三類醫學圖書館：研究型以中研院生命科學圖書館與國家衛生研究院圖書館為代表，學術型以醫學校院圖書館為主，應用型指一般醫療院所之圖書館室為主，其又可依據醫院評鑑等級細分為醫學中心、區域醫院、地區醫院三級及精神專科醫院，其中教學醫院規定要設立圖書館，因此保守估計臺灣約有 300 多所的醫學圖書館。而生命科學家、健康科學相關科系師生、醫學研究人員與臨床人員是醫學圖書館的主要使用者，他們的工作攸關人類的健康福祉與生命，因此醫學圖書館員所提供的資訊服務之品質是相當重要的。

美國國家醫學圖書館（National Library of Medicine，簡稱 NLM）的生物醫學書目資料庫 PubMed，收錄全世界 1960 年代以來至今的核心生物醫學論文書目超過 2 千萬筆，且每天都增加 2000-4000 筆的書目。然而臺灣的醫學圖書館館員大多不具醫學相關學科背景，雖然熟知醫學相關資料庫的檢索技巧，並在服務的各大醫學院校與醫療機構努力對使用者進行資料庫檢索的教育訓練，但面對以巨量成長的醫學知識之組織、整理與傳佈，卻有著對「知識內容」無法掌握的無力感。

PubMed 資料庫預設呈現檢索結果的方式，是依使用者檢到之文獻書目以出版年代之新穎性來排序，但我們也都知道最新的文獻不見得是品質最高、最符合使用者需求的文獻，因此使用者面對數量龐大的搜尋結果時，反而很需要依「主題」或「概念」來篩選相關資料；也就是說，面對日益龐大且複雜的數位資料，傳統呈現搜尋結果的方式已不符使用者的需求。如何讓使用者不用浪費時間去瀏覽數十、數百頁的搜尋結果，而讓他們可以依主題來呈現搜尋結果已成為未來資訊檢索系統的開發趨勢。

研究者期望能應用資訊技術，特別是人工智慧之文本探勘技術(text mining)，主動發掘文獻中隱藏的知識，藉以產生文件資料的自動分群(clustering)，並藉由分類結構進一步自動產生同質性資料中的知識架構(ontology)。因此，研究者將以文本探勘中資料自動分群技術，直接從 PubMed 資料庫擷取書目資料，建置一個「視覺輔助」、「相關回饋」的生物醫學資訊文獻檢索系統，該雛型系統在使用者檢索時即時進行歸納分群，以視覺化的方式呈現知識架構，並讓使用者依自己的相關判斷點選知識架構中的關鍵字來調整檢索結果；此外，亦將進行小規模的試用與評估，蒐集使用者對此雛型系統的回饋意見，做為未來在功能與介面上改善的參考。

貳、國外自動分群與視覺化檢索引擎

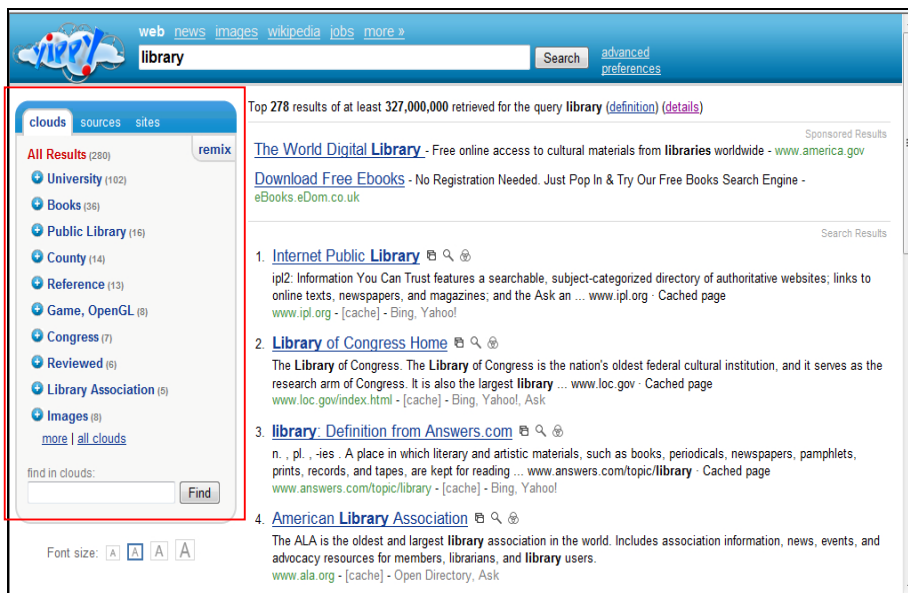
一、Vivisimo 檢索引擎

Vivisimo（網址：<http://vivisimo.com/>）成立於西元 2000 年，總公司設立在匹茲堡，是由三位卡內基美隆大學科學家設計出來的搜尋檢索引擎。當時有鑑於大家在網路上檢索時都會遭遇到資訊超載的現象，於是思考出有別於一般檢索引

擎的檢索機制。一般的檢索引擎只是著重在將搜尋結果排序，而 Vivisimo 最大的特色就是將搜尋結果依照主題類別自動分群，提供一個讓使用者可以一目了然、更直覺與方便的資訊取得與發現的資訊優化環境。^{1、2} Vivisimo 是個西班牙文，它的意思是活潑、明亮、聰明，其發明者取這個名字是想表達他們企圖優化資訊檢索的願景，讓繁複而費時的檢索經驗變得更有效率，讓瀏覽龐大而無意義的數據以及未經組織過的檢索結果變得更活化、更有意義。³

二、Clusty 和 Yippy 檢索引擎

Clusty 是 Vivisimo 在 2004 年推出的免費的網路搜尋服務，它是一個以 metasearch 與 cluster 為概念的搜尋機制，能夠分析關鍵詞搜尋結果，並進行分類的搜尋引擎，可以節省使用者比對結果與進行確認的時間。2008 年開始，Vivisimo 終止對 Clusty 檢索工具的更新服務。⁴ 2010 年 5 月 14 日，Clusty 被設於美國佛羅里達州的雲端搜尋引擎公司 Yippy Inc. 所收購⁵，目前網站名稱改為 Yippy，網址則不變，仍然是 <http://clusty.com>。Yippy 的搜尋結果出現在畫面的主畫面，而左邊的視窗則呈現出檢索結果的叢聚類別項目，它提供三種不同的分群方式，其中雲端分群 (clouds) 將搜尋結果依照主題分群；若以關鍵字“Library”為例，它分出了 35 個分群，包括“University”、“book”、“Public Library”等，點選該分群類別文字後，主畫面就會展出所有該類別的結果。若點選視窗右上的 Remix 按鍵，Clusty 則會將搜尋結果打散重新分類。（見圖 1）



¹ Vivisimo. About Us. Retrieved June 28, 2010, from: <http://vivisimo.com/about/overview.html>

² Vivisimo. Wiki. Retrieved June 28, 2010, from: <http://zh.wikipedia.org/zh-tw/Vivisimo>

³ Vivisimo. What is Vivisimo? Retrieved June 28, 2010, from: <http://vivisimo.com/about/overview.html>

⁴ Clusty. Vivisimo Inc. Retrieved June 28, 2010, from: http://www.worldlingo.com/ma/enwiki/zh_tw/Viv%C3%ADsimo

⁵ Clusty. Wikipedia. Retrieved 2010/6/28, from: <http://en.wikipedia.org/wiki/Clusty>

圖 1: Yippy 檢索結果自動分群畫面
(<http://clusty.com>，上網日期：2010 年 6 月 28 日)

三、Grokker 檢索引擎

Grokker 是個資料分群 (data clustering) 的檢索工具，其服務對象主要為企業資料用戶，是個為關鍵字或領域知識 (domain knowledge) 分類的建構工具，可用於分析與整合其它知識種類與資料庫，對於網際網路資料的瞭解與判讀也可扮演關鍵角色，但 Grokker 已於 2009 年 3 月終止服務。Grokker 檢索結果呈現方式有兩種，分別是「條列式 (outline view)」和「地圖式 (map view)」。圖 2 是條列式的呈現方式，若以 "electronic paper" 為關鍵字瀏覽其檢索結果，共計出現 240 筆資料，其中網頁中間呈現自動分群的 analysis 結果。而圖 3 是地圖式的呈現方式，使用者端必須安裝 Java 圖形瀏覽功能，此種形式可以將檢索結果依照主題分類，而且還能將這些主題視覺化，用不同的顏色代表不同的主題，讓使用者非常容易辨識也省去了讓使用者一頁一頁往下翻的尋找搜尋。地圖模式的優點為可以追蹤資料分類，與找出不同層級之分類關係，但其缺點則為類別與類別之間關係不易區分。

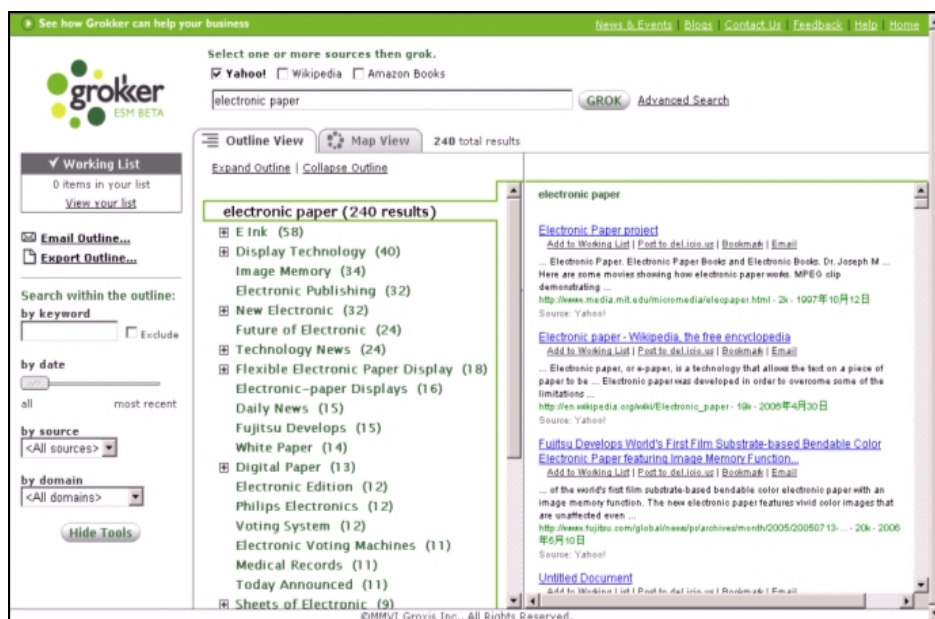


圖 2: Grokker 檢索結果條列式呈現畫面
資料來源: 財團法人國家實驗研究院科技政策研究與資訊中心
(http://cdnet.stpi.org.tw/techroom/analysis/pat_B039.htm)

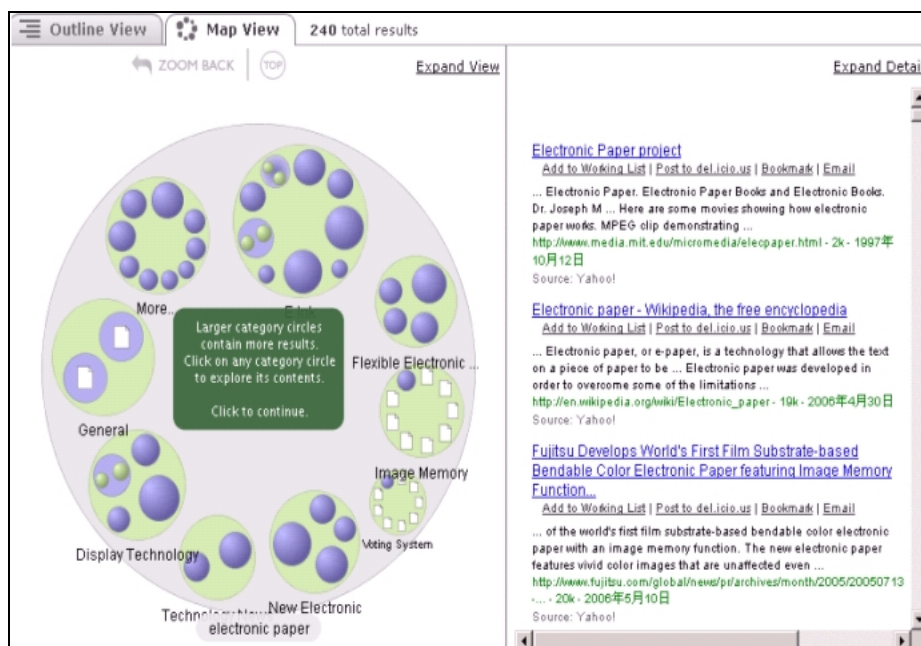


圖 3：Grokker 檢索結果地圖式呈現畫面

資料來源：財團法人國家實驗研究院科技政策研究與資訊中心
 (http://cdnet.stpi.org.tw/techroom/analysis/pat_B039.htm)

四、Carrot² 檢索引擎

Carrot² (網址：<http://search.carrot2.org/stable/search>) 是個叢聚分群的檢索引擎，其資料來源主要來自於 meta-searching。⁶它可以自動重整不同的文件集，然後按照語意法來分類。Carrot² 是一支開程式碼 (open source) 的程式，它提供現成的組件，可以直接透過 yahooAPI、GoogleAPI、MSN LiveAPI、eTools、Meta Search、Lucene、SOLR、Google Desktop 等取得檢索結果。Carrot² 的使用者介面是 Java 語言，但是非使用 Java 的平台 (如 PHP 或 Ruby)，也可以從其 REST 介面呼叫 Carrot² 的分群功能。⁷Carrot² 的檢索界面提供了各式搜索引擎的活頁頁面，選到哪一個搜索引擎的頁面就表示使用者只想從單一搜索引擎當中找資料，例如像 MSN、Yahoo、Wiki、Pubmed 等；另外還有像圖像檢索、新聞檢索、工作檢索、部落格檢索等。

Carrot² 的檢索分類結果有兩種呈現方式，分別是「樹狀式 (tree)」和「視覺式 (visualization)」。圖 4 是樹狀式的呈現方式，是將所有檢索結果依照不同類別分類，然後以相同類別資料最多的排在最上方，每個分類後面的數字就是結果數量。而圖 5 是視覺式的呈現方式，其是一種圓形圖的分群呈現，檢索結果將類別依照數量多寡排序，並以不同的顏色顯示，點選任何一個類別，畫面右邊的

⁶ Carrot2。叢集式平台的搜尋系統專案。上網日期：2010 年 6 月 28 日。檢自：
<http://www.ctimes.com.tw/Download/Show2.asp?O=200609181505034437&F=META>

⁷ Carrot2. About Carrot2. Retrieved June 28, 2010, from: <http://project.carrot2.org/>

Science.gov (網址：<http://www.science.gov/>) 是入口網站同時也是個檢索引擎，它使用的是聯邦檢索 (federated search) 技術，提供美國政府科學方面的資訊以及研究結果。目前是第五版，它可同時檢索超過 42 個資料庫，包括 14 個聯邦技術機構和 2 億個科學資訊；同時它還是 1,900 個以上的科學網站的入口。目前 science.gov 是美國提供國際科學檢索(WorldWideScience)的入口網站。⁸在檢索結果呈現方面，Science.gov 是呈樹狀式結構，依照所得到的結果相同關鍵字最多的排序。畫面中間則是得到的結果，依照檢索詞相似度排序。畫面的右邊還有即時從 wikipedia 得到的檢索詞定義，以及從 EurekAlert 即時取得相關新聞資訊。(見圖 6)



圖 6： Science.gov 檢索結果自動分群畫面
(<http://www.science.gov/>，上網日期：2010 年 6 月 28 日)

參、雛型系統之建置

本研究希望建置以「主題」呈現檢索結果的雛型系統，一般來說，可以採用「分類架構 (Taxonomy)」、「自動文本歸類 (Automated Text Categorization)」或「分群 (Clustering)」等方法來處理。從表 1 的比較可以得知，「分群」技術完全以自動化的方式定義分類架構及進行資料歸類，可自動即時產生分類架構，且建置的時間與費用成本最低，可說是組織、整理快速且巨量成長的醫學知識之利器。

表 1：以主題呈現資料之三種方法比較表

方法	Taxonomy	Automated Text	Clustering
----	----------	----------------	------------

⁸ Science.gov. Wiki. Retrieved June 28, 2010, from: <http://en.wikipedia.org/wiki/Science.gov>

項目		Categorization	
定義	以人工方法定義分類架構，並以人工方式進行資料歸類	以人工方法定義分類架構，並以 自動化 方式進行資料歸類	完全以自動化的方式定義分類架構及進行資料歸類
分類架構	有，事先定義好的	有，事先定義好的	無
分類群組命名流程	<ol style="list-style-type: none"> 1. 建立控制字彙詞表(索引典) 2. 將新的文獻依此索引 3. 定期更新控制字彙詞表 4. 檢索時，再將檢索結果分入至最適的類 	<ol style="list-style-type: none"> 1. 建立控制字彙詞表(索引典) 2. 將新的文獻依此索引 3. 定期更新控制字彙詞表 4. 檢索時，再將檢索結果分入至最適的類 	依被每次被檢索的資料內容之不同而自動產生
分類架構定義方式	人工	人工	自動即時
資料歸類方式	人工	自動	自動即時
建置費用	最高	高	低
建置時間	5~10年	2~5年	1週至1個月
未來新詞彙(如:禽流感)	無法滿足，因原有分類架構沒有新詞彙	無法滿足，因原有分類架構沒有新詞彙	可滿足，因為會隨新增資料內新詞彙而改變分群架構
分類架構維護費用	需要	需要	不需要
複雜度	最高	高	極低
整體總費用	最高	中	最低

資料來源：本研究整理

本研究的雛型系統由具資工與醫工背景的研究者負責開發，其與他先前所屬研究團隊已建置過生物醫學相關的「資料探勘系統」(Chen, Chiang, Li, Lin, Wu, & Liu, 2002; Chiang & Lin, 2001; Chiang & Hsu, 2002; Chiang, Shieh, Hsu, & Wong, 2002; Chiang, Shieh, Hsu, & Wong, 2005)、「知識建構與管理雛型系統」(Chiang, 2007; Chiang, Lin, & Hsu, 2005; Chiang, Lin, & Liu, 2005a; Chiang, Lin, & Liu, 2005b; Lin & Chiang, 2005a; Lin & Chiang, 2005b; Lin & Chiang, 2005c; Lin & Chiang, 2006; Lin, Chiang, & Hu, 2005; Liu, Chiang, Wong, Tsai, & Lin, 2006; Yen, Yang, Chiang, Wong, Wu, & Liu, 2005)以及「實證醫學臨床指引平台」(Chiang,

2005a-c)。研究團隊將本次的雛型系統暱稱為 Visual Med，主要功能是在使用者輸入檢索關鍵字的同時，系統會由 NLM 的 PubMed 生物醫學核心論文資料庫中擷取書目、摘要與主題標目(Medical Subject Headings，簡稱 MeSH)資料，立即進行後續自動分群與視覺化呈現的處理。從文字資料中摘錄重點，並區分語意上之分類是一件不容易的工作，幸好科學性(醫學)文獻之書目有一定的架構與組織，研究者可運用相關技術產生 Metadata 之資料，以下簡述執行步驟：

一、文獻結構處理 -

利用 Hidden Markov Model (HMM)或 Parser Tree 訓練並建立電腦機器學習之 Metadata 分群(segmentation) 格式，以段落出各不同單元的 Meta 區域出來，藉由不同之 Meta 區域摘錄出其中之關鍵要素。

二、索引關聯分析 -

從文獻的分類中所常出現的字詞，及其間之關聯，評斷此類醫學文獻之重要概念，以作為概念之資料搜索與查詢，協助使用者可正確從概念中歸納相關知識，如” Network” 這個字，在電子電機與資訊領域代表與電腦網路相關之知識，而於生物醫學上就代表了神經網路等，其中所運用的 terms 就完全的有差異。

三、產生知識架構 -

從同類文章中借助統計分析與文本探勘方法，主動發掘其中之內涵與詞彙間的關聯與相依現象，利用統計以求得結果，再藉助專家協助，判斷其間從屬性質(如”is a”、”part of” ...)的正確與否，進而修正以產生專業領域的特定 ontology。

四、視覺化知識呈現 -

將發掘到的知識視覺化呈現，俾利使用者迅速掌握知識的內涵。有效的可視覺化介面可以協助使用者能夠觀察、操縱、研究、瀏覽、探索、過濾、發現、理解大規模資料，從而可以極其有效地發現隱藏在資訊內部的特徵和規律，協助使用者能清楚了解其中內藏涵義。

Visual Med 的使用者界面採 Google 化查詢框之設計，當使用者輸入檢索關鍵字時，系統會即時撈取 PubMed 資料庫中近一年的相關文獻，並將檢索結果的前 500 筆呈現在主畫面；右邊的視窗呈現即時產生的知識架構，使用者若點選其中的關鍵字，系統會自動將檢索結果與此關鍵字的檢索結果再做交集；左邊的視窗則是標籤雲(tag cloud)，以字體的大小表現此頁面所出現的關鍵字之頻率，提供使用者視覺上的線索。(見圖 7)

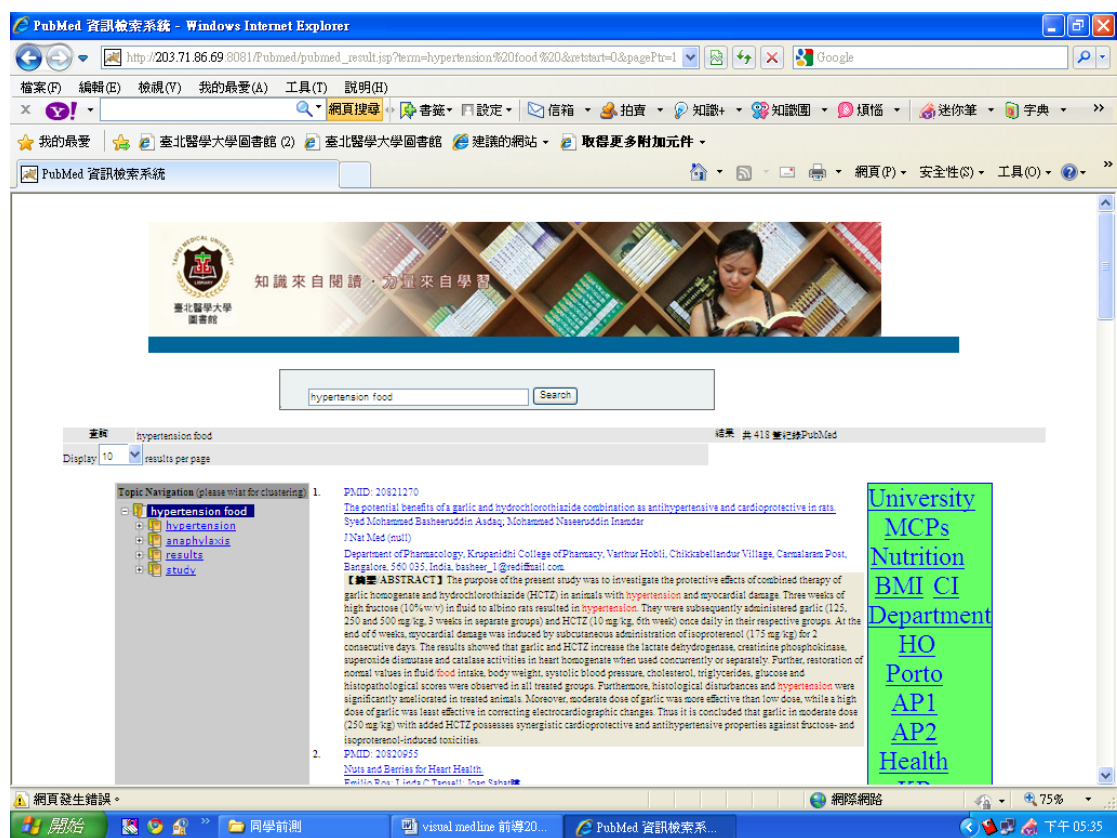


圖 7：Visual Med 雛型系統檢索結果呈現畫面

肆、雛型系統之使用者評估

此一雛型系統於 2010 年 5 月初步建置完成，之後交由具圖書資訊學背景的研究者來執行使用者試用與評估，以蒐集回饋資料。因為醫學相關科系低年級學生是醫學領域使用者對該學科內容最生疏的使用者群，最需要「智慧型」的系統來協助，因此研究者鎖定由他們試用此雛型系統。

研究者任教於臺北醫學大學通識教育中心，每年固定對大一、大二學生開設 2 學分選修的「健康科學資訊素養」通識課，該課程以講課、上機練習與作業的方式使學生熟悉各式健康科學領域之實體與數位資源，進而學習檢索、取得、評鑑、與利用該類資源之技巧，而檢索 PubMed 系統並評論之就是作業之一。研究者於 98 學年度第 2 學期開放本雛型系統給修課的 54 位同學使用，要求同學以相同的檢索主題與限制條件分別檢索這兩個系統，除以線上問卷調查的方式蒐集同學在操作功能、檢索結果呈現方面的回饋意見外，亦請同學對檢索結果之精確度（precision rate）進行評估。

在調查期間有 49 位同學上線填寫了問卷（見附錄 1），回收率為 90.7%。總體來說，在五分的量表中，受試者對 PubMed 的喜好程度平均數是 3.81，而 Visual Med 的平均數是 3.40；在檢索結果品質的滿意度方面，PubMed 的平均數是

3.85，而 Visual Med 的平均數是 3.53；當問及是否會推薦 Visual Med 給他人使用時，有 80.9% 的受測者回答「會」的選項。顯見 Visual Med 雖仍在開發階段，但測試者對其喜好程度與檢索結果品質多持正面的看法，對其發展也是樂觀其成。

在系統表現方面，因為 Visual Med 的設計是只撈取 PubMed 近一年的文獻，而 PubMed 則是收錄 1960 年至今的文獻，但其檢索結果是以出版年代由近到遠來排序，因此研究者請同學只初步對檢索結果的前 10 筆進行相關判斷，研究者之後計算同學回報的精確度(precision)之平均值，PubMed 為 61%，而 Visual Med 為 52%。受試者僅審視了兩個系統的各 10 筆書目資料與摘要，不符需求者就有 4-5 成，這個結果令研究者頗感意外，到底是因為受試學生的檢索技巧不夠成熟（參見附錄 2），導致檢索到的文獻有些與主題不符？或是因為「主題相關」與「情境相關」之間存有很大差距？都是將來可以進一步細究的議題。

問卷調查中也詢問受試者對 PubMed 資料庫各項檢索功能的熟悉程度，1 到 5 的選項分別表示從「非常不熟悉」到「非常熟悉」，以下依序呈現他們對各個項目的熟悉程度：

1. 非主題的條件限制(limit) -- 4.26
2. 自然語言關鍵字檢索(keyword) -- 4.06
3. 布林邏輯(AND、OR、NOT) -- 3.85
4. 控制字彙醫學標題檢索(Mesh) -- 3.62
5. 切截(truncation) -- 2.85
6. Clinical Query -- 2.64

而在 Visual Med 方面，由於是新開發的雛型系統，而且開發的目標即是希望使用者可以用最少的學習、以直覺性的方式來操作，因此我們詢問的是受試者對這些功能的喜好程度。研究發現這些對生物醫學學科知識較生疏的低年級學生，對 Visual Med 的各項功能中，喜愛程度則依次為：「Google 化的查詢框」（平均數是 3.96）、「視覺化知識架構之呈現」（3.68）、「標籤雲」（3.60）。

為進一步了解試用者的使用心得與建議，線上問卷調查最後一題以開放式的問題來蒐集資料。研究者在課堂上鼓勵同學發表意見，並說明他們的意見將成為日後改進此一雛型系統的重要參考。茲將試用者的回饋文字依主題整理如下：

一、正面回饋

(一) 操作簡單、直覺化

受調者普遍對於 VisualMed 這個雛型系統 Google 化的使用介面感到滿意，反應在文字方面的描述包括：「方便」、「簡單」、「容易」、「易適應」、「直接」、「平民」、「易上手」、「人性化」等。以下摘錄試用者的回饋文字。

「…相較之下 VisualMed 操作較為方便。」

「使用方法很簡易，很容易適應，雖然資料精準度與品質無法像 OvidMedline 和 PubMed 那麼高，但也還算不錯，可以接受，所以還是會推薦學弟妹做初步學習的使用。」

「VisulMed 在使用上有更合大眾口味的感覺，簡單容易不需多做說明。」

「VisualMed 的 Google 化查詢框，較 PubMed 的介面方便、直接。」

「整體來說不失為一個很平民且易於操作的系統。…相信對於不會使用 PubMed 的使用者來說，這極為簡單的系統很值得推薦的!」

「介面簡單很多很多，而且查詢出來的資料與 PubMed 不會相差很多。」

「簡單明瞭，很清楚的就可以知道自己該如何查詢，就算之前沒用過也會使用，像 PubMed 就沒辦法馬上上手。」

「對於平時常常使用 Google 搜尋的我來說，滿習慣使用『Google 化查詢框』的。」

「VisulMed 讓我覺得很人性化。」

「VisualMed 的操作方式較簡單，可如同以 Google 的搜尋方式一樣，把想打的關鍵字打上去即可，這樣的操作方式對於專業領域的人而言很方便。」

「… 我覺得可以發展一個較簡單操作的檢索介面是很令人興奮的事情，因為這樣可以讓知識更容易被廣為利用，我也很開心能參與這樣的計畫。」

「版面雖然沒有經過美化，但是明瞭簡要的介面，讓人使用起來很舒服，且讓非專科的一般社會人士使用上容易上手。」

「VisualMed 查詢的方式比較簡單，只需要把各種有相關的關鍵字都打進去就可以。」

(二) 關鍵字反紅標示

受調者對於 VisualMed 這個雛型系統在檢索結果中將使用者輸入之檢索詞反紅標示的功能大為讚賞，認為此舉有助於更快掌握關鍵字在該文獻中的脈絡，可協助他們做相關判斷。以下摘錄試用者的回饋文字。

「〔顯現出〕輸入的關鍵字在摘要中出現的位置，所以看到摘要中紅字的上下文，可能就會提到關於我們想探討的問題，因此我們可以更快決定，這篇論文是否為我們需要的，…VisulMed 將摘要中的關鍵字表現出來的功能很棒。」

「VisualMed 會把我打上去的關鍵字，以紅字出現在摘要中，而 PubMed 不會，這是 VisualMed 的優點，可以讓我比較快抓到整篇文章中我需要的段落。」

「VisualMed 有幾個貼心的設計：例如像一般搜尋引擎一樣，搜尋的關鍵字會顯示紅字。」

「用紅色標示關鍵字讓人一眼就可看到，不用再仔細找它。」

「以紅色標出檢索條件中的用詞，也可以使使用者能更明確的檢索到所需要的資料。」

(三) 標籤雲和知識架構

受調者對於 VisualMed 這個雛型系統左側畫面的知識架構與右側畫面的標籤雲等視覺輔助功能也持肯定的看法，認為自動分群之詞對與關鍵字出現頻率等資訊，有助於他們抓到相關醫學知識的重點，進而決定下一個檢索步驟。以下摘錄試用者的回饋文字。

「視覺化知識架構呈現和標籤雲皆清楚呈現其一層層架構及出現頻率高低之主題，讓我們能更容易使用，是個不錯的檢索工具。」

「標籤雲的功能，讓使用者很容易明白其相關性。」

「VisulMed 有樹狀架構，點進〔去〕可以看到明顯的關鍵字以及其摘要，非常簡厄的讓人找到想找的資料，我覺得十分不錯。」

「我覺得有標籤雲確實能縮短我在找尋相關字詞的瀏覽時間，而且

一眼就可明瞭哪些字在文獻中出現較多次。」

「在搜尋結果的旁邊就有 Topic Navigation〔指系統左側畫面的知識架構〕，很方便，可以在同一個畫面用+、- 在分類中，範圍很明確的找到所要找的類別。」

「自動分群更吻合一般大眾的需求和直覺式的思考，較不像 PubMed 是以專業學門或用專有名詞來呈現主概念和子概念的上下樹狀結構關係。…或許自動分群是提取常見的同時出現的詞對，更符合實用性和平常習慣性使用的用語。」

「標籤雲對於非專科學門的人士在使用上也有能提示關鍵語的功能。大小代表文章中用詞出現的頻率。…和 PubMed 的控制字彙比較起來，對非專科人士的使用較能上手。較容易抓到重點外也可以透過其他文章中常用用語做其他相關主題的檢索研究。」

(四) 書目之後直接呈現摘要

有些受調者表示喜愛 VisualMed 這個雛型系統將檢索到的書目與其摘要呈現在同一頁面的作法，他們認為如此不需要再按一個連結就可以直接閱讀摘要來篩選文獻，比較省時省力。以下摘錄試用者的回饋文字。

「我覺得 VisulMed 比較好用，因為它會直接顯示摘要。」

「在搜尋結果就直接 show 出摘要，很簡單明瞭。」

「〔檢索結果的呈現〕把摘要、來源及出處直接列於篇名下方，這樣不僅可以迅速了解文章的大要，且不需像 PubMed 尚須另行連結至另一頁面才能看見摘要，使時間不夠的醫護人員能加快篩檢資料的速度。」

二、改進建議

(一) 增加「限制」的功能

此次的試用對象為研究者開授「健康科學資訊素養」通識課的選課同學，經過課堂教學與檢索作業的洗禮，他們對於 PubMed 系統的操作方式已有一定程度的熟悉，問卷調查結果顯示這些受試者對「限制」這個功能的熟悉度居冠（在五分量表中平均分數為 4.26）。研究者因此推論受調者對於 VisualMed 這個雛型系統沒有提供此一功能感到不方便，而成為最大力推薦未來 VisualMed 應

增加的功能。以下摘錄試用者的回饋文字。

「VisualMed 沒有 limit 介面用起來比較簡單，但是會有一些錯誤產生，像輸入 age，卻得到 average。」

「在作查詢『限制方面』的使用時，卻不太方便。…且沒有像 PubMed 可以先查詢到醫學上所常用的專用詞彙再去作查詢，會提高查全率。」

「若能加上限制功能應該更完整。」

「如果限制文章使用語言是日文，但 VisualMed 是直接在內文篇名、摘要和內容做搜尋找交集，而這樣會造成找到文章中出現 Japanese，但實際上卻不是用日文寫成的結果。所以建議可以增加類似 PubMed 的『限制』功能鍵。」

研究團隊討論之後，決議參考 PubMed 「限制」功能中的選項，初步選擇較常使用的項目，包括 Article Types、Evidence-based、Time Frame、Species、Gender、Ages、Languages 等層面（見附錄 3），在 VisualMed 的下一個版本中增加 refine 的功能，讓使用者在輸入關鍵字得到第一次的檢索結果後，再依需求從非主題條件再次限縮檢索結果。

(二) 調整知識架構的呈現方式

有一位試用者對 VisualMed 雛型系統左側畫面知識架構的呈現方式提出看法，他寫到：「視覺化知識架構將資料自動分群，可以從中知道其它相關的研究，但起初使用時點開成樹狀圖，我覺得有點容易被不同層次中相同的詞搞混，但多點幾次就比較瞭解了。」

研究團隊認為這個意見很有價值，樹狀的知識架構可能會誤導使用者這些詞對有從上下位的關係，因此決議在 VisualMed 的下一個版本，將改以網絡狀的方式呈現。

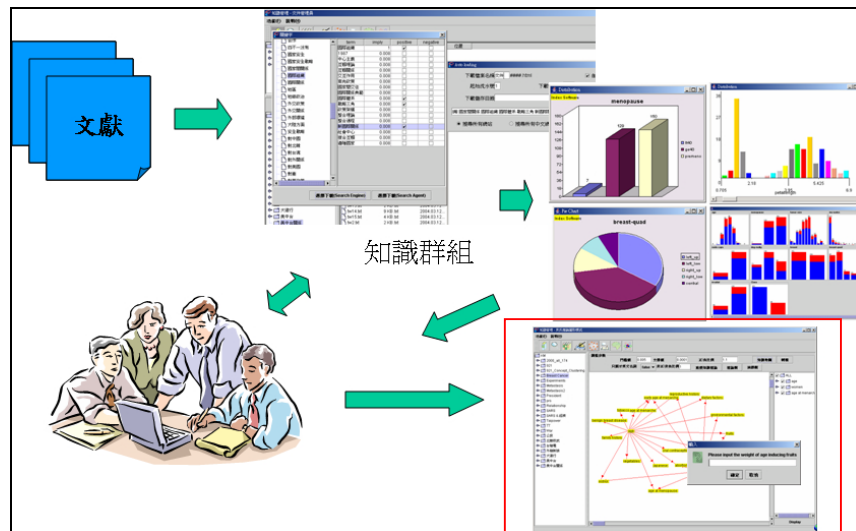


圖 8：視覺化知識呈現示意圖 - 網絡狀知識架構

伍、未來之規劃

一、進行大規模使用者測試

研究團隊規劃參考此次小型使用者評估之回饋意見，對此系統之功能進行改善，預計 2011 年上半年將初步改善後的系統，掛在本校圖書館的網頁上，透過宣導邀請本校所有系所教師與學生、以及醫療體系三個附屬醫院的醫事人員上網使用，並請他們在使用後填寫線上滿意度調查問卷，以大規模蒐集各類型使用者對本系統的建議。再次改善後的系統，將邀請使用者參與一對一的訪談，並以“think aloud”的方式實際了解使用者在操作本系統時的行為、問題與思考模式，再詢問其建議，做為再次改進的參考。

二、使用者介面開發與美化

2011 年下半年，研究團隊將依據前述使用者測試與回饋的資料，設計好用、易懂，且可以商品化的使用者界面。之後，研究團隊亦規劃邀請不同身份之使用者（教師、大學生、研究生、醫事人員）組成四個焦點團體，分別了解不同使用者群對此系統與介面的使用經驗與期望，並做最後的修正與調整。

三、產學合作，發展成商品

本研究開發出來的雛型系統，是一個「視覺輔助相關回饋」之生物醫學資訊檢索系統，因此預定命名為 Visual Medline。本系統之重點在於將巨量的生物醫學資訊自動地轉化成 ontology，以形成有組織的規則知識，進而協助生醫領域專業人員快速地建構自己本身所需的知識；而視覺化的知識呈現方式，更利於使用者迅速掌握知識的內涵，獲得更多線索進行相關判斷。研究團隊期望未來可以尋求和生醫相關的資料庫廠商進行產學合作，將此雛型系統發展成商品，進行市場之推廣。（誌謝：本系統獲國科會產學合作計畫 NSC 99-2622-H-038 -001 -CC3 支持，執行日期為民國 99 年 11 -100 年 10 月，特此誌謝）

【參考文獻】

- Chen, R.J., Chiang, I.J., Li, Y.C., Lin, Y.M., Wu, H.C., Liu, C.H. (2002). Utilizing Decision Tree Classifier with Optimal ROC Cut-off Values in the Analysis of Determining Factors for Nonoperative Failure Patients after Major Blunt Hepatic Trauma. In Workshop of the foundation of data mining and discovery in the 2002 IEEE International Conference on Data Mining (ICDM), Maebashi, Japan.
- Chiang, I. J. (2005a). A Clustering Search through Medical Evidence Sources for CPG creation. In Asian Pacific Conference on BM-CPG, Taipei, Taiwan.
- Chiang, I. J. (2005b). An Idea of Interactive Web-based Evidence-Based Clinical Practice Guideline System. In Asian Pacific Conference on BM-CPG, Taipei, Taiwan.
- Chiang, I.J. (2005c). SECIG: A Visualized Shareable Evidence-Based Computer-Interpretable Guideline System. In Asian Pacific Conference on BM-CPG, Taipei, Taiwan.
- Chiang, I.J.(2007). Discover the Semantic Topology in High-Dimensional Data. *Expert Systems with Applications*, 33 (1), 256-262.
- Chiang, I.J., Hsu, J. Y.J., (2002). Fuzzy Classification Trees for Data Analysis. *Fuzzy Sets and Systems*, 103(1), 87-99.
- Chiang, I.J., Lin T.Y., Liu, Y. (2005b). Topological Relational Tables. *Proceedings of The Tenth International Conference on Rough Sets, Fuzzy Sets, Data Mining, RSFDGrC(1)*, Regina, Canada. 728-737.
- Chiang, I.J., Lin, T.Y. (2001). Index Miner: a data mining system. *Proceedings of the Annual International Computer Software and Applications Conference (COMPSAC)*, Chicago, IL. 613-614.
- Chiang, I.J., Lin, T.Y., Hsu, J.Y.J. (2005). Generating Hypergraph of Term Associations for Automatic Document Concept Clustering. *Proceedings of the 8th IASTED International Conference on Artificial Intelligence and Soft Computing*, Marbella, Spain. 182-186.
- Chiang, I.J., Lin, T.Y., Liu, Y. (2005a). Table Representations of Granulations Revisited Pre-topological Information Tables. *Lecture Notes in Computer Science*, 3641, 728-737.
- Chiang, I.J., Shieh, M.J., Hsu, J. Y.J., Wong, J. M. (2002). Apply Fuzzy Classification to Colon Polyp Screening. *Communications of IICM*, 5 (2), 7-12.

- Chiang, I.J., Shieh, M.J., Hsu, J. Y.J., Wong, J. M. (2005). Building a Medical Decision Support System for Colon Polyp Screening by Using Fuzzy Classification Trees. *Applied Intelligence*, 22, 61-75.
- Lin, T. Y., Chiang, I.J., Hu, X. (2005). Latent Semantic Space for Web Clustering. *Proceedings of 2005 IEEE International Conference on Granular Computing*.
- Lin, T.Y., Chiang, I.J. (2005a). Latent Semantic Space for Web Clustering. *Proceedings of the Tenth International Conference on Rough Sets, Fuzzy Sets, Data Mining*. Regina, Canada.
- Lin, T.Y., Chiang, I.J. (2005b). Semantic Based Clustering of Web Documents, *Proceeding of the 2005 IEEE International Conference on Granular Computing*, Beijing, China. 189-192.
- Lin, T.Y., Chiang, I.J. (2005c). A Simplicial Complex, a Hypergraph, Structure in the latent Semantic Space of Document Clustering. *International Journal of Approximate Reasoning*, 40, 55-80.
- Lin, T.Y., Chiang, I.J. (2006). Formal Concept Analysis and Document Clustering via Granular Computing. *Proceedings of the 2006 IEEE International Conference on System, Man, and Cybernetics*, Taipei, Taiwan.
- Liu, C.C.H., Chiang, I. J., Wong, G.J.M., Tsai, H.C., Lin, H.C. (2006). A Complex Bio-Networks of the Function Profile of Genes. *LNCS Transactions on Computational Systems Biology V*, 4070, 9-24.
- Yen, Y. H., Yang, Y. C., Chiang, I.J., Wong, J.M., Wu, S.J.H., Liu, C.C.H. (2005). Exploring the Knowledge Structure of “Breast Reconstruction” – Evidence-based Plastic Surgery by Online Cluster Analysis of Literature. *Journal of Plastic Surgical Association*, 14(4), 303-315.

【附錄 1】 線上問卷調查題目

一、基本資料

姓 名 _____

學 號 _____

二、對 PubMed 的看法

I. PubMed 的使用狀況

請對以下問題圈選出您個人的同意程度

題 目	非常不同意	→	非常同意		
1. 我喜歡使用 PudMed	1	2	3	4	5
2. 我對 PudMed 檢索結果品質感到滿意	1	2	3	4	5

II. 對 PubMed 檢索介面的使用習慣

請圈選對各項使用介面的使用經驗

題 目	非常不熟悉	→	非常熟悉		
1. 自然語言關鍵字 (Nature Language)	1	2	3	4	5
2. 醫學標題 (MeSH)	1	2	3	4	5
3. Clinical Query	1	2	3	4	5
4. 限制 (Limit)	1	2	3	4	5
5. 布林邏輯 (And / Or / Not)	1	2	3	4	5
6. 切截 (Truncate)	1	2	3	4	5

三、對 VisualMed 介面的使用態度與看法

I. 對 VisualMed 的介面使用態度

請對以下問題圈選出您個人的同意程度

題 目	非常不同意	→	非常同意		
1. 我喜歡使用 VisualMed	1	2	3	4	5
2. 我對 VisualMed 檢索結果品質感到滿意	1	2	3	4	5
3. 您會推薦本介面給其他人嗎?	<input type="checkbox"/>	會	<input type="checkbox"/>	不會	
	<input type="checkbox"/>	不一定,原因是:	_____		

II. 對 VisualMed 各項功能的喜好程度

請圈選對以下各項功能的看法

題 目	非常不喜歡	→	非常喜歡		
1. Google 化查詢框	1	2	3	4	5
2. 視覺化知識架構呈現(Ontology)	1	2	3	4	5
3. 標籤雲 (Tag Cloud)	1	2	3	4	5

III. VisualMed 使用心得 (與 OvidMedline 和 PubMed 之比較)

【附錄 2】受試同學檢索策略舉要

Pub Med	Hits	Visual Med	Hits
G6PD deficiency Limits: Humans, English, Publication Date from 2007/01/01 to 2010/12/30	247	G6PD deficiency English man	6
Epilepsy Limits: Chinese, published in last year	27	Epilepsy Chinese	500
“Diabetes Mellitus” [Mesh] Limits: Core clinical journals	33274	Diabetes Mellitus	500
Anemia therapy prevention Limits: Human, Male, English, Middle Aged + Aged 45+ years, published in the last 1 year	42	Anemia male therapy prevention middle age	39
Leukemia therapy Limits: published in the last 1 year	3281	Leukemia therapy	500
Pneumonia Limits: Humans, Female, Clinical Trial, English, 80 and over: 80+ years, published in the last 1 year	45	Pneumonia, female, clinical trial, English, human, over: 80 years,	9
Anemia Limits: published in the last 1 year	4807	Anemia, Chinese, clinical trial	14
(lung neoplasm) And smoking Limits: Middle Aged: 45-64 years, Middle Aged + Aged: 45+ years, published in the last 1 year	236	Lung neoplasm, smoking, middle aged	189
Breast cancer therapy Limits: Humans, Male, English, published in the last 1 year	515	Male breast cancer therapy	500
“Heart Diseases” [Mesh] Limits: Humans, English, All Infant: birth-23 months, published in the last 5 years	6709	Cardiac disease therapy male infant Taiwan Taipei	7
Tourette	9	Tourette, humans, English, 2-5	80

Limits: Humans, English, Preschool Child: 2-5 years, published in the last 1 year		years	
Lung cancer prevention & control Limits: Humans, Female, Middle Aged: 45-64 years, published in the last 1 year	75	Lung cancer middle-aged female prevention control	73
Morbid obesity Limits: Humans, Male, English, Aged 65+ years, published in the last 1 year	70	Morbid obesity English male 65	20
Obesity genetic Limits: Humans, Male, English, Adolescent: 13-18 years, published in the last 1 year	59	Obesity genetic Humans, Male, English, Adolescent	54
Anorexia therapy Limits: Humans, Female, English, Preschool Child: 2-5 years, published in the last 1 year	11	Anorexia therapy Humans, Female, English, Preschool Child	10
SARS treatment Limits: Humans, Adult: 19-44 years, published in the last 1 year	13	SARS treatment for adults	13
Diabetes therapy Limits: Humans, English, All Child: 0-18 years, published in the last 1 year	772	Diabetes therapy child English	433
(lung neoplasm) AND therapy Limits: Humans, Female, All adult: 19+ years, published in the last 1 year	1467	Lung neoplasm therapy	500
Diabetes therapy Limits: Humans, English, Child: 6-12 years, published in the last 1 year	371	Diabetes therapy Humans English Child	418
(ANEMIA) AND (TREATMENT) Limits: published in the last 1 year	2799	ANEMIA TREATMENT LAST 1 YEAR	6

Fatty liver Limits: Humans, Male, English, Middle Aged+ Aged:45+ years, published in the last 1 year	254	Fatty liver liver cancer 45 male English	4
Chinese obesity limits: Female, published in the last 1 year	189	Female Chinese obesity	186
Cholesterol AND atherosclerosis AND hypertension Limits: Humans, Aged: 65+ years, published in the last 1 year	46	Cholesterol, atherosclerosis, hypertension, aged 65, human	6
Alzheimer's disease therapy Limits: Published in the last 1 year	1309	Alzheimer's disease therapy	500
Gout AND drug therapy Limits: Male, All adult: 19+ years, published in the last 1 year	43	Gout AND drug therapy male adult	38
Spinocerebellar degeneration Limits: published in the last 1 year	173	Spinocerebellar degeneration	148
Dementia AND therapy Limits: Humans, Male, English, Young Adult: 19-24 years, Adult: 19-44 years, published in the last 1 year	124	Dementia therapy adult young adult male	41
(Heart Disease) AND treatment Limits: published in the last 1 year	17448	Heart disease treatment last one year	21
((“Carcinoma, Non-Small-Cell Lung” [Mesh]) AND (“Mutation” [Mesh]) AND target Limits: only items with abstracts, Humans, Female, English, Middle Aged: 45-64 years, published in the last 1 year	5	Carcinoma, Non-Small-Cell Lung mutation target female middle aged	6
Smoking AND “Lung disease” Limits: All Adult: 19+years, Young Adult:19-24 years, Adult: 19-44 years, Middle Aged: 45-64	62	Smoking lung diseases middle aged	450

years, Middle Aged+ Aged: 45+ years, published in the last 1 year			
Diabetes AD eye disease Limits: Humans, Female, English, Adult: 19-44 years, published in the last 1 year	146	Diabetes eye disease English female adult	266
(Cervical Cancer) AND vaccine AND treatment Limits: Female, All Adult: 19+ years, published in the last 1 year	109	Cervical Cancer vaccine treatment all adult females	107
Anemia AND therapy Limits: Humans, Male, Young Adult: 19-24 years, published in the last 1 year	190	Anemia therapy Humans male young adult	187
Liver cancer AND treatment Limits: Humans, English, published in the last 1 year	2397	Liver cancer treatment	500
(Parkinson's disease) AND (drug therapy) Limits: Female, English, Adult: 19-44 years, published in the last 1 year	54	Parkinson's disease drug therapy English 1 year female	11
Mammography utilization Limits: Humans, Female, English, Middle Aged: 45-64 years, published in the last 1 year	44	Mammography utilization humans female age 45-64 English	77
Glycogen storage disease Limits: Adolescent: 13-18 years, published in the last 1 year	21	Glycogen storage disease adolescent	20
Stroke hypotension Limits: Male, Middle Aged: 45-64 years, published in the last 1 year	12	Stroke hypotension middle age male	14
(liver neoplasm) AND therapeutics Limits: Humans, Child: 6-12 years, published in the last 1 year	23	Liver neoplasm therapeutics 6-12 child	11
Pneumonia AND therapeutics Limits: All Child: 0-18 years,	182	Pneumonia therapeutics child	107

published I the last 1 year			
Weight gain Limits: Male, Young Adult: 19-24 years, published in the last 1 year	114	Young male adult weight gain	114
AIDS tuberculosis English young adult Limits: published in the last 1 year	47	AIDS tuberculosis English young adult	47
Diabetes AND therapeutics Limits: Female, 80 and over: 80+ years, published in the last 1 year	245	Diabetes therapeutics female over 80 years	139
Hypertension treatment Limits: Humans, Male, Aged: 65+ years, 80 and over: 80+ years, published in the last 1 year NOT English	85	Hypertension treatment male 65	247
Asthma AND medicine Limits: Humans, English, All Child: 0-18 years, published in the last 1 year	360	Asthma medicine 0-18 English	7
(Metabolomics NMR) AND magnetic resonance spectroscopy Limits: Animals, published in the last 1 year	44	Metabolomics animals NMR	46
Congenital Diabetes Therapy Limits: Male, Preschool Child: 2-5 years, published in the last 1 year	5	Congenital Diabetes Therapy male 3	7

【附錄 3】新版 VisualMed 之「限制」功能選項規劃表

Article Types

- Journal articles
- Editorial
- Letter
- Technical report
- Clinical trail

Evidence-based

- Systematic reviews [從原 PubMed 中 subsets 欄位]
- Meta-analysis [從原 PubMed 中 types of articles 欄位]
- RCT [從原 PubMed 中 types of articles 欄位]
- Review [從原 PubMed 中 types of articles 欄位]

Time Frame

- 30days
- 60 days
- 90days
- 180 days
- 1yeas
- 2years
- 3 years
- 5 years
- 10 yearas
- From [yyyy] to [yyyy]

Species

Humans Animals

Gender

Male Female

Ages

All Infant: birth-23 months All Child: 0-18 years All Adult: 19+
years Newborn: birth-1 month Infant: 1-23 months Preschool Child: 2-5
years Child: 6-12 years Adolescent: 13-18 years Young Adult: 19-24
years Adult: 19-44 years Middle Aged: 45-64 years Middle Aged + Aged:
45+ years Aged: 65+ years 80 and over: 80+ years

Languages

English French German Italian Japanese Russian Spanish