

# 中國大陸引用文獻索引 建置機構之參訪

陳光華 梁瓊方

## 摘要

本文介紹中國大陸目前從事引文索引資料庫的建置工作，並說明相關機構的選刊原則、資料建置流程及模式、資料校正模式、資料庫系統、面臨的困難及問題、以及管理的模式。期望藉此作為建置臺灣引用文獻資料庫的參考。

## 一、源起

學術發展的歷史是各學科或研究的不斷創新所累積的，多數的研究與創新都是建立在前人研究成果的基礎之上，資訊學者將引用前人文獻和經驗的行為，稱為「Citation」，一般譯為「引用文獻」或「引文」。學術論文的作者對於引用文獻的列舉，是一種尊重他人學術成果的表示，同時，亦可作為讀者進一步研究的指引。引用文獻是資訊被使用過後所留下的標示與紀錄，透過這些資料的分析研究，可以了解研究者使用文獻的情況。藉由引用文獻的分析，我們可以了解某一學科領域的發展現況、文獻使用特性、文獻間的關聯性，以及未來研究趨勢。因此，引用文獻索引資料庫對於學術研究的發展扮演相當重要的角色。

---

關鍵詞 (Keywords) : 引用文獻索引 ; 中文社會科學引文索引 ; 中國科學引文數據庫  
Citation Index ; CSSCI ; CSCD

陳光華 : 國科會人文學研究中心研究員、國立臺灣大學圖書資訊學系副教授 ; E-mail: khchen@ntu.edu.tw

梁瓊方 : 國科會人文學研究中心研究助理 ; E-mail: r91126012@ms91.ntu.edu.tw

臺灣圖書資訊學界很早即重視資料庫的建置工作，且建置相當多的資料庫，如：中文圖書資訊學文獻摘要資料庫 (CLISA)、中文報紙論文索引 (ICN)。然而，在 1999 年以前，臺灣除了一個科技期刊引用文獻索引資料庫的試行計劃外，並無針對其他學科所做的引用文獻索引資料庫。然而，引用文獻索引資料庫的功用不僅可提供使用者查詢來源文獻與被引文獻間所串起的關係，亦可透過引用文獻資料庫瞭解某一學科主題的研究成果及學科間的關係，此外更有篇名、作者、主題等的相關資料，以提供一般資訊檢索的功能。<sup>[1]</sup>有鑑於引用文獻資料庫的重要性，而目前臺灣也沒有人文學引用文獻索引資料庫，因此國科會在 1999 年成立人文學研究中心 (Humanities Research Center) 之後，隨之展開臺灣人文學引用文獻資料庫的建置工作，期盼累積一定的資料量後，能夠提供各界使用，從事基本資訊檢索與引文分析的相關檢索以及學術研究的工作。

在資料庫真正開始進行建檔工作之前，有許多的前置工作，如：期刊的選擇、資料庫的設計、資料庫建置的政策及規範等，當然在此同時也遇到了一些困難。至於開始進行資料庫建檔工作之後，同樣的，也遇到了許多問題，如：期刊的來源問題，實際建檔後與當初所訂定的建檔規範有許多出入及期刊文章中引用文格式的不同增加判斷是否實際引用的困難等。因此，我們開始思考是否需要訪問中國大陸從事引用文獻資料庫的學者專家，實際觀察並研究他們建置資料庫的情形，相互交流以獲取前人的經驗。中國大陸從事引用文獻索引資料庫的建置工作，已有一段時日。例如：南京大學成立中國社會科學評價中心，從事社會科學引文索引的工作；中國科學院文獻情報中心則是進行科學引文索引資料庫的建置工作；中國科技信息研究所則是著重於工程技術的引文索引資料庫的建置工作。因而，我們決定訪問上述三個單位，並與實際從事資料庫建置工作的學者專家們座談。以下本文就針對此次訪問後的心得，依照訪問的行程，作一個簡單的報告並提出建議。

## 二、南京大學中文社會科學引文索引

在 60 年代美國費城科學資訊研究所 (Institute for Scientific Information, 簡稱 ISI) 相繼研發《科學引文索引》(Science Citation Index, 簡稱 SCI)、社會科學

[1] 陳光華，《「臺灣人文學引用文獻資料庫」建立之研究》。行政院國家科學委員會人文學研究中心研究計畫成果報告，2000 年。

引文索引 (*Social Science Citation Index*, 簡稱 SSCI)、藝術與人文學引用文獻資料庫 (*Arts & Humanities Citation Index*, 簡稱 A&HCI) 等專業的引用文獻索引資料庫, 現在已成為國際間公認的一個對科學成果 (包括對於個人及機構的研究) 進行客觀評估的一個權威工具。在中國大陸, 中國科學院文獻情報中心針對科學方面的情報訊息研發了《中國科學引文索引》(*Chinese Science Citation Database*, 簡稱 CSCD), 但在社會科學方面卻無此一相關的評價機構, 且由於社會科學方面的研究在整個科學範疇、社會經濟及科學技術發展中的作用越來越重要, 因此對於社會科學研究進行一全面的評價研究實為必要。有鑑於此, 南京大學決定於 1997 年成立「南京大學中國社會科學研究評價中心」, 由鄒志仁、楊克義、蘇新寧、吳向東、袁培國及華薇娜開發《中文社會科學引文索引》。1998 年 5 月, 南京大學邀請香港科技大學共同參與《中文社會科學引文索引》, 於 1999 年 4 月 23 日兩校校長共同訂定正式協議。1999 年 8 月 20 日, 中國大陸教育部正式行文南京大學, 批准《中文社會科學引文索引》(*Chinese Social Science Citation Index*, 簡稱 CSSCI) 為中國大陸教育部人文社會科學研究「九五」規劃重大項目。<sup>[2]</sup>

#### (一) 來源期刊

CSSCI 最初以北京大學所出版的《中文核心期刊要目總覽》羅列的 666 種期刊為基礎, 但由於考量的觀點不同, 因此在期刊的選擇上最後根據 (1) 期刊必須要有 ISSN 標準號碼; (2) 不收錄二次文獻; (3) 期刊無列出引用文獻者不收錄; (4) 擴大原本社會科學之範圍, 另包括了軍事、地理學、環境科學等類等四個原則, 經由知名專家學者經嚴格的標準評價選定了 505 種作為來源期刊。而來源期刊的取得則可分為三個時期說明: 初期, 向圖書館外借相關的期刊建檔, 但因期刊為借閱必須加以保護, 因此在建檔資料方面增加了一些困難, 如無法清楚地將引文標示出, 必須另行抄寫, 增加工作人員的困難; 另一方面因為期刊主要是提供讀者使用, 而借閱也會造成流通的困難。然而, 為避免造成圖書館的作業困難, 中期之後, 改由社會科學研究評價中心自行購買, 每年用於期刊購買的費用約為兩萬元人民幣。由於社會科學研究評價中心於 1999 年發行了《1998 年中國社會科學引文索引》數據庫的光碟版, 許多期刊發行單位知道了此

<sup>[2]</sup> [http://202.119.47.137/cssci\\_jianjie.htm](http://202.119.47.137/cssci_jianjie.htm)

項研究的重要性，因此，現況是各期刊發行單位願意主動提供期刊，如此一來省下不少費用。

目前中國大陸期刊發行總量為 3,500 多種，扣除雜誌等非學術性期刊後，餘 2,800 多種學術性期刊。而 CSSCI 現在所收錄期刊為全國所發行之學術性期刊的 15%，將來預計再增加期刊之數量為目前所收錄之 5%。此外，對於所收錄之期刊發予證書，另外不將之稱為「核心期刊」而以「來源期刊」表示。根據 1998 年 CSSCI 光碟所收錄之期刊，數量共計 497 種，詳細的資料請參見表一。<sup>[3]</sup> 1999 年增為 506 種，2000 年減少為 419 種。

表一 1998 年 CSSCI 光碟收錄期刊

類 別	種數	類 別	種數
馬克思主義	14	民族學	18
中國文學類	17	藝術類	20
外國文學類	6	語言學類	28
管理學 含科學	16	經濟學	73
社會學 含人口學	7	歷史學、考古學類	27
哲學 含宗教	13	政治學 含軍事類	25
心理學	7	國際問題研究	14
綜合性社會科學	60	法學類	17
統計學	4	新聞學與傳播學類 含文化研究	16
教育學類	20	圖書、情報、文獻學類 含檔案學	25
地理學、環境科學	14	綜合性高校學報	47
體育類	9		

## (二) 工作流程

一筆資料的完成而後得以提供使用者查詢，需經過標引、審校、著錄、文字校對、修改、二校、總校等七項工作<sup>[4]</sup>，下面則針對著錄、標引及校對人員之來

<sup>[3]</sup> 參見 [http://www.cssci.com.cn/cssci/c\\_lyqk98.htm](http://www.cssci.com.cn/cssci/c_lyqk98.htm)，網頁宣稱共有 496 種，經實際計算應為 497 種。

<sup>[4]</sup> 本文於說明特定機構時，使用該機構的術語，若不牽涉特定機構或產品，則使用臺灣的術

源及工作內容做一簡單的說明。

### 1. 標引

目前有 20 位工作同仁負責標引的工作，工作同仁是經由一般徵才方式招募後，再接受嚴謹的訓練才正式錄用。工作內容則包括：

- (1) 關鍵詞：文章中所提供的關鍵詞並不一定準確，必須提供額外標引出的關鍵詞 1-2 個不等，以供查詢。若文章無關鍵詞時，則標引人員需檢示文章內容後給予適當的關鍵詞，而關鍵詞必須是有實際意義、具有檢索意義的，切勿濫用。關鍵詞包括文章作者所提供的及標引人員所給的至多只能有五個。各關鍵詞間以「 / 」分隔。
- (2) 學科分類：學科代碼根據大陸教育部的 26 種分類為主。遇有兩個以上的學科則以至多分為兩種，而學科代碼亦自動產生兩個，也是使用「/」將其分隔。
- (3) 建檔資料的書目劃分：將期刊資料的引用文章書目依各著錄格式標示出，以利建檔人員進行資料的著錄。
- (4) 對於已建檔完成之資料進行文字校對。

### 2. 著錄

人員的選擇原則上並無特殊要求，而在正式建檔前會提供訓練課程，以提高其對資料的熟悉度。目前建檔人員由委託南京大學圖書館提供及專職建檔人員共 7 位。工作內容主要為基本的資料建檔及在標引人員文字校對後，再次進行修改。

### 3. 審校、二校、總校

校對人員均為與圖書館學系相關的人員，目前由教師及研究生來執行此項工作，共約 10 人。工作內容因校對的層次不同而有所區分：

- (1) 審校：即針對標引人員所標出的各項著錄格式做一校對工作，確認後才由建檔人員進行著錄工作。
- (2) 二校：指在建檔人員進行修改後，校對人員於資料庫建檔介面直接進行校對修改。
- (3) 總校：於二校後，將所有二校過的資料集合起來於存入資料庫前再校對

---

語。如「標引」可理解為以人工的方式在文件中標示出輸入人員應輸入的資料；談論特定資料庫使用其產品名稱，如「中國引文索引數據庫」，一般名詞則使用資料庫。

一次。被引文獻出現於不同來源文獻之狀況，於總校時將其合併為一筆資料，以避免重覆相同的被引文獻而佔據相當的資料庫容量。外文資料較易出錯，因此對於外國的資料於總校時會特別的注意。

### (三) 來源文獻

來源文獻為資料建置的重要資源，亦決定了資料庫的權威性，「中文社會科學引文索引」來源文獻的收錄類型與收錄原則說明如下：

1. 收錄類型：學術論文、綜述（包括會議綜述、述評等）評論（包括書評）報告（包括領導人的重要講話、調查報告等）傳記資料及學術水平較高的其他文獻（如：筆談、筆會、座談會筆錄、訪談錄等）
2. 收錄原則：當某期刊全年各期的文章均無引用文獻時，則暫不收錄此期刊，留待觀察是否仍有引用價值時再予以收錄。此外，若一期刊僅一篇文獻無引文時，則照錄。

### (四) 被引文獻

被引文獻的收錄情形，影響未來引文索引資料庫的使用模式，例如：僅收錄中文被引文獻，便無法進行完整的引用文獻分析的研究，在資料庫發展初期，便應訂出詳細的政策。<sup>[5]</sup> 以下說明「中文社會科學引文索引」的被引文獻。

1. 收錄類型：期刊、圖書、報紙、會議論文、匯編（叢書、文集）學位論文、信件、報告、法規或司法文書、電子文獻、古籍、其他（檔案、專利）標準（如：ISO）
2. 收錄之語言種類：漢文、譯文（中譯文為主）英、德、日、俄、法、西文（凡非英法德文之西方語言均屬此）及其他語種（少數民族語言）
3. 被引用之形式：參考文獻、文內注、腳注、文後注、文中引（即文中所引用的其他文獻，但又不屬於文章的注或引的範圍）
4. 著錄原則：原則的制定是由南京地區及南京大學相關學者專家共同討論而成，經過12次的修改，於2000年3月6日再次修訂完成。除客觀的著錄引文的資料外，有些地方仍需經由人工介入進行規範，以達到資料的統一完整性。以下就針對幾項較重要的著錄原則做一說明：

(1) 一般至少要有「著者」、「篇名（或書名）」兩者。若只有其中一項時則

<sup>[5]</sup> 中國社會科學評價中心編，「CSSCI簡易標引細則」。

需另有一輔助項目，如：出版年、出版者、卷期等才可予以收錄，如此才可提供讀者一查詢的基本資料。

- (2) 外國的譯音依據中國大陸國家統一標準。
- (3) 打不出來的字由兩個問號(??)代替，此外亦對此做一統計數據。
- (4) 被引文獻有多個著者時，只標引第一個作者。而外文著者名應按姓在前名在後之原則著錄。
- (5) 外文的著錄，俄文、日文均可建檔，而其他西文則以作者、篇名、出版者為主，其他不必要之資料則不著錄，因對引文並不會有太大的影響。
- (6) 當一篇文章無引文時，標引人員檢視全文找出書中引用的書目，最多列舉兩個。
- (7) 同一篇文章，若引用的頁碼不同時則只著錄一次資料。

#### (五) 被引文獻可能產生之錯誤

建置資料時會發生各種型式的錯誤，也應是極力避免的，根據CSSCI的工作經驗，最常發生的錯誤為：<sup>[6]</sup>

1. 機構：將部門名稱誤認為機構名稱著錄，如：將「南京大學管理學院」當機構來著錄，而機構應為「南京大學」，而「管理學院」為部門。
2. 作者：在被引文獻中因著者的錯誤或印刷的錯誤，而造成被引作者的錯誤著錄，如：「黃宗忠」應為「黃忠宗」。遇此問題若建檔者能清楚確定錯誤時則即時改正。
3. 篇名：同一篇文章被引用時標題不一，而造成許多不同的標題，但卻是指向同一篇文章。

#### (六) 資料建檔之現況

根據1998年的統計來源文獻共6萬筆，被引文獻約28萬筆，其中每一期刊平均每年約112篇，而每一篇平均有4.8篇的引文。以1998年的計算，每年花費於資料建檔的經費約30萬，包括了標引、建檔及技術人員。而1999年之期刊於今年5月才開始進行建檔工作，扣除暑假休息一個月，預計於2000年11月底將1999年度之期刊4,000多冊建檔完成。此外，依據1998年建檔完成之期刊，可得到有關期刊的多種統計數據：(1) 作者發文統計：51,376篇；(2) 發文機

6 蘇新寧，「引文索引數據庫質量控制研究」，未刊稿。

構：8,254 個；(3) 作者被引次數：90,183 次；(4) 被引文獻 54,141 篇；(5) 被引期刊 12,864 種。

#### (七) 資料建檔工作所遇的困難

1. 規範的制定：因為每一期刊編者有其編排的原則，因此對於各種不同的引文方式要給予一個完整周全的建檔規範需花一些心力，加上經驗的累積一再修改才能制定完善的規範。
2. 規範的執行：因標引人員、建檔人員及校對人員之學科背景、學習程度不同，對於規範的熟悉度亦有不同，因此需花一些時間訓練才能順利的執行規範。
3. 標引人員的素質及培訓問題：標引的工作是分析資料的第一步驟，因此標引人員的素質也相對重要，如何培訓一個有素質對於資料各類型相當熟悉的標引人員，是一個極重要且困難的工作。
4. 人員的流動率：工作人員可能因各項因素，如：遇考試期間或者其他狀況而不願再繼續工作，如此一來還得再重新培訓一批新的人員，對於工作團隊的負擔更為加大。

#### (八) 光碟資料庫

CSSCI 光碟版的發行希望可以為廣大的人文社會科學領域的研究者提供一重要的資料查詢工具，因此為將光碟產品的製作與資料庫的建檔工作分開，於 2000 年 5 月成立了「江蘇南大社科信息諮詢有限公司」，以專門負責光碟資料庫的銷售。成立此公司所花費的人事、設備、光碟製作等經費為一百萬元，期望往後可透過公司的收入來提供資料庫建檔的經費，如此一來即可成為一自給自足的評價中心，另一方面也期望 CSSCI 系統可以提供更多種的評價統計訊息，以滿足各領域學科專家學者的需求。

### 三、中國科學院文獻情報中心

「中國科學引文數據庫」是由中國科學院文獻情報中心承建的一個重要的文獻資料庫，也是目前可以提供一實際服務的引文資料庫。「中國科學引文數據庫」得以開始進行建置的工作，乃是在 1989 年得到國家自然科學基金委員會和中國科學院的資助，在此之後 CSCD 的工作才正式開始。

#### (一) 來源期刊

CSCD 自 1989 年成立至 1993 年之間，收錄之期刊總數為 315 種，乃是透過初步的統計結果及諮詢專家的基礎而來，1995 年則增為 582 種，乃是對引文資料庫於 1989-1995 年間的引文資料進行期刊被引用次數統計，以此為基礎再扣除一些學術性與基礎理論性較差的期刊；同時亦參考了國內外主要的學術性期刊。然而期刊選擇的指標仍以期刊被引用的次數的為主要的依據。1996 年出版的光碟資料庫所收錄之期刊數量已擴展至 582 種，於 1999 年與北京清華大學「全文數據庫」合作後期刊再擴至 1300 多種，但是目前仍以 582 種為核心期刊。至於期刊的取得，早期因為中科院有自己的圖書館，因此許多期刊乃從圖書館借閱而得，有些則向期刊的出版單位請求贈閱或購買，而隨著光碟資料庫的發行及各界漸漸瞭解此項工作對其單位的助益後，則爭相請求加入，亦願意主動提供期刊以利建檔工作的進行。在 1999 年與北京清華大學全文數據庫合作後，雖然資料是由北京清華大學提供的資料欄位直接進行轉換，但由於仍會有錯誤的資料，因此目前仍需要紙本的期刊以供校對，而早期已經建檔校對完畢的期刊資料，因無空間存放，所以有些提供給圖書館補齊缺期的期刊，有些則贈送給前來索取的單位。隨著光碟資料庫的出版，有些期刊出版單位前來請求收錄，因此針對來源期刊的收錄是否增加是必須考量的因素。一方面，因出版光碟資料庫打響了名聲，因此各大院校對於其所發表的論文評價逐漸重視，所以增加期刊量是為滿足各大院校所需的評價而做；另一方面，期刊的數據統計量仍以近 600 多種來進行統計分析為佳。但除考慮期刊收錄量的增加外，對於淘汰的部分因涉及有些資料的零散，因此到目前為止並未進行任何淘汰的動作。

## (二) 人員的招募及工作流程

目前此項資料庫的工作由 11 個人負責，工作同仁並非經由特殊的遴選或是招聘而來，乃是經過辦公室的合併而自然產生，在工作的分配上並無特別規定各項專職的工作，只有針對較為細項的工作時，才分配較為專精的人員執行。此外，因對於校對人員的要求不一，因此也會給予不同輕重的工作項目。而在工作流程方面，主要為標引、著錄及校對三項工作。

1. 標引：早期的標引工作是由工作人員在期刊上畫線標出，至 1999 年後與北京清華大學的「全文數據庫」合作後，人工索引的工作即已暫停。
2. 著錄：早期將資料著錄交由外包公司承擔，每年花費近 30 萬元，現在由於與北京清華大學的「全文數據庫」合作後，則直接進行資料的轉換。

3. 校對：早期至少要校對四至五次，如：列印校對、電腦校對、合併校對，與北京清華大學合作後雖減少許多校對的工作，但因北京清華大學提供的資料欄位並非全為 CSCD 所需，加上有些文字的錯誤或譯文等仍需工作同仁花費額外的時間進行校對的工作，而後才能進行轉換的工作。此外，在統計分析時仍會再進行最後一次的資料校對工作。

### (三) 來源文獻

「中國科學引文數據庫」為中國大陸第一個建置的引文索引資料庫，其來源文獻的收錄原則影響後續建置的引文索引資料庫，說明如下：<sup>[7]</sup>

1. 收錄類型：研究論文（含基礎研究、基礎論著等）、會議紀要、研究簡報（研究報告、專題會議報告等）及綜述。
2. 收錄的項目：文章篇名、著者（最多著錄20位）、關鍵詞、學科類別、基金類別、文章出處（包括期刊的刊名、卷期、起訖頁數）等。
3. 收錄學科類型：數學、物理、化學、天文、地理、生物等自然科學為主，兼有醫藥衛生、農林化學、工程技術、航空科學及環境科學等。

### (四) 被引文獻

在引文方面，只收錄中文的引文而不收錄外文的引文，而所謂中文引文則包括以中文發表的文章、中國人在中國大陸以外地區所發表的所有文章（含中、英文）及在中國大陸所發表的文章（含外國作者發表的外文文章）。此外，因經費及技術等條件的限制，對於外國的引文，只處理外文引文的語種及出版年。語種包括了有西（所有西方語言及拉丁語系）、日、俄及其他四種。<sup>[8]</sup>

### (五) 與北京清華大學「全文數據庫」之合作

北京清華大學「全文數據庫」實際上已由北京清華大學成立公司自行營運並與北京清華大學圖書館合作，此中心單位名稱為「中國學術期刊光盤中心」屬於北京清華大學下的一個中心單位。

其所收錄的期刊乃針對大陸所出版的科技、社會科學等期刊進行收集，目前大陸每年出版 3,000 多種期刊，但其收錄之期刊目前僅佔全國出版量之一半，而為使所收集的期刊更完整，目前北京清華大學計劃進行全面性的收集，只是遇到

[7] 孟連生，中國科學引文數據庫建設的現況與發展，載於：《21世紀資訊科學與技術的發展國際學術研討會論文集》（臺北市；世界新聞傳播學院圖書資訊學系，1996），頁-。

[8] 同註7。

兩個難處：一為編輯問題，有些期刊的編排方式不符合北京清華大學的要求；另一為版權問題，因期刊全文提供北京清華大學掃描後，北京清華大學出版的光碟版權屬北京清華大學所有，而有些出版社希望保留此項權力因此不願意提供。

CSCD 與北京清華大學的合作於 1999 年後半年才正式開始，而由於北京清華大學亦與中國社會科學院合作，為滿足兩個合作單位的不同要求，因此在提供的資料欄位上，三方面必須要進行協調及並適度妥協，也因此有些取得的資料欄位是中國社會科學院所要求的，並非中國科學院本身的需求，所以中科院必須在資料轉換後針對取得的資料再進行校對的工作。此外，北京清華大學的「全文數據庫」早期並沒有做索引的工作，而是近期才開始進行索引的工作，因此在這一領域上仍無法確實達到中科院的要求，所以中科院在與北京清華大學合作之初便與北京清華大學作一清楚詳細的規範，以期取得的資料更為接近中科院本身的需求。而由於北京清華大學所提供的資料與中科院所收錄的期刊有所出入，並非全包含了中科院所收錄的期刊，針對這一部份的期刊，將來需再進行原始的索引建檔工作，但由於此類期刊並不多，所以現在尚未對此類資料進行索引的工作。而原本進行的索引工作也因在取得北京清華大學「全文數據庫」的資料後，全面進行資料的校對而暫停。

#### (六) WEB 的查詢功能

網路的查詢功能已經完成，提供 1991-1994 年的資料，來源資料約 13 萬條，引文資料有 45 萬條，可提供上網免費查詢。至於其他的資料，由於中國科學院文獻情報中心現在缺乏此類人才可以解決相關問題，以及收費問題仍未擬出較合理的方案，所以現在無法提供線上查詢。

## 四、中國科技信息研究所

中國科技信息研究所於 1988 年由中國工程院院長努力推動而成立的，主要工作項目為中國科學技術索引的建置工作，此外與中國科學院的「中國科學引文數據庫」在收錄的期刊方面仍有些許的差別，中國科學技術索引以收錄工程技術類的期刊為主，而中科院此類的資料則較少，當然兩者所收錄的期刊仍會有所重複。當初資料庫建置的目的是想統計中國大陸期刊發行的論文總量，當完成了部分的工作，而資料庫也累積相當豐富的資料後，資料統計分析的工作則佔相當大的比重，而中國大陸官方亦要求工作同仁統計中國大陸學者專家在國際上的論文

發表的相關資料。

### (一) 來源期刊

期刊最初的選擇依據下列三個基本原則：參考北京大學所出版的《中文核心期刊要目總覽》羅列的期刊；有無為國際檢索工具所收錄（如SCI、ERA等）；期刊被引用的頻率。而期刊的選定乃是在1988年開過多次會議後決定，之後則由中國科技信息研究所為主要的增修者，每年均會進行小幅的修訂，對於不合規定的期刊暫不收錄，待其符合要求後才再度收錄。目前所收錄的期刊量為1999年統計之資料，共計1,372種。在期刊的取得方面，最初因為中國科技信息研究所所有圖書館，因此早期有些期刊是由圖書館借閱，有的期刊則由期刊發行者贈送。而在發行光碟資料庫後，各期刊發行者瞭解此項工作的重要性，以及對期刊本身的影響性，已都自願提供期刊以利進行期刊建檔工作。

### (二) 工作內容

現有10位全職的工作人員負責此資料庫的建置工作，茲就主要工作項目做一概述。

1. 建檔：目前資料庫的建檔工作包括了標引後建檔、直接建檔及資料的直接轉換三個方式，但未對三者做過比較，不過未來仍希望以自動化為目標開發一完善的資料庫，以利建檔工作。
2. 標引：目前資料庫的建檔仍有部分是需要經由標引期刊後才進行建檔工作。
3. 校對：由於建檔工作有三個不同的方式，因此校對人員的工作也變得較為沉重，直接建檔的資料需花較多的時間審查，而資料的轉換一樣需要時間進行校對的工作。
4. 技術人員：負責的項目主要為開發統計軟體及接收其他單位的委託統計其所需的資料項目。
5. 統計報告：1989年開始發行統計報告，每一年必須花4-5個月進行資料的建檔工作，3-5個月進行統計分析的工作，才能將統計報告完成出書，此外每一年的資料均會進行小幅的修訂。

### (三) 著錄原則

最早在1988年開始建置資料庫時只是做全文資料庫，直至1994年才開始進行引用文獻的建置工作。在語言方面，因為中國大陸發行的期刊文章中，以中英文為主，因此較少遇到其他語言，但有關日文資料則不輸入亦不統計；而在作者欄

位上，早期是只建入第一作者項，直至近期才將其他作者輸入資料庫中，此外，資料庫亦提供了作者本身的資料項，如：年齡、服務單位等以提供查詢，另外也可以就此分析文章作者的年齡分布。

## 五、結 論

人文學研究中心成立的目標乃在推動國內人文學領域相關研究，目前已進行了多項研究計畫，「臺灣人文學引用文獻資料庫」之建置也是眾多研究計畫之一，為了解中國大陸的引用文獻相關工作的推展情形，故參訪南京大學中國社會科學評價中心、中國科學院文獻情報中心及中國科技信息研究所。基本上，前述機構的收錄原則各有不同，著重點亦不同。南京大學為社會科學引文索引，並涵蓋人文學相關領域。雖然，中國科學院與中國科技信息研究所接將目標放在科學、科技、與工程相關領域的期刊，但是，前者偏向科學與科技，而後者偏重於科技與工程，縱有重疊的期刊，還是有其不同的應用方向。中國科學院是由引用文獻索引出發，而中國科技信息研究所是由期刊書目資料出發，再發展為引用文獻資料庫。南京大學則因建置時期較晚，考慮較為周詳。然而，三者共同的特性是有較多人力可從事校正的工作，中國科技信息研究所甚至表示可以幫我們進行資料建置的工作。

參訪的成果讓我們對於引用文獻資料庫的建置工作有了更多的瞭解。由於人文學研究中心的「臺灣人文學引用文獻資料庫」真正進行建檔工作是在2000年的二月，因此對於資料庫的建置工作，我們可以說尚在學習階段。針對目前面臨的一些課題，提供下列建議。

### (一) 重視資料建置者的背景知識

資料建置者是影響引文索引資料庫品質重要的一環，因而，資料建置者的聘任與訓練顯得十分重要。中國大陸的建檔人員，相對的流動率較低，因為是長期工作的人員，資料錯誤率因而降低。以目前人文學中心的資料建置費用而言，僅能聘任一般的工讀生，工讀生的背景知識有限，有時無法正確判讀引文資料。未來若能獲得國家的支持，應以聘任專職人員為目標。

### (二) 訂定一致的引用文獻格式

國內人文學期刊引用的格式並沒有一定或是標準的規範，甚至同一本期刊的引用格式也不固定，造成引用文獻判讀的困難。人文學研究中心執行委員會也曾

經討論這個問題，並希望舉行說明會或討論會，推動引用格式一致化的工作。中國大陸由於引用文獻資料庫已具有相當的聲望與重要性，凡是引用格式不受規範的期刊，則不予收錄，因此，多數期刊為了能夠進入資料庫，會接受各引用文獻資料庫的規定，長此以往，引用文獻格式的問題漸漸減少。

### (三) 建構視覺化的工作環境與流程

以我們迄今的經驗而言，資料的輸入是非常重要的瓶頸，也是影響資料庫品質的重要因素。輸入的介面必須符合輸入者的直覺，同時必須儘可能降低輸入者的判讀次數，各欄位的規範與資料建置的政策必須簡單易懂，經驗豐富的諮詢者極為重要。我們也逐步修正輸入介面，但整體的視覺效果並沒有大幅度的修改，主要目的也是希望工作人員能夠繼續使用熟悉的使用者介面。

### (四) 擬定合理的期刊收錄原則

期刊的蒐集對於引用文獻資料庫的建置可說是重要的第一步。不同的資料庫建置者對於期刊的擇訂有其不同的作法與觀點。以臺灣人文學引用文獻資料庫而言，基本原則是廣泛蒐羅，因此採用比較寬鬆的標準圈選期刊；至於臺灣社會科學引用文獻資料庫則是以比較嚴謹的標準；中國大陸各機構亦有其收錄的政策。不同的作法有其背後的考量因素，我們初期以寬鬆的標準，而後分析資料庫內的資料，在未來將進一步增刪期刊。

### (五) 降低資料校正流程的成本

資料的正確性是建置資料庫非常重要的課題，然而引用文獻資料的校正卻非常困難。南京大學的資料經過五次校正的過程，仍然無法確保資料的正確性，我們根本沒有能力進行五校，因為經費過於龐大。未來將計畫撰寫一個輔助工具程式，協助資料的校正，最後僅能依賴使用者隨時回報錯誤資料，進行即時的更正。

# Interview of Research Institutes for Citation Index Construction in Mainland China

Kuang-hua Chen

Chiung-fang Liang

## Abstract

The article introduces the research institutes in Mainland China, which have been constructing citation index databases. It describes as well the principles of journal selection, the procedure and model of data construction, the model of data modification, the database system, the problems and difficulties, and the model of management. We hope this article could be a good reference for the construction of citation index databases in Taiwan.

---

**Keywords ( 關鍵詞 ):** Citation Index ; CSSCI ; CSCD

引用文獻索引 ; 中文社會科學引文索引 ; 中國科學引文數據庫

Kuang-hua Chen : Research Fellow, Humanities Research Center, The National Science Council;  
Associate Professor, Department of Library and Information Science, National  
Taiwan University ; E-mail: E-mail: khchen@ntu.edu.tw

Chiung-fang Liang : Research Assistant, Humanities Research Center, The National Science  
Council ; e-mail: r91126012@ms91.ntu.edu.tw