

控制詞彙與非控制詞彙 主題索引問題之探討

Problems of Controlled Vocabulary versus Uncontrolled Vocabulary in Subject Indexing

陳昭珍*

Chao-chen Chen

摘 要

資料庫的設計，應該使得資料庫中的資料，能達到最佳化的檢索效果，而要達到最佳的效果，就應有好的索引及檢索模式，索引與檢索可說是一體之兩面。遺憾的是，最佳的索引與檢索模式並不易獲得，致使實務工作只能在理想與現實之間擺盪，其中最根本的問題還是在於主題語詞的語意問題相當多元，作者、索引者與檢索者三者所用詞彙不一定相同。所以是否該用控制詞彙掌握語意，或應節省成本並呈現作者所用的語詞而用自然語言來表達本文之主題，至今仍無法定論。不過，資訊界對此問題的追求卻是契而不捨的，由於他們的努力，使我們能對這個問題有更深入的了解。本文主要以文獻分析法，說明近百年來資訊檢索界對於控制詞彙與非控制詞彙之爭辯的相關研究、中文資訊檢索相關研究及問題，並分析影響資訊檢索效能之相關因素有那些，進而提出幾點建議供國內相關單位參考。

關鍵字：主題索引、控制詞彙、非控制詞彙、資訊檢索

*國立台灣師範大學社教系副教授

Abstract

This article is mainly to survey the centennial controversy between controlled vocabulary v. uncontrolled vocabulary of subject indexing in the western library and information society. We also discuss the related problems in Chinese information retrieval systems and analyze the factors affecting their performance.

Keyword: Subject indexing, Controlled vocabulary, Uncontrolled vocabulary, Information retrieval

壹、前言

資料庫的設計，應該使得資料庫中的資料，能達到最佳化的檢索效果，而要達到最佳的效果，就應有好的索引及檢索模式，索引與檢索可說是一體之兩面。遺憾的是，最佳的索引與檢索模式並不易獲得，致使實務工作只能在理想與現實之間擺盪，其中最根本的問題還是在於主題語詞的語意上，所以是否該用控制詞彙掌握語意，或應節省成本並呈現作者所用的語詞而用自然語言，至今仍無法定論。不過，資訊界對此問題的追求卻是契而不捨的，由於他們的努力，使我們能對這個問題有更深入的了解。本文主要以文獻分析法，說明近百年來資訊檢索界對於控制詞彙與非控制詞彙之爭辯的相關研究、中文資訊檢索相關研究及問題，並分析影響資訊檢索效能之相關因素有那些，進而提出幾點建議供國內相關單位參考。

貳、自然語言與控制詞彙爭論之相關研究

索引語言乃指檢索者用來做為檢索條件的標籤，如作者、主題、期刊名稱等，都是使用者會使用到的檢索點。本文主要將討論重點放在「主題」的索引與檢索上。主題索引語言到底應該使用控制詞彙還是自然語言，那一種索引模式效能較好，這個爭論從十九世紀就已開啓，至今仍無定論，羅理(Jennifer Rowley)將這長達一世紀的爭論分為四個時期，(註1)以下即依羅理之分期方式，分別針對這四個時期主要的研究及論點做一介紹：

一、第一期研究

十九世紀主要的控制詞彙是分類表，當時的目錄乃是分類式的目錄，資料不

多時，分類式目錄還算好用，但當資料量越來越多時，使用者往往無法在一分類式目錄中找到某一特定的書。再者，目錄中所用的類目也不夠具親和性，所以開始有題名索引(title indexing)，並實際應用於British Catalogue of Books等目錄的編排上。但是卡特(Charles Ami Cutter)認為題名索引有兩個問題：(註2)

- (1) 有些題名用詞並未呈現真正的內容；
- (2) 相同主題的書可能因題名用字不同，而被分散排列。

針對Cutter的說法，Crestadoro認為可以做參照以連結同義詞(註3)。不過Cutter認為參照仍不能解決問題，所以另提出字順式的主題目錄來解決此一問題，所以開始了Library of Congress Subject Headings的編訂。

二、第二期研究

第二期的爭論和電腦資訊系統的發展有關，1959年，Hans Peter Luhn開發了KWIC(Key Word in Context)索引法，而不將資料做控制詞彙的處理，這是應用電腦工具來做索引的開端，當這個想法被提出時，很多人認為KWIC索引法無法做到完全的回收。而資訊科學家也認為，電腦是可以做一些事，但是若要做更複雜的索引，則需要考慮語意及語法結構等問題，才能設計電腦化的索引系統。(註4)

此時期另一個重要的特色是有大型資訊檢索實驗的進行。如Cranfield所進行的不同索引語言在檢索成效上的評估研究，以及控制詞彙的效果之探討。透過實驗，Cranfield認為：(註5)

- (1) 增加了索引的深度與詳盡性，將會提高檢索的回收率並降低精確度
- (2) 詞彙的專指性(specificity)會左右檢索的精確度與回收率
- (3) 詞彙間的參照會提升回收率但降低精確度
- (4) 輸入階段的詳盡、專指與連結層次會因查詢情況而影響其效果
- (5) 必需在輸入資料時，為回收率與精確度找出較佳的平衡點

Cranfield的另一項重要成就，就是提出以回收率與精確度，做為資訊檢索效能評估的操作型定義。不過，也有人質疑此實驗的統計設計是否合理？自變項和依變項應如何定義？此外，該精確度與回收率的計算完全基於檢索所得之資料與問題的相關情況來判定，但所謂的「相關」與否，其實是很主觀的判斷。

Aitchison等人也做了相似的實驗，他們探討INSPEC資料庫的索引，分別以title、title加abstract、及紙本的Science Abstract所用的索引詞、自由語詞、控制詞彙等去檢索，比較其結果(註6)；Keen and Digger也在1972年比較資訊科學中不同索引詞彙的檢索效果。他們認為：非控制詞彙的效果不比控制詞彙差，兩者的差異不顯著。(註7)

總而言之，這個時期的研究主要在探討：那一種索引語言較好，而從實驗中發現，自然語言的表現即使不比控制詞彙好，但也不會太差。此外，這個時期的研究主要是實驗性的研究，而非對實際的使用者進行調查。

三、第三期研究

第三期的研究約始於1970年代中期，此時期的研究主要是個案性的研究，傾向於特定環境，如針對某線上資料庫做測試。由於1970年代的技術環境已有不同，所以在不同環境下所做的實驗不一定能互相比較。

1977年，Carrow and Nugent以National Criminal Justice Reference database進行索引語言(即控制詞彙)與自由詞進行比較研究，他們發現二者的精確度相當，但索引詞檢索的回收率較高。所以他們認為這二種索引模式應該相輔相成。(註8) 1978年，Henzler也以CANCERNET資料庫比較索引詞及自由語詞，她發現有35%的題名用詞無法在控制詞彙表中找到意思相近的語詞，相對的，也有50%的指定索引詞無法表現在自由語詞中。所以她也認為，控制詞彙與自由語詞二者應是相輔相成的工具。(註9)類似的研究，Markey等人也會用在ERIC資料庫的查詢上。Markey等人收集了165個正在查尋ERIC之查詢者所用的自由語詞查尋敘述，看看它們是否可以用ERIC descriptors表示。結果發現，大約每八個就有一個查詢敘述無法以控制詞彙表示。此外，他們也同時用控制詞彙及自由語詞查尋六個問題，發現自由語詞查尋的回收率較高但精確度較控制詞彙低。(註10) Tenopir針對Harvard Business Review Online這一資料庫，以36個問題分別查尋全文、摘要、題名、敘述語(descriptor)，發現全文檢索的回收率較高但精確度較低。(註11)

全文資料庫的特色主要在於其檢索點多，但當資料量越大時，就越難達到可以忍受的精確度。Dubois針對此問題，將全文檢索與控制詞彙檢索的優缺點條列如下：(註12)

自由語詞檢索的優點

- (a) 成本低
- (b) 簡化查尋
- (c) 全部的資訊內容都可檢索
- (d) 每一個字都有相等的檢索價值
- (e) 無人為的索引錯誤問題
- (f) 新詞彙也可用來檢索

自由語詞檢索的缺點

- (a) 檢索者的負擔較重，因為系統會找到很多不相關的資料，需由檢索者自給

一筆一筆過濾

- (b) 未呈現在文字中但隱含在文內的主題，無法被檢索出來
- (c) 缺少由一般性主題連結到專指性主題之功能
- (d) 使用者需先知道相關詞彙才能查尋

控制詞彙的優點

- (a) 解決很多語意上的問題(如同義詞、同形異義字等)
- (b) 可表達主題之類屬(上層主題、下層主題、相關主題)間的關係
- (c) 是一種知識地圖的呈現

控制詞彙的缺點

- (a) 成本高
- (b) 所包含的範圍未必適當
- (c) 會有人為的錯誤
- (d) 可能會有過時的語詞，新詞彙的加入速度太慢
- (e) 詞彙間的關係，難以系統性的建立

綜合而言，第三期的研究多針對某一資料庫檢索效能之研究，且認為較好的系統是控制詞彙與自然語言並用的系統，同時也再次肯定檢索點的多寡將影響檢索效果的說法。

四、第四期研究

透過第三期在不同學科資料庫的檢索研究，發現要達到最有效率的研究，必需結合自由語詞及控制詞彙，因為這二者是互補的工具。而第四期的研究則自1980年代開始，這時期主要的研究特色，是對真正的使用者查尋進行調查，同時從這時期起，電腦與網路技術的發展也進入另一階段。相關技術如 OPACs, CD-ROM, GUI, Hypertext, Internet 都已相當普及。以下即說明此時期之相關研究重點。

a. Online searching and CD-ROM

Saracevic 等人曾請40個使用者提供問題，並請39個查尋者來查尋這些問題，每一個問題都有9個人來查尋，所以總共做了360個不同的查尋，查尋結果根據相關性與有用性加以分析，發現彼此間差異極大，此研究最大的發現可能是每一個人對問題的闡釋都不相同，每一個查尋者也都有查到和別人不同的資料。(註13) McCue 也針對公共圖書館的查尋能力做調查，他請21個公共圖書館針對相同的問題，分別查尋兩個資料庫，查尋者並不知到他們在被做測試，查尋結果也是跟據所用的策略、一般技術及結果打分數，結果分數從419分到155者都有，同時也發

現，資料找到越多的分數越高。(註14)

Fidel 曾針對47個專業查尋者所做的281個查詢，調查其使用控制詞彙與自然語言的情形，在這281個查詢中，總共使用了3189個查尋詞彙，其中1607個詞彙為控制詞彙，1582個詞彙為自然語言，深入分析後，發現影響查詢詞彙的選擇因素主要有三方面：(註15)

- (a) 問題相關因素(request-related factors)，占32%，如所用的詞彙已相當專指或對回收率的需求程度等，
 - (b) 資料庫相關因素(database-related factors)，占48%，如該資料庫可能沒有提供索引典，或使用者的問題需要查尋多個資料庫方能滿足等，
 - (c) 查詢者相關因素(searcher-related factors)，占20%，如個別的喜好。
- 若再進一步分析，到底什麼原因會影響使用者呢？主要有：(註16)
- (a) 查尋者的查詢風格，
 - (b) 問題所牽涉的主題領域，如科學界的查詢者，使用自然語言較多，因為其詞彙意義比較精確，
 - (c) 其問題需要查詢數個資料庫方能解決：所需查詢的資料庫越多者，使用自然語言的也比較多，
 - (d) 索引典及索引的品質。

b. OPACs

在OPAC環境中，自然語言大多只能來自title，而控制詞彙可能來自LCSH或PRECIS索引語言。專家們認為如果增加OPAC的主題資料，如在MARC中，加上目次或摘要，將會提升檢索效能。所以Cousins曾以三個MARC資料庫來測試此一問題，每一個都做自然語言及PRECIS的查詢處理，發現以PRECIS來加強檢索，效能較佳，若再加上自動切截技術，則自然語言查詢，效能較佳。不過，其中有一個問題是很明顯的，即自然語言將控制詞彙的負荷推給了使用者，自動切截或參照的使用雖然提高了回收率，但是並未解決使用者選擇詞彙的問題。(註17)

Peters and Kurth's analysis 也曾分析學術圖書館線上目錄查詢記錄，發現有66%以上的線上查詢者同時使用兩種查詢語言，而有58%的查尋者，是以非控制詞彙(題名的關鍵字)開始查尋。平均而言，使用者使用非控制詞彙的比例較高。所以研究者建議應加強題名關鍵字的查尋。(註18)

c. 新介面及自然語言查詢技術的發展

系統設計者在此時期也不斷的努力以改善使用者介面，因為若有親和性的介

面，將能吸引使用者自己做查尋。而且介面的設計也會影響自然語言與控制詞彙的呈現。漸漸的 GUI(Graphic User Interface)及 Windows 介面取代文字介面，這樣的介面更便於線上顯示索引典或控制詞彙，解除了控制詞彙難以被了解的限制，轉而成爲輔助使用者得知檢索詞彙的好幫手。

另一個重要的發展是自然語言查詢，自然語言查詢的發展有很多面，如 CAIRS 系統，以前是一控制詞彙系統，現在則以索引典爲知識庫，發展智慧型的檢索，索引典定義了詞彙之關係，讓使用者可以擴大或縮小其查尋範圍。研究性的專家系統也漸漸商業化，如 TOPIC 是英國 Verity Inc. 開發的系統，原爲一研究性的分類索引專家系統，可用來做概念導向查詢(concept-based retrieval)，目前已商業化。

綜合上述研究，我們可以得知，歐美國家經過近百年的爭論，研究者並未解決孰優孰劣的問題，不過倒是得到一個共識，即控制詞彙與自然語言這兩種索引模式，應該是可以互相截長補短。而最近幾年的研究重點已轉移到如何改善衍生索引(derived indexing，即自原文中截取索引詞)及自動索引的品質上，以減少控制詞彙索引的必要性，因爲控制詞彙索引模式成本昂貴，所以，任何有經濟效益的替代方案永遠吸引人，而且，對全文資料庫而言，衍生索引確有其必要。然而，對 CD-ROM、OPAC 及線上資料庫的查詢而言，要使用者想到相關的、正確的或完整的詞彙，來建立檢索策略還是有困難，此時，控制詞彙應可發揮其效果，如：可在螢幕上顯示控制詞彙，以協助使用者找到合適的檢索詞。

參、中文資訊檢索系統相關研究及問題

有關中文資訊系統檢索之相關研究，主要可分爲三方面來說之：

一、有關使用者查詢行爲的研究

民國78年吳明德曾針對9所大學圖書館、10所專門圖書館、7所公共圖書館，調查其目錄使用情形，其中分析了讀者檢索時握有的書目資料，在大學圖書館部份，讀者握有的書目資料分別爲：著者32.0%，書名44.5%，主題字彙39.2%，完整標題2.4%，完整索書號0.9%，分類號2.7%，其他8.9%；在公共圖書館部份，讀者檢索時手中握有的書目資料分別爲：著者18.8%，書名30.2%，主題字彙51.8%，完整標題0.4%，完整索書號0.0%，分類號4.5%，其他9.4%；而專門圖書館的情況爲：著者29.0%，書名50.5%，主題字彙42.1%，完整標題3.7%，完整索書號0.0%，分類號9.3%，其他7.5%。(註19)民國83年，邱韻玲的碩士論文曾調查清

大OPAC 使用情形，在3663次檢索中，使用關鍵字查詢者占40.4%，書名占34%，作者占13%，標題占2.9%。(註20)民國85年，李宜容也調查中央研究院圖書館的人文及社會科學之讀者使用OPAC的情形，發現在592次的檢索中，利用書名查詢的比例占34%，作者占26.9%，關鍵字占20.8%，標題占18%。(註21)吳美美曾以「中華民國期刊論文檢索系統」的使用者做為研究對象，其中有關使用者檢索點特質的調查，發現檢索者以關鍵字為檢索點者最多，占83.1%(註22)，筆者也曾經針對使用者查網路資源之相關問題做過調查及訪談，在訪談的對象中，發現大部份的使用者只使用蒐尋引擎中最簡單的關鍵字查詢，使用其他檢索功能者不多。(註23)中華民國期刊論文索引及網路資源都是未做控制詞彙索引之系統。民國86年，曾繁娟徵求48位使用者使用標題檢索台大TULIP系統，在進行取樣期間，曾觀察使用者查尋OPAC的情況，發現主動使用中文標題檢索的人相當少。而在所徵求來的查詢者中，以前曾使用過中文標題檢索者占50%。(註24)曾元顯曾調查模糊搜尋、相關詞提示與相關回饋在OPAC系統中的成效評估，發現在116位受訪者中，使用近似字串檢索系統查詢者占73%，選用欄位式查詢系統者占27%，而在近似字串檢索功能中，有82.4%的使用者認為自動排序對瀏覽檢索有幫助，有70.6%的使用者認為相關詞提示對選擇檢索點有幫助，有58.9%的人認為相關詞回饋對進一步發現資料有幫助。(註25)

二、有關控制詞彙與非控制詞彙效能之研究

至於針對中文控制詞彙與非控制詞彙檢索效能之研究者較少，1997汪先生等(Wan, Tian-Long et. al.)收集1993年6月至1994年4月，555則電腦與資訊科學方面的科技簡訊摘要為測試系統，並以科技簡訊索引典為工具，作為自動索引的基礎，以處理斷詞的問題，並探討有索引典及無索引典之人工索引及自動索引之系統的檢索效能，測試結果如下：(註26)

回收率	精確度	
	無索引典	有索引點
0.1	0.8279	0.9104
0.2	0.8088	0.8795
0.3	0.7199	0.8117
0.4	0.6367	0.7707
0.5	0.5138	0.7115
0.6	0.3917	0.6236
0.7	0.2782	0.5539
0.8	0.1901	0.4834
0.9	0.1168	0.3935
1.0	0.0584	0.2953

回收率	精確度	
	無索引典	有索引典
0.1	0.8667	0.9916
0.2	0.8356	0.9667
0.3	0.7439	0.9430
0.4	0.6386	0.8890
0.5	0.4268	0.7371
0.6	0.2878	0.5869
0.7	0.1606	0.5045
0.8	0.0970	0.4041
0.9	0.0618	0.2728
1.0	0.0309	0.1807

Wan 等人透過此研究發現，(1) 在自動索引環境中，索引典確實可提高檢索效能，(2) 在人工索引的環境中，索引典也確實可提高檢索效能，(3) 此研究也發現自動索引與人工索引的效果一樣好

三、有關中文檢索系統之設計與研究

中文資訊檢索系統所遭遇到的問題及使用的理論與技術，大致而言，和西文資訊系統相似。有差異是中文斷詞問題。中文斷詞困難造成書目資料或全文資料，無法以程式自動擷取有意義的索引詞，進行關鍵字檢索。所以多數的 OPAC 系統以「字」為索引詞，少數是以人工標出關鍵字，做「詞彙」索引。前者往往造成檢索詞與查詢所得之詞彙顛倒的情況(如預查「國中」，但卻查出很多內含「中國」一詞的文獻)，後者仍需耗費人力，且也和控制詞彙索引一樣，會有標詞不一致的問題存在。所以近年來有些全文檢索系統採用 N-gram 索引法來解決斷詞的問題。此外，大多數人仍有中文輸入的困難，所以國內也有很多學者針對這些問題進行研究，並獲得不錯的結果。1988 年，Lian, Y 使用倒置索引法處理中文科技期刊系統，除停用字外，將文中每一個字都做倒置索引，所以整個倒置檔相當大。(註 27) 1989 年 Shi, Y 以常規式查詢方式(regular expression searching)，試圖解決中文字輸入困難的問題。(註 28) 1992 年 Lee 等人介紹電信局採用澳洲發展的 STATUS 系統，建立電傳視訊查尋的全文檢索系統。(註 29) 1992 年 Chang and Chen 等人以霍夫曼(Huffman coding)，有效的查尋壓縮的中英字典。(註 30) 1993 年 Chang and Lee 等人在交大自然語言實驗室進行中文斷詞及中文人名辨識研究。(註 31)

此外，實際的中文檢索系統也相當多，比較有特色的有簡立峰博士的尋易(Csmart)檢索系統(註32)、吳昇教授的蓋世檢索系統(GAIS)(註33)、曾元顯教授的水晶全文檢索系統(CRYSTAL)(註34)、昶陽資訊的龍捲風系統(註)。Csmart及CRYSTAL系統的功能相似，如模糊檢索(fuzzy search)、相關回饋、N-gram索引模式等，GAIS的查尋特色為可設定查尋條件(matching criteria specification)、可設定檔案屬性(file attribute specification)，此功能可設定要查尋的檔案為何、查尋範圍設定(matching domain specification)，此功能可機動性的設定只要查尋檔案中的某一部份、輸出控制設定(output control specification)、其他設定(miscellaneous options)，如可設定要連結查尋那一個伺服器、此外GAIS允許查尋語詞可以是關鍵字(keyword)、萬用字元查尋(limited expression)、常規式查尋(regular expression)、或結構式的布林邏輯(ordered boolean expression)。而龍捲風較不一樣的是有容錯查尋及相同語音查尋功能，此外它也能代理蒐尋、抓取網路資源，及檢索本系統之檔案，換言之，系統內外之資料的整合查詢是該系統的特色。

肆、國內相關資料庫主題檢索問題

由國外有關控制詞彙與非控制詞彙相關文獻之分析可知，歐美國家認為控制詞彙與非控制詞彙各有優劣，而好的系統通常二者同時採用，筆者文也對國內圖書館使用國內外光碟系統是否有用控制詞彙做一調查，發現國外的系統大都有用控制詞彙，而國內的系統大都未使用控制詞彙。(參見附錄)筆者在教授技術服務專題時，也曾請學生分別就不同主題之資料庫，以系統提供之控制詞彙及非控制詞彙欄加以檢索，比較其檢索結果。在這些檢索過程當中，我們發現下列一些問題：

- (a) 有的學生分不清楚那一個檢索是控制詞彙，那一個是非控制詞彙欄；
- (b) 有些系統的非控制詞彙欄不只一個，如LISA有title, title keyword, full-text, keyword等欄位，URICA在WebPac上有題名、作者、標題等檢索點，而其關鍵字查尋無轉屬的欄位，只以附註方式說明，若要以關鍵字檢索，可在檢索點後面加*號。Dynix有題名-字順，題名-關鍵字，主題-字順，主題-關鍵字，當然每一種檢索結果都不相同，但即使是研究所的同學，有的也分不清其間的差別，以及他該用那一種來檢索；
- (c) 在和主題相關的欄位中，我們發現有的系統的題名查詢和關鍵字查詢所採用的比對模式不同，如有的系統之題名查詢是字順比對，即檢索詞一定要

出現在題名前面，有的不論檢索詞出現在題名何處都可檢索到；而其關鍵字查尋部份，有的系統是按字順比對(即是以詞彙爲主的比對)，有的系統則將詞分爲單字來比對(即以字元爲主的比對)；題名是採字順比對的系統其關鍵字可能採單字比對；

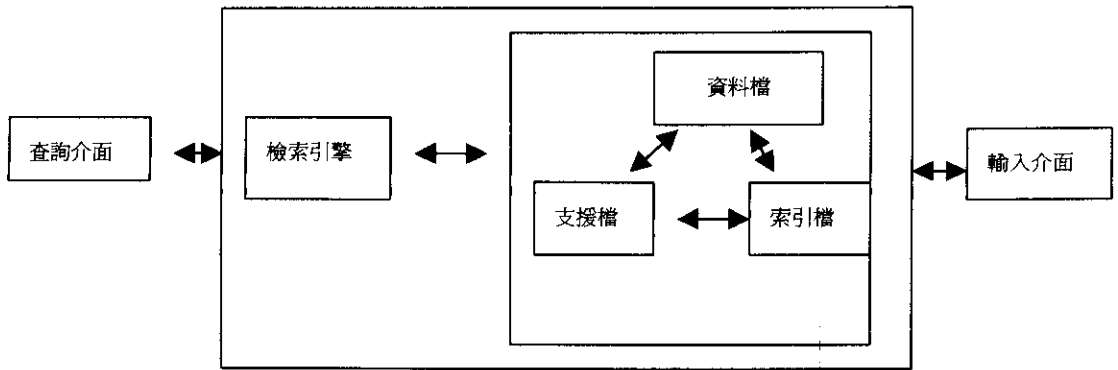
- (d) 大多數的圖書館 OPAC 所提供的中文標題檢索之查得率都相當低，這和中文標題的近幾年才開始採用有關，如台大圖書館的中文資料，大都沒有給標題，此外，中文主題詞語表的語詞並不多也是重要的因素。
- (e) 大多數的中文資料庫都沒有使用控制詞彙，即使有使用控制詞彙，也沒提供線上顯示(及提示)控制詞彙表的功能。
- (f) 經過深入的討論，同學們認爲較好的系統是：有提供索引典瀏覽、介面親和的系統。這和使用者不太能建構其檢索詞彙有關，尤其對西文資料庫而言，若英文或專業程度不好，要找到相關的詞彙實在有困難。
- (g) 關鍵詞索引所包含的欄位各系統不盡相同，但一般而言，關鍵字欄查尋所得的筆數應多於題名查詢所得筆數，但在我們的查詢中，「中華民國期刊論文影像索引系統」這個資料庫卻有不一樣的結果，我們分別用「教改」、「國中」、「腦科」三個檢索詞是查尋篇名及關鍵字，三個檢索所得到的結果都是篇名查詢所得的筆數多於關鍵字。

伍、結論與建議

一、影響系統檢索效能之因素分析

綜合上述研究，筆者認爲，事實上並無那一個研究可以讓我們拿來做爲是否採用控制詞彙或不採用控制詞彙的決定。因爲這其中牽涉因素甚廣。以汪先生等人的研究而言，雖然其結果認爲有控制詞彙的效果比無控制詞彙的效果好，但是這個研究的樣本是資訊科學文獻，其控制詞彙也是由專家給予，整個研究過程都做了控制，它是否適用於一般圖書館的情況，則不得而知。

資訊系統的檢索效能，不能只看是否使用控制詞彙，或資訊系統本身設計的好不好，乃是受到三種因素的影響，一是系統設計、二是資訊處理者(含資訊徵集者、組織者)、三是檢索者。而系統設計實則又包括三個部份：資料庫、檢索引擎、使用者介面，如下圖所示。茲分別說明如下：



1. 系統設計部份

(1) 資料庫部份

在資料庫部份主要是處理資料的輸入、資料檔的管理與儲存、索引檔的建立與維護、支援檔(如權威檔、索引典、語料庫等)的建立與維護等工作。

(2) 檢索引擎部份

在檢索引擎部份，主要在提供布林邏輯、切截檢索、拼字核對、鄰近字檢索、乏析檢索 (fuzzy search)、多語文檢索、索引典參照檢索、或其他智慧型檢索等功能。

(3) 使用者介面

使用者介面是一人機介面，系統的功能應能由此介面表現出來。此外，使用者介面的設計必須具有親和性，介面上的用語必須簡單明瞭，讓使用者對各項功能一目了然，當使用者碰上困難時，系統也應該給予指示訊息，使其不致亂無頭緒，無法處理而不再使用。

這三個部份息息相關，當我們決定了系統的檔案類型時，也必須同時考慮其檢索引擎的功能，如資料檔是一書目性檔案或為全文資料檔。若為書目檔，系統需考慮要描述資料的那些屬性，而在主題屬性部份，要作人工索引或自動索引，是否要採用控制詞彙，提供那些檢索功能等。若為全文資料檔，系統也需考慮是否要建索引，是否採用控制詞彙，提供那些檢索功能。

2. 資訊處理者部份

資訊的處理者主要包括資訊的徵集者及資訊組織者。一資訊系統不可能包含所有的資訊，而會按資訊的良窳、主題、程度等條件來選擇資訊。而資訊的組織者主要在對資訊做加值的工作，包括描述資料、分析資料、建立資訊間的關係，建立索引典等，換言之，即為索引的工作。資訊系統的索引工作可分兩方面來談，一為欄位化的資訊檢索系統，二為全文資訊檢索系統。此外，超文件系統就

文件性質而言，也可算是全文系統，但其使用方式和一般的全文系統有所不同，茲分別說明之：

(1). 欄位化的資訊系統：欄位化資訊的索引相關問題包括：

- a. 需要描述資訊的那些屬性？如圖書館界描述各類型資訊的題名及著者敘述項、版本項、特殊資料細節項、出版項、稽核項、叢書項、附註項、標準號碼等八大項屬性。而目前電子圖書館所處理的資料，如博物館資料、地理資料、檔案資料等所描述的資料屬性往往比圖書館界所描述者更加深入。
- b. 各項屬性的描述是否需遵照標準？如書目資料的描述遵守中國編目規則，主題分析遵守分類表及標題表、索引典等。
- c. 是否做主題分析？欄位化的資料乃以書目代表資料全文，有關全文的內涵屬性必需透過主題分析來描述，找出可以代表全文意義的主題語詞。
主題分析的方法一般而言有兩種：人工索引或自動索引
 - (a) 人工索引：以人工來分析資料的內容最大的優點，是索引人員可以將文件本身所隱含的主題表現出來。但其缺點是費時費力，相關的問題是索引者的品質是否有能力了解文件內涵？是否清楚要描述些什麼？是否能以精確的語詞描述資料？主題分析的深度與詳盡度如何？索引者本身或不同的索引者之間對相同主題是否能維持一致性的描述？
 - (b) 自動索引：自動索引是指以程式找出可代表全文內涵的主題語詞，其相關方法另說明於全文檢索系統中。
- d. 是否使用控制詞彙：無論是人工索引或是自動索引，其代表文件之主題內涵的語詞可以是不做控制的關鍵字，也可以是控制詞彙，如索引典等。索引典的功能主要有：闡明專門術語(clarify terminology)、提供參見(provide cross referencing)、提供同義字(supply synonyms)，換言之，索引典主要在定義「詞」之語意，以及相關的概念層級。
- e. 片語詞彙如何建索引檔索引者用來描述文件的主題語詞可能是單字詞也可能是複合與詞，將複合語詞建成索引，其方法有三種：一是分解成單字詞，二是以複合詞做索引，三是做單字詞索引也做複合詞索引
- f. 是否提供全文檢索功能：欄位化的資料庫除可提供欄位化的檢索以外，也可以將整筆書目視為非結構性的資料，提供全文檢索。

(2). 全文檢索系統

全文檢索系統是指可以用一種檢索機制，提供使用者以任一字串或字串邏輯組合為條件，以檢索相關文件段落。所要檢索的全文可以完全不做索引

處理，也可以做自動索引。全文檢索相關方法主要有：

a. 全文掃描 (Full-text scanning)

對於比較小量的文件資料庫而言，全文掃描是搜尋檔案內資訊不錯的方法。這種方法主要是在檔案中，以精確比對(exact match)方式將使用者的查詢語詞與原文字串進行快速的比對；假如比對程式發現有符合條件的文獻，即為成功的搜尋。目前大部份的文書處理軟體皆提供全文掃描的功能，讓使用者搜尋檔案或本文。全文掃描法可以省略人工索引的人力，也不必配合索引典等控制詞彙，同時也節省了儲存索引檔案所需的硬碟空間，但缺點就是速度較慢。(註 36)

b. 逐項反轉索引法 (Inversion of terms)

逐項反轉索引法乃是除了停用字(如冠詞、定冠詞等)外，擷取文獻中每一個字，依字典方式排列，記錄其在文獻中的位置，製做成索引檔。這種索引檔又稱為倒置檔。有了這個索引檔，當使用者打入檢索詞時，系統即將之與此索引檔做比對，若有配符(match)者，即透過位置指標，檢索出資料檔的資料。

逐項反轉索引法的缺點在於需要編製索引的時間，及大量的儲存索引檔的硬碟空間，但其查尋速度較全文掃描快很多。比較適用於大型、靜態、或資訊只需周期性更新的資料庫。(註 37)

c. 簽名檔 (Signature file)

簽名檔利用重疊編碼 (Superimposed Coding) 的原則，將文件轉換成一固定長度的簽名 (Signature) 以加速字串比對。查尋時也將查尋字串建立查尋簽名 (Query Signature)，再將查尋簽名與所有文件一一比對。簽名檔所需索引空間較逐項反轉法節省，但比全文掃描耗費時間；簽名檔法的檢索速度比逐項反轉法慢，但比全文掃描快。(註 38)

d. 聚叢法 (Clustering)

聚叢法主要在利用統計方法，計算文件之特徵值，將特徵相似的文件歸為一類，取一最具代表性的特徵 (Centroid) 做為這類文件的特徵，查尋時則找尋與查尋項最接近之文件特徵做為初步結果，之後再進一步比對，並歸屬這類文件的所有文件特徵以找出最相近文件。(註 39)

e. 字典式索引法 (dictionary-based indexing)

即事先建好一主題語詞庫，將主題語詞庫與全文配對後，以配符的主題語詞代表該文件，加到索引檔中，另有一些專家系統是採用字典/推論規則式索引法 (dictionary/rule-based indexing)，這是除了上述的主題語

詞庫外，再加上相關推論規則，以程式判斷文件之主題。(註40)

(3). 超文件/超媒體系統

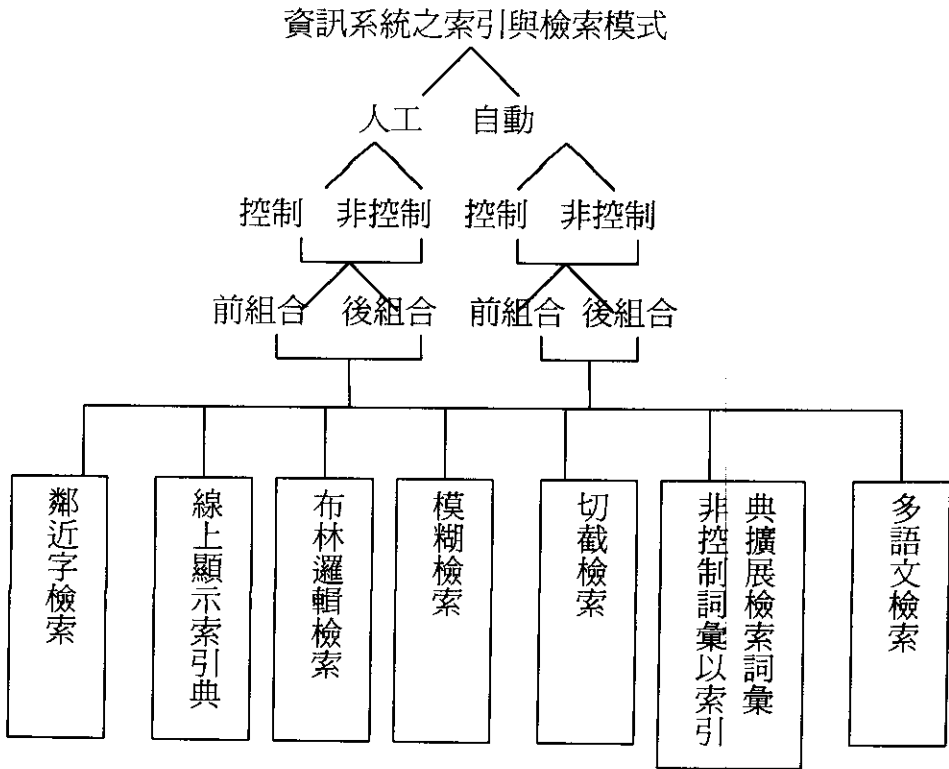
超文件/超媒體系統也是一種重要的檢索方式，因為它提供連結功能，讓我們可以找到相關資料，此外，也可以利用連結提供瀏覽功能，這是在Internet環境中，任何檢索系統(無論是欄位化檢索或全文檢索)都應該加上的功能。也是將各種資訊類型，如影像、聲音、視訊、文件等整合在一起的重要工具。

3. 系統使用者

使用者也是影響系統檢索效能的因素之一，使用者方面的問題包括：

- (1) 使用者是否能正確的闡釋其需求？並非所有的使用者都能精確的闡釋其主題，將其心中的需求化為檢索的問題。他的闡釋可能太廣、也可能太窄，或不著邊際。
- (2) 是否能正確的將問題以詞彙描述？即使使用者可以正確的說出其問題，但可否將問題化成數個合適的查尋語詞？我們從下列的一些研究中可以知道，不同的人對相同的問題往往有不同的闡釋，會用不同的語詞來查尋，所以查尋的結果也就不盡相同。
- (3) 是否知道控制詞彙和非控制詞彙的差別？如果系統同時提供了控制詞彙及非控制詞彙，使用者是否了解這二者的差別。跟據我們曾做過的調查，使用者大都不了解圖書館線上公共目錄上所提供的主題欄之意義為何，檢索時也大都使用關鍵字，而非主題欄。
- (4) 是否願意使用控制詞彙？如果系統有提供控制詞彙檢索，使用者是否願意使用。
- (5) 是否知道其他檢索功能(truncate, boolean, proximity, ...)之意義及使用方法？
- (6) 精確度與回收率的需求：每位使用者在查尋資料時，不一定都要求全或求準，有的人只要有找到相關的資料就好，有的人希望找全，需求不一。

如果我們將上述因素一起考慮，可以下圖來表示索引與檢索這一體兩面的關係：



二、建議

綜合上述之研究，有關中文資訊系統主題檢索問題，筆者提出下列之建議

1. 近似字串查詢系統已有不錯的效果，建議圖書館考慮採用，以解決使用者無法輸入精確之主題詞的困擾；
2. 全文系統無法解決主題詞的語意問題，控制詞彙仍有其必要，它也是一種知識地圖，國內還是應有國家級的機構來建立索引典等控制詞彙表；
3. 由於人工索引成本極高，故人工索引整體效益應再仔細評估；
4. 有採用控制詞彙之系統，應加上線上顯示控制詞彙的功能，以協助使用者建立查尋語詞，並能建立語詞間之參照關係，將使用者所用之同義詞轉成系統用詞；
5. 未採用控制詞彙的系統，應可用控制詞彙表擴充使用者之檢索詞，讓同一主題之同義詞都變成檢索詞；
6. 圖書館界所使用之中文主題詞表示尚不完整，應參考農資中心、立法院、科資中心已建立之索引典，擴充中文主題詞表，加快增詞的速度，並建立主題

詞表合作機制；

7. 中文主題詞表應採中英對照方式建立；
8. 中文書目資料庫應加上摘要或目次，以增加檢索點。
9. 圖書館界持續已久的主題編目成效如何，應有更深入的調查研究，並配合合作機制、人力成本、新檢索技術、檢索效能等整體考慮，以決定未來發展的方向。

附註

- 註1：Jennifer Rowley, "The controlled versus natural indexing languages debate revisited: a perspective on information retrieval practice and research," Journal of Information Science, 20(2) 1994 p: 110.
- 註2：Ibid. p: 111.
- 註3：Ibid.
- 註4：Ibid.
- 註5：E. M. Keen, Aspects of Computers in Libraries International '91: Proceedings of the 5th Annual Conference on Computers in Libraries, held in London in February 1991. (Meckler, London, 1991)
- 註6：T. M. Aithchison et al. Comparative Evaluation of Index Languages 2 vols (Institution of Electrical Engineers, London, 1969-1970)
- 註7：E. M. Keen and J. A. Digger, Report of an Information Science Index Languages Test (Department of Information Retrieval Studies, CLW, Aberystwyth, 1972)
- 註8：D. Carrow and Nugent, "Comparison of free text and system." In: Information Management in the 1980s: Science 40th Annual Meeting, September 26- October 1, 1977 (Knowledge Industry Publications, White Plains, NY, 1981) : 131-138.
- 註9：R. G. Henzler, "Free or Controlled Vocabularies," International Classification 5 (1) (1978) : 21-26.
- 註10：K. Markey, et al., "An analysis of Controlled Vocabulary and Free Text Search Statements in Online Searches," Online Review 4 (1982): 225-235.
- 註11：C. Tenopir, "Full Text Database Retrieval Performance," Online Review 2 (1985):149-164.
- 註12：C. P. R. Dubois, "Free Text Versus Controlled Vocabulary: a Reassessment," Online Review 11(10) (August 1987): 243-253.
- 註13：Tefko Saracevic, Paul Kantor, Alice Y. Chamis, Donna Trivison. "A Study of

- Information Seeking and Retrieval," Journal of the American Society for Information Science, 39(3) (May, 1988): 161-216.
- 註 14 : Janice H. McCue. Online Searching in Public Libraries: a Comparative Study of Performance. Metuchen, NJ: Scarecrow Press, 1988. 272p.
- 註 15 : Raya Fidel, " Who Needs Controlled Vocabulary?" Special Libraries (Winter 1992): 1-9.
- 註 16 : Ibid. 5-7.
- 註 17 : S. A. Cousins, " Enhancing Subject Access to OPACs: Controlled Vocabulary Versus Natural Language," Journal of Documentation 48(3) (1992) : 291-309.
- 註 18 : T. A. Peters and M. Kruth, " Controlled and Uncontrolled Vocabulary Subject Searching in an Academic Library Online Catalog," Information Technology and Libraries (September 1991)
- 註 19 : 吳明德, 我國圖書館目錄使用之研究, 行政院國家科學委員會專題研究計劃成果報告, 民79。NSC 79-0301-H002-17
- 註 20 : 邱韻玲, 「國立清華大學線上公用目錄使用調查 -使用者查詢過程之研究」, 台灣大學圖書館研究所碩士論文, 民國83年。
- 註 21 : 李宜容, 「人文及社會學科讀者使用線上公用目錄檢索詞彙之研究」, 淡江大學教育資料科學研究碩士論文, 民國85年。
- 註 22 : 吳美美, 「中文檢索詞彙初探」, 在 21世紀資訊科學與技術的展望學術研討會論文集 (民國87年), 頁167-191。
- 註 23 : 陳昭珍, 虛擬圖書館書目控制與資訊組織模式之探討 (I) 行政院國家科學委員會專題研究計劃成果報告 民國86年。NSC 86-2418-H-030-001
- 註 24 : 曾繁娟, 「中文標題檢索效益之研究」,
<http://www.lib.ntu.edu.tw/pub/univj/uj2-1/uj2-1-9.html>
- 註 25 : 曾元顯, 「模糊搜尋、相關詞提示與相關詞回饋在 OPAC 系統中的成效評估」, 中國圖書館學會會報 第61期 (1998), 頁: 103-125。
- 註 26 : Tian-Long Wan, Martha Evens, Yeun-Wen Wan, and Yuen-Yuan Pao, " Experiments with Automatic Indexing and a Relational Thesaurus in a Chinese Information Retrieval System," Journal of the American Society for Information Science, 48(12) (1997): 1086-1096.
- 註 27 : Y. Lian, " Technical Periodicals Data Base Retrieval System," Proceedings of 1988 International Conference on Computer Processing of Chinese and Oriental Languages, Toronto, Canada. (1988) 367-370.

- 註28 : Y. Shi, Regular Expression Searching in Chinese Information Retrieval. Ph. D. Dissertation. Berkeley, CA: Library and Information Studies, University of California at Berkeley, CA. (1989)
- 註29 : Y. S. Lee, B. R. Lo, S. F. Huang, S. W. Lian, and Y. C. Chen, "The Planning and Designing of Full Text Retrieval System Through Chinese Videotex System," The Bimonthly Journal of Telecommunication Research, Taiwan. 22, (1992) 461-469.
- 註30 : C. C. Chang and W. Y. Chen, " Looking up words in a compressed Chinese-English Dictionary," Journal of Electrical Engineering, 35 (1992): 145-154.
- 註31 : C. S. Chang, & P. C. Lee. "The Situation and Development of Information Retrieval," Computers and Communication, Taiwan, 17(1993): 3-6.
- 註32 : Lee-Feng Chien, " Csmart Project: Toward Intelligent Chinese Information Retrieval," <http://scmart.iis.sinica.edu.tw/research.html>
- 註33 : GAIS Project, "Man Page: gais,"
<http://133.1.96.43/~seke/WEBgais/docs/gais.1.html>
- 註34 : 曾元顯, 「模糊搜尋、相關詞提示與相關詞回饋在OPAC系統中的成效評估,」中國圖書館學會會報第61期(1998), 頁: 103-125。
- 註35 : 請參見龍捲風軟體功能(光碟片)
- 註36 : 小蝶, 圖書資訊檢索技術, (台北: 文華, 民國85年): 90-92.
- 註37 : 同上註, 頁92-95。
- 註38 : 同上註, 頁96。
- 註39 : 同上註, 頁97。
- 註40 : Wolfgang Von Keitz, "Automating Indexing and the Dissemination of Information," INSPEL 20(1) (1986): 47-67.

資料庫主題索引現況

資料庫名稱	形式	控制語言	索引典/標題表
中西文參考書選介系統	on-line	無	
中研院古籍全文檢索系統	on-line	無	
中華博碩士論文索引檢索系統	on-line	無	
中華博碩士論文摘要檢索系統	on-line	無	
中文期刊目次檢索系統	on-line	無	
教育論文摘要資料庫	on-line	無	
教育新聞剪輯資料庫	on-line	無	
教職員著作目錄資料庫	on-line	無	
中文期刊目次索引資料庫	on-line	無	
中文報紙論文索引資料庫	on-line	無	
中文博碩士論文目錄資料庫	on-line	無	
國立政治大學教職員著作目錄	on-line	無	
中文圖書資訊文獻摘要資料庫	on-line	無	
卓越商情資料庫	on-line	無	
美國 OCLC FirstSearch 檢索系統	on-line		
EBSCO Host 全文資料庫	on-line		
Cambridge Scientific Abstracts 資料庫	on-line	無	
UMI ProQuestDirect 全文影像資料庫	on-line	無	
中華民國期刊論文索引	CD-ROM	無	
全國西文科技圖書與西文期刊聯合目錄	CD-ROM	無	
報紙合訂本資料庫	CD-ROM	無	
ERIC 教育研究資料庫	on-line	有	Thesaurus of ERIC Descriptors
Academic Index	CD-ROM	有	LC Subject Headings
Art Index 全球藝術發展史文獻索引	CD-ROM	有	

A-V online	CD-ROM	有	NICEM Category and Subject Headings
ABI/INFORM(Abstracted Business Info.)	CD-ROM	有	Thesaurus
Applied Science & Technology Index	CD-ROM	有	Thesaurus
Books In Print 美國出版社發行中圖書資料	CD-ROM	有	
Boston Spa Book 大英圖書館文獻供應中心圖書目錄	CD-ROM	有	
Boston Spa Conferences 大英圖書館文獻供應中心會議論文目錄	CD-ROM	有	
Boston Spa Serials 大英圖書館文獻供應中心期刊目錄	CD-ROM	有	
Chemical Abstracts on CD-ROM 1998	CD-ROM	無	
Cross-Cultural	CD-ROM	有	Culture Materials Subject Codes
Dao 美加歐博碩士論文摘要資料庫	CD-ROM		
DAO (Dissertation Abstracts on Disc)	CD-ROM	有	Thesaurus
EconLit	CD-ROM	有	Economic Literature and Index of Economic Articales
ERL 資料庫	CD-ROM	有	Thesaurus
EMBASE CD: Drugs & Pharmacology	CD-ROM	有	MALIMET(Master List of Medical Terms)
EMBASE CD: Neurosciences	CD-ROM	有	EMTREE Thesaurus
Humanities Index 人文科學期刊文獻索引	CD-ROM	有	
International ERIC 澳英加教育文獻資料庫	CD-ROM		
LegalTrac	CD-ROM	有	EasyTrac subject heading
Library Literature	CD-ROM	有	Thesaurus
LISA (Library and Information Science Abstract)	CD-ROM	有	Thesaurus
MLA International Bibliography	CD-ROM	有	The MLA International Bibliography Thesaurus
Mathsci Disc	CD-ROM	無	
Newspaper Abstracts on Disc	CD-ROM	有	Thesaurus
PAIS (Public Affairs Information Service)	CD-ROM	有	Thesaurus
Periodical Abstracts 一般性期刊文獻索摘	CD-ROM	有	

控制詞彙與非控制詞彙主題索引問題之探討

PsycLIT	CD-ROM	有	Thesaurus
R&D in Russia	CD-ROM	無	
Sociofile	CD-ROM	有	Thesaurus of Sociological Indexing Terms
Social Science Citation Index	CD-ROM	無	
Sport Discus	CD-ROM	有	Thesaurus
Ulrich's Plus 全球65000家公司所發行定期及不定期之期刊目錄	CD-ROM	有	

資料庫名稱	網址	類型	主題檢索		控制詞彙
			控制詞彙工具	關鍵詞	
當代文學史料影像全文系統	http://192.83.186.7/cgi/m_writer	online	✓		
國立台灣大學佛學研究中心佛學資料庫	http://140.112.2.89/SEARCH/srch.htm	online	✓		
漢籍全文資料庫 (中央研究院)	http://www.sinica.edu.tw/ftms-bin/ftmsw3	online	✓		
中央研究院平衡語料庫	http://www.sinica.edu.tw/ftms-bin/kiwi.sh	online	✓		
文物圖象研究室資料庫	http://saturn.ihp.sinica.edu.tw/~wenwu/ww.htm	online	✓		
台灣地質文獻目錄查詢	http://140.115.123.52/query/gpquery.htm	online	✓		
中央研究院植物研究所標本館資料庫	http://euler2.sinica.edu.tw:8080/hast/index_c.html/	online	✓	✓	台灣植物誌 (Flora of Taiwan)
魚類資料庫	http://com5.iis.sinica.edu.tw/cgi-bin/accept-query	online	✓		
台灣魚類資料庫	http://140.109.48.161/	online	✓		

水產資料庫全文檢 索查詢	http://www.tfrin.gov.tw/D-FI/lab.fii/database/index_c.html	online	✓		
造船工程名詞資料 庫	http://itis.usddc.org.tw/shipterm.htm	online	✓		
染化技術資料庫	http://www.dfm.com.tw/te-data.htm	online	✓		
中醫藥文獻資料庫 (國立中國醫藥研 究所)	http://www.nricm.edu.tw/index1.htm	online	✓		
中醫藥資訊暨文獻 資料庫	http://www.nricm.edu.tw/cmr2ndocs/cmr2n2-7.html	online	✓		
勞工安全衛生研究 所資料庫	http://192.192.46.66/DATABASE.HTM	online	✓		
建築資訊服務系統	http://arch.org.tw/	online	✓		
亞太全球建築專業 資訊網資料庫查詢 系統	http://abba.com.tw/database/S-expert.htm	online	✓		
造船輪機術語釋義 資料庫	http://itis.usddc.org.tw/name-term.htm	online	✓		
台灣船期航運整合 資料庫	http://www.ivnet.com.tw/marine.htm	online	✓		
專利公報資料庫 (中央標準局)	http://patent.seed.net.tw/	online	✓		
永然法律資料庫	http://www.ivnet.com.tw/lee/bhtm/search.htm	online	✓		
人事法規全文資料 庫	http://www.sinica.edu.tw/fms/law.html	online	✓		
證券專業資料庫線 上服務	http://sfi.anjes.com.tw/sfi/service.htm	online	✓		
卓越商情資料庫 (ebds)	http://info.anjes.com.tw/ebds	online	✓		

中華民國企管文獻摘要資料庫 (mars)	http://ebds.anjes.com.tw/mars	online	✓		
台灣工商名錄資料庫	http://www.ivnet.com.tw/search/typa.htm	online	✓		
Virtual Trade Mart	http://www.tptaiwan.org.tw/tto/tto.htm	online	✓		
工廠登記資料檢索	http://www.moeaidb.gov.tw/~sed/2000.htm	online	✓		
台灣化工廠商檢索查詢	http://www.ivnet.com.tw/chemnet/me1.htm	online	✓		
建材與設備電話簿資料庫	http://203.66.111.10/build/search.cfm	online	✓		
慶宜工商電話簿資料庫	http://203.66.111.10/trade/search.cfm	online	✓		
台灣禮品業供應商檢索資料庫	http://www.gift.com.tw/page5c.htm	online	✓		
食品CAS認證廠商資料庫線上查詢	http://203.72.130.5/CAS/CasSFb3.asp	online	✓		
CAS廠商資料查詢	http://www.firdi.org.tw/cgi-bin/cas.pl	online	✓		
行政院公報全文資料庫	http://www.gazettes.com.tw/	online	✓		
台灣攝影業界資料庫查詢	http://www.thenet.com.tw/photo/6/6.htm	online	✓		
當代藝術作家系統	http://192.83.186.7/cgi/m_art	online	✓		
電影海報資料庫查詢系統	http://www.cs.nccu.edu.tw/~s8325/db/movie.htm	online	✓		
Cyber MIC資訊新聞資料庫標題查詢	http://mic.iii.org.tw/michtml/news.html	online	✓		

ChemNET 化工資訊網	http://www.chemnet.com.tw/scripts/data/frame.asp	online	✓		
房屋租售資料交換	http://www.bullet.com.tw/usr-bin/run8.html	online	✓		
房車買賣資料交換	http://www.bullet.com.tw/usr-bin/run3.html	online	✓		
ITIS 經濟部科技專案計畫執行成果摘要資料庫	http://www.tier.org.tw/tier/querydb.htm	online	✓		
工業技術研究院進行中研究資料庫	http://itrinews.itri.org.tw/info/database/moeadb-1.htm	online	✓		
國科會研究人才資料查詢	http://www.nsc.gov.tw/nsc/rs/rs_query.htm	online	✓		
交通大學思源基金會人力資料庫	http://db2serv.cc.nctu.edu.tw:1975/humanresource/htmls/right.html	online	✓		
NBINet Web OPAC 聯合目錄	http://nbinet.ncl.edu.tw/screens/opacmenu_chi.html	online	✓	✓	中國圖書分類法、中國圖書十進分類法、美國國會圖書館分類法
中華民國政府出版品目錄系統	http://192.83.186.115/cgi/m_ncl5?1	online	✓	✓	中國圖書分類法、中文圖書標題表
中文現期期刊目次資料庫_漢珍資訊系統	http://www.tbmc.com.tw/journals.htm	online	✓		
UMI 博碩士論文索引資料庫_漢珍資訊系統	http://140.119.115.31/online/ttsweb?@0:0:1:/disk1/dai/dai:回主頁:/	online	✓		
中華博碩士論文線上資料庫_飛資得資訊	http://www.lib.ccu.edu.tw/cgi-bin/flyweb/flyweb.cgi?o=1	online	✓		
全國科技資訊網路 (STICNET)	http://sticnet.stic.gov.tw/	online	✓	✓	科技分類典、人文社會分類總表、科技索引典
博碩士論文資料庫	http://192.83.176.196/theses/index.html	online	✓		
政府研究計畫成果	http://192.83.176.196/grb/index.htm	online	✓		

大陸最新期刊目次	http://192.83.176.196/ccChina/index.htm	online	✓		
學術研究機構名錄	http://192.83.176.196/InstDir/index.htm	online	✓		
電子詞典－土木建築類	http://192.83.176.196/eDic/eDic_Query.htm	online	✓		
Internet快訊	http://www.stic.gov.tw/stic/1/107.htm	online	✓		
台灣常見魚類資料庫	http://192.83.176.196/fish/fish.html	online cdrom	✓		
台灣常見植物資料庫	http://192.83.176.196/plant/plant.html	online cdrom	✓		
中華民國年鑑(八十五年版)	http://www.gio.gov.tw:7008/COVER0.HTM	online	✓		