

國家圖書館分類 / 標題資料庫概述

—兼論規範化問題—

許令華 阮靜玲 陳友民

摘要

本文係針對本館與飛資得資訊公司共同研發建置之「國家圖書館分類 / 標題資料庫」(Window 版), 就其功能及特性、資料之整建、資料庫檢討等環節, 加以探討; 同時對於由資料庫建置及檢索所衍生之「規範化問題」, 也一併加以討論。

全文分為六部分, 除前言及結語外, 正文分四個單元敘述: 一、功能及特性: 簡述系統之瀏覽、查詢、編輯、列印、通訊、工具列等功能; 並歸納出視窗切換的靈活性、查詢方式的多樣性、查詢設計的親切性、內容文字的規範性、修訂管理的即時性等特性。二、資料之整建: 以《中國圖書分類法》、《國家圖書館增訂類目表》、《中國圖書標題表》等分編規範工具為基礎, 建置成 Excel 電子檔欄位, 這些資料均經規範化、格式化、整理、比對等處理過程。三、資料庫檢討: 提出資料庫應與書目資料庫相聯結, 以及自動帶出分類號或標題詞的功能構想。四、規範化問題: 針對內容體例再完善、檢索詞的規範化、分類主題一體化、規範修訂的參考四個課題, 提出改進的方向及方法。

關鍵詞 (Keywords): 分類; 標題; 資料庫; 規範化

Classification; Subject Heading; Database; Standardization

許令華: 國家圖書館編目組編輯; E-mail: clhhsu@msg.ncl.edu.tw

阮靜玲: 國家圖書館編目組助理編輯; E-mail: cljuan@msg.ncl.edu.tw

陳友民: 國家圖書館編目組組員; E-mail: clasub@msg.ncl.edu.tw

一、前言

文獻電子化是現代出版趨勢之一。在全球電子化浪潮衝擊、席捲下，本館為了迎接現代化的歷史浪潮，乃於民國 70 年代，開始朝電子化方向發展、邁進，先後開發製作了許多電子化文獻產品，建置了多種多樣的資料庫，以提供讀者查檢與利用，提高圖書館服務的職能。

在編目方面，先後研製了「中華民國出版圖書目錄彙編電子書光碟系統」及「中華民國出版圖書目錄光碟系統」等兩種資料庫，普獲好評。^[1] 後者資料庫，於民國 83 年更曾榮獲行政院「十大中文資訊產品創作獎」。^[2] 民國 87 年與飛資得資訊公司又合作研製「中文圖書標題表 / 中國圖書分類法資料庫」，推出 Web 版（以瀏覽器為檢索介面的資料庫），以表格式畫面加以呈現，提供於國家圖書館編目園地網頁上，供各界檢索利用。

由於 Web 版分類及標題資料庫不盡完善，本館乃於民國 88 年 11 月起與飛資得公司再度合作研製開發分類及標題資料庫。針對前一次 Web 版的缺失加以改進，並參考美國國會圖書館分類法光碟版的作法，改成 Window 版（微軟公司視窗作業系統版本，即俗稱之「視窗版」），使得分類及標題呈現的形式與紙本完全一樣；同時在功能上作了改進，除具有多種查檢方式外，在視窗間可相互切換。Window 版資料庫的製作，前後花費了一年多光景，分成三階段驗收，最後一次驗收時間為民國 89 年 9 月。全部驗收完畢後，民國 89 年 11 月展開資料庫使用教育訓練，至此可說真正開啟了標引規範工具電子化的新頁。以下即針對本資料庫之功能及特性、資料之整建、資料庫檢討等數方面加以論述。同時，對於因建置資料庫所引發之相關問題——資料及檢索規範化的問題，也一併在此加以探討。

二、功能與特性

本系統設計之初，即參酌國外類似系統的作法，以 Windows（視窗）環境為基礎，以 Java 為程式架構，以微軟公司 Excel 應用軟體將「中國圖書分類法」及

[1] 參閱：「國家圖書館資訊檢索服務」摺頁及「國家圖書館遠距圖書服務系統」摺頁。

[2] 參閱：《資訊傳真》編輯部，'94 年中文資訊產品創作獎特寫，《資訊傳真》1994：11=118，（1994.11），頁 48。

「中文圖書標題表」的每個標目逐一建檔，雖分屬兩個資料庫，但共用同一個查詢、顯示、編輯等系統功能。目前系統建置於本館的館內網路系統上，供館內需要參考分類表或標題表的同仁使用。^[3]茲將資料庫主要功能簡述如下：

(一) 瀏覽功能

為使系統更容易讓使用者接受，標目的呈現方式盡量與書本式的類表及標題表類似，故進入系統後，螢幕左方可由分類表的類號層次或標題表的筆畫順序依照樹狀結構逐層展開顯示，展開後使用者可直接點選某類號或某字起首的標題，即可顯示該類號或標題的標目及以下的說明文字。同時，兩個資料庫間可快速切換，以便靈活運用。

本資料庫之標目主表，係以不同字體、顏色或底線區別主標目、參照標目、不用標目、可超連結 (hyperlink) 的標目以及一般說明文字，讓使用者易於辨識區別。在分類資料庫主表方面，為顯示其層屬關係，將上位類號及下位類號的類名以往內縮格方式顯示，以符合一般使用者的習慣。

若以標目查詢，螢幕下方的視窗可顯示此次查詢的結果，亦即可瀏覽與查詢標目相關的類號或標題，使用者可直接點選所要的標目，避免需逐一瀏覽造成的不便。若選定一標目，螢幕上方的視窗可顯示該標目的上層路徑 (例如：查到「892 (新聞) 組織」，則上方視窗顯示 8 語文, 89 新聞學)^[4]，特別是該標目若在各學科領域都被使用時，可便於使用者判斷該標目是否確實合用。

(二) 查詢功能

在查詢資料庫的過程中，可於各自不同的情況下使用到下列查詢或搜尋方式：

1. 全文查詢：以類號查詢時，可直接查到此類號的類名或與此類號相關的標題；以標目查詢時，則以全文方式檢索包含此標目的類名或標題。無論查詢的類號或標目出現在主要標目或參照標目或說明文字中，皆可被檢索 (見圖一)。

^[3] 目前暫建於本館網內網路 (Flylink intranet) 上，將來視需求情況，或擬製成網路版、光碟版對外發行。

^[4] 類號原採「小數制」，前三位整數位自成一節。前三級每一級類目均用三位數表示，例如「100」表示哲學類 (一級類目)，「120」表示中國哲學類 (二級類目)，「121」表示先秦哲學 (三級類目) 本資料庫以一位數表示一級類目，二位數表示二級類目，三位數表示三級類目，只是讓前三級類目顯示時，易於辨識而已。

2. 區間搜尋：在分類表中，若欲查詢的標目極為普遍，為縮小查詢範圍，可以限制在某類號區間內查詢該標目，以提高檢索的效率。
3. 合併搜尋：若欲以同一標目查詢兩個資料庫，為避免重複鍵入檢索款目，可以合併搜尋功能同時查詢兩個資料庫，只要以切換視窗的方式即可分別顯示兩個資料庫的檢索結果。
4. 超連結 (hyperlink) 查詢：標目之下若有參照款目 (見、參見、反見、反參見、入)，表示可直接以超連結的方式連接至同一資料庫的參照款目。標題後的參考類號可以超連結的方式連接至分類表；而分類表的類名若顯示底線，則表示標題資料庫中有與此類名完全一致的標題，可提供超連結連至該標題。
5. 複分表查詢：系統將可適用於整個分類法或整個標題表的複分表置於「工具列」中，當查詢某標目的說明文字中指示需使用複分表時，可隨時開啟工具列點選適用的複分表 (說明內容詳見「工具列」)，亦可在該標目下直接以 hyperlink 方式連接複分表。進入該複分表視窗後，可全文檢索該表內容，縮短逐一瀏覽的時間 (見圖二、三)。
6. 圖示 (icon)：查詢某標目的說明文字中所指示使用的複分表非通用性，而是限定於某類號區間使用時，系統在可使用該複分表的類名後加一圖示 (icon)，使用者點選後即可展開此專論複分表視窗。



圖一：類表全文檢索畫面



圖二：類表 hyperlink 複分表



圖三：標題表 hyperlink 複分表

(三) 編輯功能

任何使用者皆可於其本身的 PC 中新增標目，或選定一標目後進行修改或刪除，但增修的結果都僅暫存於各自的暫存檔內，需由特定授權人員統一彙整，並綜合各方意見且討論決定後，將結果正式轉進資料主檔內，系統經「re-index」後方完成資料更新。

若新增及修改的標目需使用到複分表，可由下拉方式選取合用的複分表，以避免錯用字彙而無法超連結。若欲修改或刪除的標目本身與另一標目超連結，而修改或刪除此標目將切斷此超連結，系統會自動提示，以避免不當的修改。

增修或刪除的標目皆以「Excel」檔案暫存，特定的授權人員可直接以「Excel」維護標目；外部檔案若以相同的檔案格式鍵標目，亦可批次轉入。

(四) 列印功能

本資料庫之列印功能，既具有選取列印範圍的功能，也具有排序列印內容的功能。就前者而言，兩個資料庫皆可選擇列印某特定日期區間內所更新的標目；在分類法部分，可選擇列印特定類號或某類號區間內所有的類目；在標題表方面，可選擇列印特定某單一字首或首字為某筆畫數內的所有標題。就後者而言，將欲列印的標目選擇出來後，可依更新日期先後或依標目的類號前後排序。假如欲列印的是標題，還可依標題詞首字的筆畫序排列。

(五) 通訊功能

為廣納各方使用的意見，蒐集應增刪修改的標目，系統提供以 E-mail 聯繫國圖編目組的功能，使用者可自由表達意見，作為修訂分類表與標題表的參考，使兩表更合乎時代的需求。

(六) 工具列

利用一般 Window 環境中常用之「工具列」概念，將使用者在資料庫任一處皆可能使用到的集中於工具列表單下，利於使用者隨時點選使用，包括下列表單：

1. 通用複分表

將分類及標題表中適用於全表的複分表集中一處，包括：總論複分表、中國時代表、西洋時代表、日本韓國時代表、中國省區縣市表、世界分國表、各國史地複分表、機關出版品排列表、中國作家時代區分表及標題複分表；且結合互有關係者（如：中國省區表與中國縣市詳表結合為中國省區縣市表），除方便使用者直接點選外，並以檢索功能增加其使用的便利性。

2. 模式標題一覽表

模式標題指的是數個同屬性標題之下展開的複分標題大致相同，為了節省篇幅並避免掛一漏萬，只在其中一個標題下詳列其複分標題，並註明其他同屬性標題皆可仿此複分的註釋文字。主表中雖已詳列該主要模式標題的複分標題，但為

避免使用者於主表中遍尋不著，故將模式標題集中於工具列中，包括：禽畜模式標題、教育模式標題、語言模式標題、佛教宗派模式標題及農作物模式標題，方便與主表中的其他同屬性主標題互相參看。

3. 中文異體字一覽表

紙本分類表及標題表中有同義字但不同形（內碼不同）的情況，如：「砲」與「炮」等，為使標目間可正確的超連結且檢索不致遺漏，故將異體字規範化；但為讓使用者明瞭系統所採用的字形以及與紙本式的差異，特明列此表以供參照。

4. 中文特殊字一覽表

與上述異體字表功能一致，不同點在於「特殊字」之處，是中文視窗環境的字集中無此字，以致系統無法鍵入，必須以圖形造字的方法顯示。在此一覽表中逐一顯示字形，並給予編號；在主表中因無法顯示字形，僅能顯示該字編號，但可超連結至本表，使用者即可看到該字字形。

綜合上述功能，本資料庫較之紙本式類表標題表及 Web 版資料庫，具備以下特性：

1. 兩個資料庫共用一套系統功能

讓使用者學習一套操作方式即可運用於兩個資料庫，增加使用效率。

2. 靈活運用視窗功能

不僅兩個資料庫可自由切換，且可用「hyperlink」方式交叉查詢，即是可用標題連接到其參考的類號，亦可由分類表連接到與類目相同的標題，使得查詢更為便利，並增加選用標目的正確性。此外，可隨時開啟複分表或其他工具視窗與主畫面並列參看，使標引更為完整與精確。

3. 查詢方式多樣化

系統可以瀏覽，或以特定類號或標目查詢；可在某類號區間以標目查詢，亦可以一個標目同時查詢兩個資料庫。既可廣博又可精確的查詢方式幾乎涵蓋使用者所有的查詢需求。

4. 複分表點選與查詢便利

通用複分表獨立於主表之外，可於主表的任一處隨時點選；專論複分表則僅出現於可使用的類目後，可避免引用錯誤，兼具指引功能。此外，任一複分表皆可全文查詢檢索，特別是複分表篇幅較大（如：中國省區縣市表）時，

複分表的檢索即十分必要。

5. 標目下參照用語規格化

書本式分類表中的參照或說明文字多有分歧，例如「參見」關係有用「參見」、「參看」等用語，但系統卻因此無法明確界定標目間的相互參照關係以做正確的連結。故建檔時需將參照標目鍵入規格化的欄位中，系統即可據以正確的連結參照款目，並顯示一致的參照用語。

6. 編輯修訂快速，是管理主題標目的利器

隨著時代演進及學科發展，主題標目亦應隨時增補，然過去紙本式的類表難於隨時將增補的資料加入並公告周知。現利用系統可快速編輯修訂標目，且可運用程式檢核標目之間參照關係是否正確，同時可在網路上立即顯示修訂的內容，隨時保持標目的正確性及新穎性，實為建置及管理我國主題標目資料庫的利器。

7. 異體字及特殊字的顯示

類表及標題表內容廣泛，以紙本形式出版時標目可能有前後不一或稍有疏失的情形；鍵入系統時，現階段的中文視窗環境又無法支援所有字形。所以在逐筆查核標目內容時，即將此類字形列出，以工具列及超連結的系統功能輔助使用者，以彌補紙本式及系統式資料庫在檢索或顯示上的缺失。

三、資料整建

資料庫的建置，系統功能為其骨幹，而資料本身為資料庫的實體內容，至為重要。本資料庫雖有紙本式分類表及標題表為依據，然建置過程中，需考慮建檔格式與系統功能配合，標目亦藉由人工及程式比對的方式查核校訂其正確性，以提高資料庫的品質。茲將資料整理的過程敘述於下：

(一) 資料依據

本資料庫建立的主要目的是提供編目員工作時之立即參考工具，因此，以現階段所使用之規範工具為依據，亦即《中國圖書分類法》(賴永祥編訂，增訂七版，1989。簡稱賴表)^[5]、《國家圖書館增訂類目表》(國家圖書館編目組編訂，2000。簡稱增訂類目表)^[6]、《中文圖書標題表》(國立中央圖書館編目組編訂，

[5] 賴永祥編訂，《中國圖書分類法》增訂七版(臺北市：三民書局經售，1989)。

[6] 國家圖書館編目組編訂，《國家圖書館增訂類目表》(臺北：國家圖書館，2000)。

修訂版，1995。簡稱標題表）^[7]除了上述三種紙本工具書外，也考量建立資料庫的目的之一是要能夠獲得即時的資料，因此也將數年來國家圖書館因實際需要增補之類號類目及標題納入資料庫。

(二) 建檔格式與資料整理

資料庫以民國88年建置之「中國圖書分類法/標題表資料庫」Web版的資料內容為基礎，由於格式不同，因此由開發廠商根據本館所訂定之欄位將該資料轉為Excel檔案。本資料庫分由分類及標題兩部分組成，資料建檔之始，考慮標題表標目之間的相互參照關係較為單純，也就先行處理標題表，其次為分類表。

1. 標題資料建檔格式與資料整理

就紙本資料而言，標題是由標目詞、參考類號、註釋文字、參照說明四部分構成，標目詞又包括正式標題詞與非正式標題詞兩種；資料庫也就根據紙本資料內容，確定標題表資料庫的資料架構包括：標題、參考類號、參考類號（試用本）、見、反見、參見、反參見、一般注等欄位。

建檔方式採取兩部分進行，首先是將紙本標題表依筆劃數分成數個Excel檔案，將標題逐字鍵入各欄位中，其中一般複分標題單獨為一個檔案，另外為使資料格式一致，將仿分標題與模式標題同時歸類在模式標題中，與一般複分標題一樣置於資料庫工具列之下，使用者可以透過超連結看到所需之模式標題。另一部分則是將歷年來本館增補之標題（增補時已鍵入Excel檔中，並定期公布於編目園地全球資訊網中）轉入。

2. 類表資料建檔格式與資料整理

紙本類表資料中，每一條標目包括標記符號（類號）、中文類目、英文類目、類目註釋、參照說明等數個要素。前三者構成標目（條頭），後兩者構成註釋（註文），中文類目除正式類目外，有時還包括類目之異名（即同義詞）。資料庫也根據類目所包括的要素，定義了類號、類目、英文、同義詞、範圍註、一般註、書例、入此、入、宜入、參見、附表等十二個欄位。範圍註即所有小字註文內容，其後再根據內容性質依照各個欄位所定義的範圍分入一般註、書例、入此、入、宜入、參見、附表等欄位，範圍註並於最後資料校核完成後刪除。茲以

^[7] 國立中央圖書館編目組編訂，《中文圖書標題表》修訂版（臺北：國立中央圖書館，1995）。

「357.51 天然元素」類為例，舉示其樣式如下：^[8]

原紙本分類表

357.51 天然元素 (天然礦) Native elements 標目 (條頭)
其經濟用途宜入於356.2.-4 注釋 (註文)

Excel 電子檔案比對說明：

類號	類目	英文	同義詞	範圍註	一般註	書例	入此	入	宜入	參見	附表
357.51	天然元素	Native elements	天然礦	其經濟用途宜入於356.2.-4					其經濟用途 / 356.2.-4		

為了讓資料庫一致化，資料輸入時有些特定欄位是依據紙本照錄，有些欄位則需經人工判定及刪除前導用語鍵入。一般而言，直接鍵入的欄位有類號、中文類目、英文類目、範圍註。至於同義詞、一般註、書例、入此、入、參見等則是需經人工判定的欄位，至於所刪除的前導用語包括「論」、「例」、「依」等，但若是「泛」、「泛論」、「其」等則予以保留。

類表中由兩個或兩個以上類目組成的「類組類目」，在紙本中多以標點符號隔開，例如：「，」（逗號）「、」（頓號）「；」（分號）「」（圓點）等，因類組類目各類目可以個別分開，所以資料鍵入時一律以「/」（斜線）區隔，替代原有標點符號。若類目間的語義不可分割，或分開後會產生詞義不完整時，則資料照錄。類組類目中有「：」（冒號）照錄，不判斷各類目語義間的關係。

註文包括同義詞、範圍註（類目註釋）參照說明、書例等。在類表中同義詞格式不統一，有用圓括號加註於類目之後，也有出現在類目註釋，以「又稱」、「又名」等引出，只要判斷是同義詞都一律放置於「同義詞」欄位。

參照說明包括「入此」、「入」（含「見」）、「宜入」、「參見」（含「互見」、「參照」等）四個欄位。

類書、普通論叢、普通叢書、群經、古算經、中國史、中國傳記、考古、中國語言、中國文學等類目之說明文字中出現書名，著錄於「書例」。凡書名後出

[8] 參閱：賴永祥編訂，《中國圖書分類法》增訂七版（臺北市：三民書局經售，1979），頁174。

現類號或「入 xxx 類號」，若是非下位類號，則改放在「入」欄位，以利進行超連結。若是書例中的類目是「書名 + (朝代, 作者)」或「書名 + (作者)」照錄。

通用複分與「專類複分」如經書複分表、中國哲學家著作複分表、佛教各宗派複分表等，另行鍵入，不與類表正表相混淆，以超連結至專類複分表。正表中各類下之仿分不視為類目註文，而仿其意另行鍵入本類之類號及類名。

工具列下共有九種複分表，是將紙本資料中的附表鍵入，包括：(1) 總論複分表；(2) 中國時代表；(3) 西洋時代表；(4) 日本韓國時代表；(5) 中國省區縣市表；(6) 世界分國表；(7) 各國史地複分表；(8) 機關出版品排列表；(9) 中國作家時代區分表。

(三) 資料比對

雖然本資料庫在鍵檔之初即已建立一套建檔格式，但所依據的規範標準是紙本資料，都是以人工方式處理，資料錯誤在所難免，又加上分類部分有參照注釋用語的不統一、類名的不一致等問題，因此本組在進行資料庫鍵檔時，除了有一套建檔原則，及採用人工逐字校對外，也考慮以 Access 進行資料比對，提高資料正確性。整個資料鍵檔流程如後：(1) 制定輸入格式；(2) 資料輸入；(3) 逐字校對；(4) 根據需求程式比對；(5) 將矛盾歧異或錯誤部分修改及新增；(6) 程式重新比對；(7) 最後校核。

1. 標題資料比對

「見」與「反見」及「參見」與「反參見」需要互相對應，資料庫在程式設計部分也將此關係建立，因此將來只要在執行資料庫新增修改刪除時，如無正確對應關係，系統會出現警告訊息（見圖四）。但資料庫現存資料，則無法自動顯示，為了讓資料本身矛盾及對應關係錯誤減少至最低，需藉程式將其中有問題的資料篩選出來。

首先處理的是「見」與「反見」，亦即只要在「見」欄位有資料，該資料就與「標題」比對，如果是一樣的內容，則將二者同時列出，經由編目員比對判斷並加以修正。「參見」與「反參見」處理方式亦同，如此一來可以將錯誤降至最低。



圖四：警告訊息

2. 分類資料比對

資料經過人工逐字校對後，本組就假設資料庫資料內容與紙本資料完全一致，但在處理過程中，發現紙本資料格式與體例並未完全一致，雖然事先曾規劃輸入格式，但人工鍵檔時，難免會產生欄位誤判情形，為了解決此一問題，就需要利用程式判定將錯誤降至最低，分類資料以程式比對的欄位，包括入、宜入、參見等欄位，除此之外，也利用程式將同義詞、入此、書例欄位資料列印出來逐筆判斷。

首先分別處理書例、入此兩個欄位，這兩個欄位與其他欄位或是類目並無對應關係，因此只藉由程式將此二欄位有資料者列出，由工作人員就其資料內容判斷是否有錯誤或是鍵檔人員誤判情形。

入、宜入、參見三個欄位資料與其他類目有相關性，可以進行超連結，也就可以程式比對資料，欄位比對部分首先處理的是「入」，由入欄位中的類號與分類類號比對，以確定入類號是否存在，如果比對不到類號，則判斷該類是屬於未展開類號或不存在類號，確定二者對應關係是否正確，有三者可能產生：一所入類號為正確，二鍵入時錯誤，三依據之資料本身錯誤，後二者則可藉此加以更正。「宜入」欄位處理方式亦同。「參見」欄位處理方式也大致相同，但為求資料一致性，也藉資料比對時將「參見xxxx 類號」中的名稱修改與類目一致。

3. 標題與分類資料比對

標題的參考類號除了有助於辨識與確認標目詞的含意外，也可與類表聯繫，尤其本資料庫有提供分類與標題的超連結，因此參考類號的正確與否也就格外重要，雖然根據紙本標題表將資料內容鍵入資料庫後有經過人工的逐筆校對，但仍然無法達到百分之百與紙本資料完全一致，又加上所依據民國4版的標題完全是人工作業，其本身亦有錯誤產生。本資料庫涵括二個來源的參考類號-類表與試用本，資料庫另一部分為類表內容，因此就利用程式將標題表中的類表參考類號與分類檔內容比對。

由於本資料庫涵括標題檔與分類檔，而研發本資料庫的最終目的是希望能夠節省編目員時間，編目員可由所決定的標題來確定類號，同時也希望藉由某類號而獲得適用之標題，因此資料處理部分除了標題表本身見反見與參見反參見的程式比對，分類檔中的入、參見的比對外，更是花費相當多的人力與時間做標題與分類的比對。首先，標題中的參考類號與類表中的類號比對，如果是存在類號，則再進行標題與類目比對。最後將標題與類目完全一樣者，輸出一個檔案，標題與類目有兩字相同以上者另以一檔處理，而兩個字以下不同者另存一檔。三個檔案分別將標目、參考類號（類表類號）、類目的格式列印，分由組內十一位編目員校核，校核的規則如下：（1）以標題表與清單核對，並將標題表中不用的標題在清單中刪除；（2）清單中的標題若無對應的類號，查出可對應的類號，並補上類目；（3）標題若有對應的類號，但無對應的類目，則判斷是因類表中該區段類號展開？或是所對應類號錯誤？並加以修正；（4）標題與類目用字的統一（如雕與彫）；（5）判斷標題對應的類號是否有錯或矛盾之處；（6）以人工判斷是類號錯誤或是類表中無此類號。

資料庫資料數量中標題達 10,218 筆，分類達 21,513 筆（未包括工具列中的複分表），鍵檔之初為了將錯誤減至最低及建立相互間連結關係，已訂定輸入原則，但資料本身有太多的例外情況，因此人工判定時難免會有錯誤產生，而資料庫製作目的主要是提供分類標題相互間之連結，甚至是分類與標題本身之間的連結，因此有時會產生為適應現有規格而有所取捨，包括類號是否展開等的問題。

資料庫資料內容除委託廠商處理外，亦由數位編目員就資料內容加以校核，但資料庫仍無法達到百分之百的正確，主要原因有如下：（1）鍵檔錯誤，未能校對出來；（2）依據之紙本資料本身有錯誤；（3）程式比對時所產生之盲點，

而無法將錯誤資料比對出來。雖然希望能在高達三萬多筆資料資料庫達到零錯誤的目標，但就實際的情況並無法達到，唯有期望未來在使用時能陸續發現將其改正，以減少錯誤至最低。

四、資料庫檢討

本資料庫之建置，由最早的 Web 版，進展到 Window 版，畫面與紙本接近，檢閱上非常 Friendly，本是一大進步。但就其作用來看，目前形態還有改進之處，應該更進一步發展成與書目資料庫相聯結的形式，這樣才真正把資料庫的強大功能發揮出來。

目前資料庫分為分類資料庫和標題資料庫兩部分，這兩者之資料是分別建置的，唯透過系統軟體設計，兩者間之資料可以交互檢索而互通有無。雖然如此，本資料庫系統還是存在著缺點。因為本資料庫之建置，仍只是紙本《中國圖書分類法》、《國家圖書館增訂類目表》、《中文圖書標題表》等規範工具書的電子化而已。因為這層理由，使我們聯想到假如在建分類及標題資料庫之外，另與已建置書目資料庫相聯結，那就會是非常完善的措施了。如此設計，不僅可作為編目人員文獻標引時的參考，同時也可作為讀者檢索文獻之需。

其實這種構思、做法，據我們所知香光尼眾佛學院圖書館即已在「香光資訊網」網頁上設置「電子資料庫」，開闢「佛教文獻結合索引典」視窗，以檢索「香光尼眾佛學院文獻資料庫」，主題詞表資料庫與書目文獻資料庫密切結合，奠定下良好的基礎。^[9]

與書目資料聯結是充實措施的第一步，接下來即可研議分類號或標題詞自動帶出的功能。所謂「自動帶出功能」，即指編目人員只需鍵入分類號或標題詞其中一個檢索符，系統即可自動帶出相關的另一個檢索符，如此可以減少編目人員的工作量。再來，可建立資料庫的管理功能，自動累計類號標引及標題標引的次數，以作為日後修訂分類表及標題表的參考。

^[9] 香光尼眾佛學院圖書館文獻資料庫網址：www.gaya.org.tw/paper_index/index.asp

五、規範化問題

(一) 內容體例再完善

1. 內容增補的問題

除了依據《中國圖書分類法》、《國家圖書館增訂類目表》、《中文圖書標題表》建檔外，同時，也利用這次建檔之便，增補了頗多的類目及標題。例如檔案學（028.9）、應用心理學（177）、佛教文藝及佛教語文（224.5）、電腦工程（312.9）、醫經及中醫理論（413.1）、環境工程（445.9）、中國兒童文學（859）、織繡服裝工藝美術（966）等。但需要增補修訂的，則遠超過前面所舉，譬如資訊科學、企業管理、各地文學、各族文學、美術等，都可經進一步討論後予以增補收錄，以解決無類可歸或無標題可標引的缺失。

2. 「內部一致性」的問題

所謂「內部一致性」是指包括交替類目，類目及其注釋文字，標題及其注釋文字，類目與標題數者間，所用之類目詞、標題詞、體例用語一致性而言。換句話說，即指涉及同一事物或概念，應該使用相同的詞彙來指稱或敘述，不可存在「各自為政、各自表述」的現象，以免增加使用者的困擾。

在這次建檔過程中，我們已努力朝這個方向進行。例如類目詞與標題詞，即統一採用以「尚書」、「詩經」、「周易」、「理則學」等詞彙為規範詞，而把「書經」、「詩」、「易經」、「論理學」、「邏輯」等同義詞，作為類目的注釋文字或標題的「見」參照。即使這樣，需要改進的地方還是很多，例如類目詞及標題詞「古典主義（音樂）古典派（文學）古典派（繪畫）」三者之「主義」與「派」詞素；以及類目詞「英語」，標題詞「英國語言」等，均宜進一步加以規範，以求統一。

3. 參照符號問題

分類資料庫的參照符號有「入」、「入此」、「宜入」、「參見」四種，標題資料庫的參照符號也有「見」、「」（反見）、「參見」、「」（反參見）四種。^[10] 有關它們的涵意及用法以及二者相互間的關係還有深入探析的必要。例如，分類資料庫「入」、「入此」、「宜入」、「參見」四種參照符號，其區別的

^[10] 有關參照符號，各種主題法、分類法所用符號及用法不盡相同，本文所論，以《中國圖書分類法》及《中文圖書標題表》為準。

疆界何在，應給予釐清，否則造成「甲乙互用、彼此不分」的局面。在建檔過程之中，我們即發現用「甲」用「乙」均可的現象，這就說明了是資料庫未曾嚴格定義四種參照符號所致。

標題資料庫「見」、「反見」、「參見」、「反參見」四種參照符號，其中「反見」與「反參見」直觀性較差，使用者不易認知與識別；如採用「反見」與「反參見」作參照符號，也有完全人為、用詞生硬、不合構詞規律的問題存在。因此，將來修訂宜改為合乎構詞法、直觀性良好、意義明白顯豁的詞彙，例如採用「用、代、狹、廣」。

同時，標題資料庫中原使用之「參見」參照，宜進一步依功能再區分為「狹義詞」(狹)及「相關詞」(關)二者。還有，標題資料庫之參照注釋宜進一步增加「族首詞」(族)，以方便日後詞表整個體質的改換，以接近自然語言作為文獻檢索的基礎，即由「前組合式」的標題詞表，變換為「後組合式」的主題詞表。

4. 「參見」參照語意問題

分類資料庫之「參見」參照，其含意是說明兩類目間相關的意思，與標題資料庫之「參見」參照其含意是指狹義詞，不論含意及用法均有差別。這二者實在應該加以統一為宜。例如：將分類資料庫中之「參見」參照和標題資料庫中之詞意相關的「參見」參照，均配合改為「相關」(關)參照，使得二者的參照符號能合一致性的原則。

(二) 檢索符的規範化

資料庫是否實用好用、盡善盡美，取決於許許多多內外因素的配合，而檢索符的規範化問題，應該也是考慮重要因素之一。只有在實行了檢索符標準化和規範化之後，才能使資料庫與使用者之間建立良好的互動關係，也才能使資料庫檢索更合用、更人性化，進而提高並增進資料庫的使用功能及效率。

就符號的類型來分，資料庫的檢索符可分為符號式及文字式兩種。對於以數字和字母為主的符號檢索符，例如：分類號、化學符號、代碼等，相對的規範化問題比較小。但對於以文字為主的檢索符，例如：標題詞、單元詞、敘述詞等檢索符，尚存在下列兩項規範化問題：詞語本身的規範和詞語書寫的規範。關於檢

[11] 所謂「檢索詞語規範化」，是指檢索詞彙的控制，主要包括詞量控制、詞類控制、詞形控制、詞義控制、詞間關係控制、專指度控制、先組度控制等。參閱：馬張華，侯漢清編著，《文獻分類法主題法導論》(北京：北京圖書館出版社，1999)，頁 19-23。

索詞語規範化問題，又包括同義規範和詞義規範兩者。^[11] 有關同義規範和詞義規範等問題，通常在編製詞表時，會透過「見」參照以規範同義詞和準同義詞，例如「理則學、論理學、邏輯、名學」等詞，可以選「理則學」為正式標題詞，以其餘的為非正式標題詞。透過加註「限定語」以區分同形異義詞，例如「指揮（音樂）」與「指揮（軍事）」兩標題詞，即分別加註「音樂」和「軍事」兩限定語以區別之；至於「內科」與「內科（中醫）」兩標題詞，只需加註其一「限定語」，即可形成區分詞義作用。

但關於詞語書寫之規範，是個牽連頗為廣泛又不屬於檢索詞語本身的問題，向來為多數資料庫建立者與使用者所忽略，尚未引起大家應有的重視。不過，覈按實際，很顯然的對於資料庫的利用，具有很深遠的影響。例如某主題詞表或主題資料庫選用「傢俱」、「臺灣」、「祕書」等詞語為規範詞，但未建「家具」見「傢俱」，「台灣」見「臺灣」（其他地名「臺」字仿此），「秘」見「祕」字等檢字的「見」參照，則使用者因查不到自己心中設想的檢索詞語，以為該詞表或資料庫未收錄，如此檢索者就會停止檢索活動，連帶地使資料庫的功能無從發揮，其結果豈不與未建資料庫相同？

基於上述理由，我們堅持編製了「中文異體字一覽表」一種^[12]，作為資料庫檢索利用時查詞的輔助，以完善並加強本資料庫的檢索功能。只有承載詞語的文字的標準化和規範化以後，才能使資料建立者及檢索者所使用的檢索詞語及其字形是相同的、是一致的。其次，我們也希望引起大家對中文正字、正形的重視，從字詞典編纂，以至於一般人書寫、交流，電腦字集建置及其輸入、輸出，以及檢索規範工具書編製等各個環節，都能使用正確的、標準的、規範的字形字體，從而解決詞語書寫的規範問題。當然這是初步的試探，而所建之異體字表也未盡善盡美，猶待吾人努力之處尚多，但願大家共同努力奮進！

(三) 分類主題一體化

民國 73 年 3 月編製《中文圖書標題總目初稿》時，即以《中國圖書分類法》的分類號作為參考類號，這可以說是臺灣地區有意朝向「分類主題一體化」發展的新嘗試。這次，我們更把分類和標題二者緊密結合起來，透過軟體支援可以交

^[12] 所謂「異體字」，是指音同義同而形體不同的字。即俗體、古體、或體、帖體之類。如「嘆」、「歎」；「跡」、「蹟」、「」；「祕」、「秘」等。參閱：夏征農主編，《教育文藝辭海》（臺北：臺灣東華書局，1993），頁 92。

互檢索，可以說是向分類主題一體化邁進了一大步。

所謂「分類主題一體化」，即指分類法與主題法在相互結合向著一體化方向發展。換句話說，也就是分類語言和主題語言在概念（術語）系統、參照系統、標識系統、索引系統等方面實現了兼容，把兩種以上的詞表融合為一個統一的系統。

「分類主題一體化」最早濫觴於美國杜威 (Mevil Dewey) 《杜威十進分類法》書後所附的字順相關索引。20 世紀則有英國艾奇遜 (J. Aitchison) 所編製之《分面敘詞表》(Thesaurfacet) 及《聯合國教科文組織敘詞表》(Unesco Thesaurus) 等詞表，1996 年中國大陸也在《中國圖書館圖書分類法》和《漢語主題詞表》兩者基礎上編制了《中國分類主題表》，但在臺灣地區則尚未有較正式且成熟的一體化詞表，不能不引為一樁憾事。^[13]

分類法與標題法原是文獻檢索兩種不同的方法和不同途徑，它們各有所長，也各有所短，兩者不能互相取代。長期以來，人們一直想方設法使兩者結合起來，互相取長補短，以滿足人們進行多元檢索的需求。分類法可說源遠流長，標題法則是新起之秀。開始時，兩者並行各自發展；甚或有人認為可用標題法以取代分類法。但後來實踐證明，對於一定範圍檢索（與特定範圍檢索相對），分類法與標題法兩者有如左右手，允宜攜手合作不宜偏廢。唯有如此緊密結合在一起，才能真正發揮檢索文獻的功能。

分析而論，分類檢索具有「鳥瞰全局、觸類旁通」系統性和擴檢、縮檢運用自如的優點，而標題檢索則具有高度的專指性和直觀性的優點。假如兩者合併使用，優點之一便是使用者就可在既經濟又有效的檢索系統中進行字順主題查詢和系統分類查詢，以提高文獻的檢全率和檢准率，並節省查找時間。優點之二便是索引人員可以同時完成文獻的分類標引和主題標引。上述兩點理由，是我們這次遂行分類主題一體化的思考基點。我們希望以此為基礎、以此為契機，將來更進一步編製或建置一種真正一體化的標引規範工具書或標引規範資料庫，為這貧乏的圖書館園地，耕耘出一片的新綠與幼苗。

(四) 規範修訂的參考

規範工具一方面要求穩定性，另一方面要求靈活性或適應性，而這兩者的要求

[13] 參閱：馬張華，侯漢清編著，《文獻分類法主題法導論》（北京：北京圖書館出版社，1999），頁 179-184。

是具有矛盾的。這種同時追求穩定性與靈活性，以分類法的修訂尤為顯著。因為分類號碼是藏書分類組織及組織排架目錄的依據，因此類號的更改變動，就足以牽一髮而動全身。分類法的增補及換用，相對的往往也產生目錄改編或圖書改編的問題。但社會事物總是不斷求新求變，學術文化也總是日新月異。過一段時間後，標引規範工具就漸漸不能適應社會文化的需求，因此就需要不斷加以增補修訂。如何在靈活性與穩定性之間求取平衡點，是得我們思考並努力的方向。

在這次建置資料庫過程中，除了按三種紙本規範工具書鍵入資料外，同時也作了部分修訂、增補和統一體例的工作。在修訂和增補體例過程中，除了極少數類目採用「類目改訂法」外，大多數類目的增訂，我們運用了下列六種方法，以不影響原來分類號為原則，作為類表修訂增補之手段。這些方法，大都不產生改書、改號問題，因此可作為今後規範工具修訂，譬如編訂《中國圖書分類法》第九版的參考。茲將六種類目修訂方法，簡述如下：^[14]

1. 空號增訂法

在編製類表當時即預留類號，以備將來修訂類表增補類號之需。

2. 類組類目法

將兩個性質相近或互有關聯的類目，合組起來以構成一組新的類目，是謂「類組類目法」。例如原類目為「472.2 鑄工」，經修訂後改為「472.2 鑄工；熱處理」，即應用類組類目法將相關的兩類目組合在一個類號之下。

3. 類義擴充法

將原來類目含義加以擴充，並更換類名，以求「名實相符」，是謂「類義擴充法」。例如類號「224.5」原類名是「佛教藝術」類，修訂後因增入佛教文學、佛教語文，因此將類目涵義擴充並更改為「佛教文藝；佛教語文」。

4. 類目展開法

將原來未加細分的子目予以展開，以增補新的類目，容納新的事物，是謂「類目展開法」。例如原類目「177 應用心理學」，經展開後增補了「177.1 溝通心理學」、「177.2 成功法」、「177.3 人際關係」等12個四級和五級的子目。

5. 類目注釋法

例如原類目「201 宗教政策」，經修訂後增補類目注釋文字為「宗教法規入

[14] 參閱：國家圖書館編目組編訂，《國家圖書館增訂類目表》（臺北：國家圖書館，2000），編例。

此」，從而解決宗教法規無類可歸的窘境。

6. 借號增補法

本非某類之下位類，因性質相近或互有關連，在該類目之下借用其號碼，以增補新類目，是謂「借號增補法」。例如新增類目「314.89 組合數學」即借用類目「314.8 離散數學」以增訂之，並非謂組合數學為離散數學之下位類。

以上類目增補法，可說是這種建置分類資料庫的重大收獲之一。借此機會提出，一來請分類法的修訂者多加利用，一來也請方家不吝指教以求改進。當然上述六種方法也不是萬靈丹，可以一舉解決分類法所有的修訂增補問題。可以預見的，將來修訂增補類表之際，還是需要適當地運用「類目改訂法」以濟其窮，但應儘量少用為宜。

六、結語

分類及標題資料庫的建立，可以說是檢索語言及方法、電腦技術、網路技術等數方面知識與智慧的結晶，也是電腦科技產品在圖書館業務方面的推廣與應用。由於電腦的導入與利用，對於編目規範工具的編輯、應用與管理，也產生了深遠的影響與衝擊。分類表與主題詞表，一方面是編目人員標引文獻的必備工具，另一方面又是讀者檢索文獻的輔助工具，電腦化是未來發展必然要走的方向。

就這個趨勢而論，本館可以說已然走向時代尖端了，並踏出成功的第一步。所建置之分類及標題資料庫，雖然不夠盡善盡美，待改進之處也頗多。我們認為只要假以時日，不斷求新求變，將來定會成為編目不可或缺的利器。就改進方向來說，上述各端，諸如資料庫功能的再改進，分類表、標題表、複分表等資料的充實及體例統一，嘗試與書目資料庫的結合以及詞表自動管理功能的設計等等，都應該是我們引為注意和重視的地方。還有目前資料庫，推出的僅Window版而已，假如經費許可的話，未嘗不可研製推出光碟版，對外發行，以嘉惠圖書館編目從業人員。

在當今之世代，文獻電子化是未來必然的趨勢與方向。更有進者，《中國圖書分類法》增訂八版除發行印刷本外，同時也發行光碟版^[15]，而《中圖法》增

[15] 賴永祥編訂《中國圖書分類法》增訂八版，將於民國90年9月由文華圖書館資訊管理公司出版（臺北市民生東路5段219號；電話：2760-6492；E-mail：sewice@mandarin.com.tw），分精裝典藏版（附光碟）、實用版等版。

訂九版也擬由國家圖書館接辦，以如此的趨勢，我們怎敢走回頭路？這或許是我們目前推出分類及標題資料庫，有更深一層的意義在吧？我們希望今天辛勤地播種與耕耘，會是明日錦繡前程的保證與保障。但願資料庫能不斷求新求變，以更上層樓！

參考書目

1. 美國國會圖書館分類法，光碟版。
2. 馬張華，侯漢清編著，《文獻分類法主題法導論》（北京：北京圖書館出版社，1999）。
3. 國立中央圖書館編目組編訂，《中文圖書標題表》修訂版（臺北：國立中央圖書館，1995）。
4. 國家圖書館，飛資得資訊公司製作，「中文圖書標題表 / 中國圖書分類法資料庫」(Web版)(臺北：國家圖書館，1999)。
5. 國家圖書館，飛資得資訊公司製作，「國家圖書館分類 / 標題資料庫」(Window版)(臺北：國家圖書館，2000)。
6. 國家圖書館編目組編訂，《國家圖書館增訂類目表》（臺北：國家圖書館，2000）。
7. 賴永祥編訂，《中國圖書分類法》增訂七版（臺北市：三民書局經售，1989）。

Introduction of National Central Library Classification / Subject Headings Database: Issues of Standardization

Ling-hwa Hsu Ching-ling Juan Yeou-min Chen

Abstract

The first Classification / Subject Heading Database (Windows version) in Taiwan has been co-developed by the National Central Library and Flysheet Information Services, Inc. This article is an introduction of the functions and specifications of the database, verification of data entries and the issues of standardization.

There are 4 main parts: First, system functions and specifications of the database, including browsing, searching, editing, printing, communication, tool box and windows switching, are user-friendly and powerful. Secondly, verifying data entries

Keywords (關鍵詞): Classification ; Subject Headings ; Database ; Standardization
分類 ; 標題 ; 資料庫 ; 規範化

Ling-hwa Hsu : Editor, Cataloging Department, National Central Library, R.O.C ; E-mail: clhhsu@msg.ncl.edu.tw

Ching-ling Juan : Assistant Editor, Cataloging Department, National Central Library, R.O.C. ; E-mail: cljuan@msg.ncl.edu.tw

Yeou-min Chen : Senior Clerk, Cataloging Department, National Central Library, R.O.C ; E-mail: cla-sub@msg.ncl.edu.tw

are based on *New Classification scheme for Chinese Libraries* (1987, 7 th ed.), *NCL's Additions of Classification Scheme* (2000 ed.) and *Chinese Subject Headings* (1995 ed.). All entries are formatted by Excel and the matching is examined by computer programs. Thirdly, the future goal is the linking between classification/subject headings data entries and bibliographical database records, that is, when we catalog, the system can retrieve the right classification no. and subject headings from the database automatically. Fourthly, in developing the database and system, we find that standardization issues about the display format, retrieving keys, data entries, and revision processing are very important.