

時間的限制及熱烈的討論聲中劃下句點。

後記：

本次研討會能圓滿落幕，我們特別要感謝中國圖書館學會醫學圖書館委員會的支援、飛資得資訊公司的鼎力協助和本館同仁的配合，以及所有來賓的熱情參與！這次的演講帶給我們新的觀念、重新省思圖書館資訊服務的新方向與努力的目標，期盼相關單位多多支援圖書館舉辦類似的研討會，使來自全國各地的館員能彼此交換與分享經驗，同時掌握整個大環境的發展趨勢，進而與世界潮流並駕齊驅！

醫圖教師著作資料庫與總圖特藏組 台大教職員著作目錄光碟系統之比較

黃婷鈺

一、前言

針對特藏組於台大五十年校慶時所展示之台大教職員著作目錄光碟，醫圖收到不少讀者對其與本館教師著作資料庫之異同有所疑義，故於此簡單作一比較說明。

二、資料庫內容之比較

1. 資料庫來源：

醫圖：台大醫學院區專兼任教授及醫事同仁之著作論文抽印本。

特藏組：(1)台大教職員之著作目錄清單。

(2)未提供著作目錄清單者，則就圖書館已收錄入藏之資料予以增補。

2. 資料庫收錄標準：

醫圖：(1)以英文發表之期刊論文書目及摘要資料。

(2)非以英文發表之期刊論文書目及摘要資料，惟需附英文標題及摘要。

(3)以英文發表於期刊以外之學術著作（如單行本、叢書之一章），而其內容包含作者自行研究之成果或作者本身之觀點。

(4)國際性學術會議發表論文之全文記錄特刊（proceedings）。

特藏組：凡本校教職員之著作未曾收錄於民國六十四年「慶祝本校三十週年教職員著作目錄」、民國六十九年「慶祝本校三十五週年教職員著作目錄」（續編）及民國七十四年「慶祝本校四十週年教職員著作目錄」（第三編）者，為資料庫收錄範圍。

3. 資料庫涵蓋年代：

醫 圖：以西元 1989--1994 年資料為主。

特藏組：以西元 1985--1994 年資料為主，若有『國立台灣大學教職員著作目錄』第一至三編未及收錄之著作，亦一併著錄。

4. 資料庫收錄之文獻類型：

醫 圖：原始論著、案例報告、會議報告、研究簡報、綜合評論、書籍。

特藏組：期刊／報紙論文、書中文章、專著、會議論文、國科會計劃、研究報告、學位論文、評論、摘要、專利及其他。

5. 資料庫收錄筆數：

醫 圖：目前收錄筆數 4,517 筆（未完成統計重複筆數）

特藏組：有關醫學院區（醫學院及附設醫院）教職員著作筆數計 11,468 筆。
(未完成統計重複筆數，且尚有部份著作持續補鍵中)

6. 資料庫欄位之比較

<醫圖>		<特藏組>			
欄 位	名 称	檢索功能	欄 位	名 称	檢 索 功 能
第一作者	中文名	全文式	AU	作者	全文、片語式
第一作者	英文名	全文式			
最後作者	中文名	全文式			
最後作者	英文名	全文式			
其他作者	中文名	全文式			
其他作者	英文名	全文式			
-----	-----	-----	-----	-----	-----
第一作者	單位名	全文式	UD	所屬單位	片語式
-----	-----	-----	-----	-----	-----
中文篇名		全文式	TI	篇名 / 書名	片語式
英文篇名		全文式			
-----	-----	-----	-----	-----	-----
期刊名稱		全文式	SO	出處	全文、片語式
卷冊	*		CS	會議名稱	全文、片語式
期號	*		CL	會議地點	全文、片語式
補充期號	*		CY	會議日期	限制檢索
起迄頁數	*		PM	學位論文	全文式
發表年代	限定檢索		PD	出版日期	限定檢索
-----	-----	-----	-----	-----	-----
			CN	叢書項	全文、片語式
			PL	出版地	全文、片語式
			PR	出版者	全文、片語式
			OT	其他	僅供顯示
-----	-----	-----	-----	-----	-----
國科會編號		全文式	NT	附註	全文式

語言	限定檢索	LA 語文	片語式
著作性質	限定檢索	TY 著作性質	片語式
控制編號	全文式		
關鍵字	全文式		
摘要	全文式		
得分C	限定檢索		
得分J	限定檢索		
得分A	限定檢索		
總分	限定檢索		
參考篇數	僅供顯示		

說明 1：上述『全文式檢索』意指資料庫該欄位的檢索方式可以自由詞 (free term) 的方式表達，即資料庫中的每一個字皆可檢索。

說明 2：『片語式檢索』，全文檢索以字為索引值，而片語檢索則以整個片語為索引值，可節省大量資料比對的時間，提高檢索效率，同時亦可減少誤引的機會，提昇檢索值正確率，惟檢索時較不具彈性。

說明 3：『限定檢索』指該欄位的儲存值可用以限定資料筆數，提高檢索之精準率，亦可單獨為一檢索策略。

說明 4：醫圖與校總區雖皆提供全文式檢索，但達到的功能及使用的界面各異，本文稍後將對資料庫檢索功能作較詳盡之比較。

(1)作者部份： 醫圖組因考慮教師著作升等評估之需要，故在著作部份，區分為第一作者、其他作者、最後作者三等級，以應貢獻程度不同而分數不同之需求。此外另建作者姓名權威檔，以確保資料庫紀錄之完整，並對照作者中英文姓名，以助資料辨識及檢索。

特藏組因資料來源為著作清單，故在資料著錄的完整程度上，有其實務上的困難。且其目的乃在記錄台大教職員之出版概況，故無區分作者層次的必要性。

(2)所屬單位： 醫圖與特藏組均設有欄位以著錄發表論文之單位別，以利統計資料之收集，並得以概觀論文之發表狀況。惟二者在著錄的判定上，標準各異。醫圖的作法是以論文抽印本上所載之第一作者單位為準，若第一作者非屬台大醫學院或附設醫院教職員，則以第二作者所屬單位記錄之，以此類推。優點是區別論文發表的主要貢獻單位、重複資料易於查核。但協同著作的單位會被忽略，相對的檢索點也會較少。

特藏組則除依據清單所載單位著錄之外，並參考「國立台灣大學教職員錄」(民國 83 年版)，目的在將每一篇著作之分屬單位列明。優點是檢索點增多、考慮協同著作單位的貢獻，缺點是同一資料因分散於各單位，查核複本不易、且不區分著作發表之主要貢獻單位。

(3)題名部份： 醫圖原擬收之資料需附有英文摘要，故分中、英文篇名。惟實際運用上，因醫圖不嚴格篩選不符收錄範圍之資料，故在資料的著錄

上，共計有二個欄位用以著錄不同語文的篇名（並不以英文發表之文獻為限）。特藏組亦採相同的作法，惟欄位名稱有所差異。

- (4)資料出處： 醫圖之期刊名稱一欄，著錄文獻之來源，即包含：書名（如文獻為某書之一章）、刊名、會議名稱（含屆次、舉辦日期及地點）、碩博士論文（含年度、學校及系所）；特藏組的作法是區別各資料類型來源，故有多個欄位。換言之，特藏組在資料來源的定義較為嚴謹。

醫圖為便區分複本資料，故將文獻來源出處之卷期、頁碼區分為四個欄位。而特藏組為避免單元欄位過於冗長及求畫面精簡，故以單一欄位定義之，但在資料的檢索及輔助複本查核上，會存有若干問題。（需視鍵檔品質，如：區隔符號、標點、格式）。另有關文獻發表的日期，基於醫圖收錄文獻以期刊為主，故僅載明出版年度，而特藏組因亦收取報紙文獻，故出版日期詳記日月年。

- (5)出版及叢書項： 醫圖以收錄期刊文獻為主，故有關出版地、出版者及叢書項均無記載。而特藏組收錄的文獻類型較廣，故需載明出版事項及叢書項以助資料之辨識。

- (6)附註說明： 二單位皆設有欄位記載研究計劃編號或專利號碼，惟欄位名稱有異。醫圖以國科會編號一欄著錄相關資訊，特藏組則以附註事項記錄之。此外，對於各種未能記錄於適當欄位之書目紀錄，特藏組亦一併於附註項中記載。

- (7)語文標示： 二單位均設有欄位標示文獻發表語文，但欄名亦各異。欄位名稱請參見上表。

- (8)著作性質： 二單位均設有該欄位，然文獻性質的認定及分類程度不同。醫圖以期刊文獻為主，故著作性質依論文內容分。而特藏組則依不同出版形式來區分資料類型。（分類項目參考本文二之4）

- (9)其他說明： 特藏組備有『OT 其他』欄位，用以記載：館藏代碼、館藏註記及登錄號，以方便讀者取得原件資料，惟此一欄位僅供顯示，不提供檢索。

醫圖因自1994年起乃收論文抽印本，故除少部份資料不全外，均可依控制編號向館方索借原件。此外醫圖另可以關鍵字、摘要、論文或作者等級評介分數等為檢索途徑。尤其評介分數，可以藉此作為判斷論文價值、統計某學科或某特定作者的學術成就的參考。

三、系統功能之比較

1. 索引概念

醫圖所採用的介面是自然語言索引系統。即資料庫中的每一個單字、片語皆可用以為檢索詞。其優點是：索引詞可以較專深、使用者可自由表達所欲檢索的概念。但相對的，檢索結果不易完整。因為系統使用者較難判定論文撰寫者所採用的語詞(term)為何，故需應付同一概念不同表達方式（同義詞）及同形異義字所帶來的問題，對使用者而言，負擔較重。此外，自然語言索引系統乃將使用者所採用的term與資料庫作比對，故有時會有誤引的狀況發生。例如：當系統使用者欲找尋有關『化工』方面的資料，若直接以「化工」為檢索詞，會有『自動「化工」廠』等資料的出現，

而實際與讀者的檢索需求無關。

基本上，特藏組所採用的亦屬自然語言索引系統。但為求解決部份類似上述『化工』例子所造成的誤引，特藏組亦兼採片語式檢索方式。即將文獻主題區段儲存，如：篇名為『Technique of dynamic programmilumped-parameter system : optimization theory...』則將文獻主題分段儲存為：technique of dynamic programmilumpedparameter system; optimization theory; ...; ... 利用這樣的技術，可以減少資料誤引的機會。如：檢索者以system optimization 為檢索詞，若採自然語言式直接比對的方式，則本篇文獻會被錯誤引出；但如採片語式檢索，因system 和 optimization 分屬不同區段，故系統會去除此一文獻，提昇系統的檢索命中率。然採用片語式檢索，存在著文獻主題如何截取的問題。例如主題為『奧運金牌選手』，應區隔為「奧運金牌」、「選手」或「奧運」、「金牌選手」或「奧運」、「金牌」、「選手」或「奧運金牌選手」？區段過長，失去區隔的意義；若區段過短，相對地檢索者所能採用的詞語，彈性會較小。尤其以人工製作索引，區段取長或取短將視索引者觀點而異，控制詞彙的功能並不彰顯。

整體而言，使用自然語言索引，輔以片語式檢索仍是目前較佳的作法，尤其國內一直缺乏定義較嚴謹的索引典，自然在檢索正確率方面無法獲得較佳的系統支援。此外，特藏組所採行的系統，可提供跨欄位檢索，而醫圖所採用的系統則必須以布林邏輯作聯集方能達到相同的檢索效果，故就此點而言特藏組所採用的系統介面較為 POWERFUL 。

2. 布林邏輯

二單位所採用的系統均提供布林邏輯運算，以擴大或縮小、限定檢索範圍，使檢索策略更具彈性。

3. 切截

二單位均提供向後切截的功能，以便讀者擴大檢索。尤其英文文獻，同一字首、字根的字，通常意味該組字群具有某種程度的相關關係，具切截的功能可幫助檢索者完整檢索及解決部份單複數形或詞性變化所產生的檢索問題。

4. 相近運算

在使用自然語言索引系統時，會有誤引的情況發生，常是因為所下的檢索值過於鬆散的緣故。舉例來說，當讀者欲找出有關『台北污染問題』方面的資料，若以聯集概念『台北 and 污染』的檢索策略，固然可以找出符合檢索需求的文獻，但也會產生一些諸如討論『台北的交通及高雄的污染』等之誤引資料，尤其是在摘要部份作 free term 檢索，若不限定二概念詞之相對位置（如：前後或字距），很難對檢索主題作控制。為因應這類問題，或有系統以相近運算的方式來限定二概念詞的前後位置及字距，以提高檢索的正確率。如特藏組所採用的系統即具備相近運算的功能。惟讀者在使用相近運算元時，必須審慎考慮字距值。字距過大，通常也意味著二概念詞的關係不密切，但字距過小，可能會錯失因文法表達方式或倒置標題所造成字距較遠的文獻。

5. 線上索引詞展示

醫圖採行系統以自然索引為主，故無所謂的索引詞展示。而特藏組所

採用的系統因輔以片語式檢索，故另可以選單的方式檢索。但凡具片語檢索的欄位均可線上瀏覽索引詞，並可於瀏覽的同時，指定索引詞作檢索，對檢索者而言，應是便利的作法。

6. 連結檢索

特藏組所採行系統允許使用者於瀏覽檢索結果時，如認定有相關的索引詞可資採用，則可以 mark 的方式，在文獻的某一單字或片語上標記，轉入檢索。

7. 列印與儲存

二單位均提供頗為彈性的書目資料儲存與檢索功能。使用者可指定所欲儲存的欄位，特藏組採行系統更可依使用者需求進行儲存資料的排序。另外，二單位系統也都提供檢索策略的儲存及執行功能，使用者甚至可以一般 DOS 下的文書處理軟體編寫檢索策略，惟副檔名需使用該系統之指定名稱。

8. 跨資料庫檢索

二單位採行系統均為套裝軟體 (package)，故只需定義資料庫，即可加掛於同一系統，俾可使用同一檢索介面來檢索不同的資料庫。

四、結語：關於『比較』

醫圖在 82 年開發教師著作資料庫時，原先目標乃在整合教師升等評估、研究論文報告之出版及記錄醫學校區教職員論文發表概況等三項工作計劃。今年為慶祝台大五十週年校慶，由總圖負責規劃全校教職員出版品目錄光碟系統，做為『台大教職員著作目錄第四編』。本文旨在區別二性質相近之資料庫在內容收錄及系統功能上的差異，俾備醫圖在評估是否繼續維護本館斥資開發之資料庫及將來全校整合教師著作資料之方向的參考。

評鑑資料庫的方式很多，但一來由館方單方面評估易流於主觀，二來因為醫圖開放教師著作資料庫檢索為時尚短，使用者意見的觀察及統計未能客觀反應真實檢索狀況，且特藏組之資料庫系統尚未開放供眾使用。所以很多方面，不宜也不能在此一併考量。特別是在系統的輔助功能、使用者介面及系統的穩定性、檢索效率及書目品質、資料庫維護等方面本文未能論及，相當遺憾！

資料庫是否能提供使用者作為研究之參考，其主要關鍵在於資料庫是否易於維護及資料庫管理者有無能力長期維護。我們欣見特藏組順利完成資料庫系統的開發工作，但顯然未來的路還很長，在資料庫性質相近的前提下，整合不失為一項避免人力重複投注的方案，但也如同開發一個新的資料庫一樣，這不會是一項輕鬆容易的工作。然而值得慶幸的是，此一問題，已然彰顯，相信最終的決定必定是以讀者需求為依歸，請大家拭目以待。

*附註說明：特藏組所稱『教職員』意指根據「國立台灣大學教職員錄」83 年 11 月版所列教職員為主，文中不便一一作注，在此一併說明。

特此銘謝：總圖特藏組提供系統參考資料