

中文名稱權威檔合作計畫問題探討

吳明德

臺灣大學圖書館

鄭恆雄

國家圖書館編目組

摘要

國家圖書館與臺灣大學圖書館於民國 87 年 7 月開始進行合作建置中文名稱權威檔計畫，以期將權威檔轉上「全國圖書資訊網路系統」，供各合作館查詢取用。建檔規格採用中國編目規則及中國機讀權威記錄格式，並建立系統以供建檔、查詢、轉檔之用。計畫進行所面臨的主要問題如下：中文機讀權威記錄與 USMARC 轉換、羅馬拼音、外文人名、簡體字與繁體字等。由於兩岸三地及歐美圖書館界早有合作建立中文名稱權威資料庫之構想，因此本文也提出合作時需探討之議題：建立實體或虛擬的權威檔？各地使用之權威格式如何修改以便權威記錄之交換？各地區設計之權威記錄轉換格式如何互享？如何共同向自動化廠商提出權威控制系統之需求？

壹、前言

國家圖書館（以下簡稱「國圖」）推動之「全國圖書資訊網路」（NBINet），於 1991 年 10 月啓用以來，各合作館曾多次討論期望建立權威記錄檔，以提供各合作館作為進行合作編目之依據。各合作館亦曾達成共識，網路合作編目之權威記錄以國圖及國立臺灣大學圖書館建立之權威記錄為依據，以合作方式建置。早在 1990 年間臺灣大學圖書館採用中國機讀權威記錄格式（Chinese MARC Format for Authority Records，簡稱

CMARC) 在 NBINet 資料庫完成 11,215 筆權威記錄建檔，供各館使用。唯國圖建立的權威記錄，則因缺乏人力進行權威記錄之整理，包括重複資料之合併、補充以及建立參照關係，未能轉上網路提供使用。事實上，國圖建立之國家書目資料庫數量龐大，編目建檔時相應訂定的權威記錄也十分可觀，數量截至 1998 年 6 月止共有 56 萬筆。其中中文個人權威記錄 25 萬 7 千 5 百餘筆，團體權威記錄 1 萬 9 千 7 百餘筆，都是採用 CMARC 建置於 URICA 系統上。

另一方面，國立臺灣大學圖書館的自動化系統於 1993 年 11 月 18 日啓用，曾在資料庫進行轉錄時，就第一次轉錄之 13 萬筆西文書目，委請 LTI (Library Technologies, Inc.) 產生 7 萬 5 千筆名稱及標題權威記錄；而後，於 1994 年 9 月將權威控制作業納入編目自動化流程，進行建檔與整合書目作業。迄 1999 年 7 月底，資料庫中已建有 21 萬筆中西文權威記錄。其中的中文權威記錄約 4 萬 5 千餘筆，譯名權威記錄約 8 千 8 百餘筆。臺大建立的中西文權威記錄，都是採用 USMARC Format for Authority Data (簡稱 USMARC) 建檔。

1997 年起，國圖進行 NBINet 新系統的建置，各合作館的書目記錄重新予以整合，新系統至 1998 年 4 月開放使用，開放之初書目記錄約為 65 萬餘筆，隨後各合作館的書目記錄持續轉入，至 1999 年 6 月數量已達 124 萬 3 千餘筆。書目記錄的增加，權威記錄之重要性更形突顯，因此國圖乃邀集臺大共同擬訂合作建立中文權威檔的計畫。

貳、權威記錄檔建置計畫

合作建置權威檔計畫自 1998 年 7 月起展開，由國圖與臺大合作，以兩館權威記錄為基礎，逐年進行整編，期已形成較為完整之權威紀錄參考檔。因仍在整理階段，現階段建置於國圖「書目資訊中心」的電腦系統，亦尚未與書目檔連結，未來將轉上「全國圖書資訊網路系統」，提供各合作館查詢取用。茲將計畫現況簡述如下：

一、擬定權威整理及建置原則

由國圖及臺大共同組成「中文權威參考資料庫建置小組」定期舉行會

議，商討標目彙整進行的方式與原則。目前達成的決定有以下幾點：

1. 中文權威標目整理以人名及團體名稱為優先。
2. 建檔規格與依據：MARC 格式依據《中國機讀權威記錄格式》；編目規則中文依據《中國編目規則》、西文依據《AACR2》。
3. 臺大圖書館與國圖使用標目之合併與擇定原則：
 - (1) 依據編目規則進行標目之合併與擇定，有疑慮者再行討論；
 - (2) 只要任一館有書目就可以建立相關之權威記錄；
 - (3) 權威記錄之建立應顧及完整性。
4. 中國人名主標目以中文繁體著錄，簡體人名置於反見標目；外國人名主標目採用外文。
5. 反見標目著錄各種不同語言之中譯名、羅馬拼音，儘量收集不同的羅馬拼音。
6. 資料來源之著錄原則：非必備項目，可視情況決定是否著錄。
7. 參考工具書之收集與著錄：以臺大目前所提之參考工具為主，如有闕漏可另行增補。
8. 開發權威檔相關程式，應包括 Web 介面之查詢與維護功能。

二、進行系統開發

依據整理的原則擬定需求，並與廠商商討系統開發權威彙整及建置系統。

開發的程式包括：

1. 權威記錄轉換程式：以《中美機讀權威記錄格式對照表》為藍本，撰寫程式，提供 CMARC 與 USMARC 兩種權威記錄格式的轉換功能。
2. 權威記錄彙整程式：含權威記錄之查詢、新增、修改、刪除、轉移與合併。
3. 權威記錄查詢及轉錄：提供 Web 介面查詢，可由參照關係檢索出相關標目；並可提供 CMARC 及 USMARC 兩種權威機讀格式之轉出功能。

目前已開發完成 DOS 及 Web 兩種版。包括：資料查詢及維護、資料轉入及轉出、產生報表等功能，內容如下：

1. 資料維護

包括主標目、反見及參見、系統號查詢、瀏覽等。進行維護時會顯示該筆資料的新系統號、原系統號、MARC 格式、來源欄號、主標目、最後更新日期及資料狀態等。具有全螢幕維護功能。

2. 資料轉入及轉出

包括資料轉入、驗收筆數計算、索引檔重整等。資料轉入只接受符合 ISO 2709 格式的資料，但可為 CMARC 或 USMARC 記錄；內碼可為 CCCII 或 BIG5 碼。資料轉出可配合不同圖書館自動化系統需求，轉出 USMARC 或 CMARC；CCCII 或 BIG5 碼之權威記錄。

3. 報表

包括權威標目比對報表、批次註記刪除功能、待刪資料核對功能、待刪資料移出功能等。

三、權威記錄彙整作業

中文權威標目整理，包括權威資料的查核與權威記錄的維護兩部分工作，具體內容包括：

1. 將國圖中文個人及團體權威記錄全數分批列印。
2. 將臺大已建檔之 USMARC 格式中文人名權威記錄轉換為 CMARC 格式，並轉入資料庫。
3. 將標目依筆畫序列印，以方便格式錯誤的更正。
4. 合併：相同權威名稱的合併
 - (1) 判斷是否為同一人，不同人應建立不同記錄，並加修飾予以區分。
 - (2) 若為同一人，應彙整所有相關標目，建立一筆完整記錄。
 - (3) 註記刪除不用之重複記錄。
5. 建立參照關係：補充現有標目的參照關係及相關標目的整合。
 - (1) 「反見標目」及「反參見標目」應力求完整。
 - (2) 中國人名主標目以中文繁體著錄，簡體人名應置於參照段。
 - (3) 以參考工具書考證並查核「反見標目」及「反參見標目」之正確性。

6. 現有標目的補充說明（例如：加註學科、生卒年月、附註說明等）。
 - (1) 查詢 OCLC 及 LC Name/Subject Authority File，並增補標目。
 - (2) 每一標目應編製對應之羅馬拼音。
 - (3) 參考工具應詳列於欄位 810。
7. 線上建檔及維護：彙整後記錄應進行審核及校閱。

參、合作建立機讀權威檔的問題

雖然國圖與臺大同樣採用《中國編目規則》及 AACR2，但由於缺乏像 LC Rules Interpretation 一類的編目規則解說，在合作之前二館之間亦未有定期的編目規則討論會議，在標目的選擇及建立自然會有一些差異。藉助此次合作建置權威檔的機會，國圖與臺大得以將一些編目規則及機讀權威記錄格式的解釋予以釐清，不同的作法取得共識，有些細微的差異不致影響權威檔之運作者，則不勉強各館更改。但在合作建置權威檔的過程中，面臨了一些不僅存在於國圖與臺大的基本問題，也是未來臺灣圖書館與海外其他地區圖書館合作時必然要解決的問題，茲分述如下：

一、中國機讀權威記錄格式與 USMARC 轉換的問題

在合作建立權威檔首先面臨的問題是二個圖書館所使用的機讀權威記錄格式不同，國家圖書館使用《中國機讀權威記錄格式》，臺大圖書館使用《USMARC Authority Format》。《中國機讀權威記錄格式初稿》於 1986 年出版，經過修改於 1994 年正式出版，係以 UNIMARC 為藍本。不同的機讀編目格式間之轉換通常會無法對應或資料遺失的情形，但經過比對，二者之間的主要差異如下：¹

0--段：在 CMARC 中，欄位 005（最後更新時間），050（國立中央圖書館權威記錄識別號碼），099（國家書目中心資料庫權威記錄系統識別號）等，無法對應至 USMARC；USMARC 之 010(LC

1 吳瑠璃，劉春銀，〈中美機讀權威記錄格式轉換之探討〉，《國家圖書館館刊》，85年第1期（民國85年6月），頁77-82。

Control Number) , 014(Link to Bibliographic Record for Serial or Multiple Item) , 020(ISBN) , 022(ISSN) , 042(Authentication Code) , 045(Time Period of Heading) , 052(Geographic Classification Code) , 066(Character Sets Present) , 072(Subject Category Code) , 073(Subdivision Usage) , 09X(Local Call Number) , 在 CMARC 無適當之對應。

但如果仔細分析，則會發現這些差異並未實質上涉及標目本身（CMARC 2--段，4--段及5--段），因此，這些差異應可接受。例如，在權威記錄最主要的部份—主標目方面，CMARC 與 USMARC 內容雖有不同，從《中美機讀權威記錄格式對照表》的分析（註2），可知在轉換上並無問題：

C MARC			US MARC				
欄號	分欄	名稱	欄號	分欄	名稱	備註	
200	a	標目主體	100	a	Personal Name	a, b → a	
	b	副標目					c → c
	c	年代以外之修飾語		c	Titles and other words associated with a name		d → b s, f → d
	d	世代數		b	Numeration		g → g,
	f	生卒年代		d	Dates associated with a name		4 → e x → x
	g	副標目頭字語全名					y → z
	s	朝代		g	Miscellaneous information		z → y
	4	著作方式		e	Relator term		
	x	主題複分		x	General subdivision		
	y	地區複分		z.	Geographic subdivision		
z	時代複分	y	Chronological subdivision				
7	文字						

二、羅馬拼音的問題

國家圖書館與臺大圖書館的機讀編目記錄係採用 Wade-Giles 羅馬拼音系統，機讀權威記錄亦然。在機讀權威記錄中，所有的中文名字，不論位於主標目段、反見標目段、或反參見標目段，都另外加上 Wade-Giles

2 機讀權威記錄格式對照研訂小組，《中美機讀權威記錄格式》（臺北市：國家圖書館，民85年），頁8。

羅馬拼音，這些羅馬拼音的名字都放在反見標目，例如：

- 200 1\$a 丁 \$b 衣
- 400 1\$a Ting \$bI,\$f1925-
- 400 0\$a 尼羅 \$f1925-
- 400 0\$a Ni-lo\$f1925-
- 400 1\$a 丁 \$b 克用 \$f1925-
- 400 1\$a Ting \$bK'e-yung,\$f1925-

由於漢語拼音已成爲主流，國圖與臺大會考慮以之作爲採用之羅馬拼音系統。一個理想的權威記錄最好能夠包含各種常用的羅馬拼音，以符合不同讀者的查詢需求。然而，有些中文人名的權威記錄中，往往包含十數個反見標目（例如，有的著者有十數個筆名或其他名字），若要靠人工一一鍵入二種羅馬拼音，則頗耗費人力。目前二個圖書館所使用的圖書館自動化系統都無法自動將中文標目產生羅馬拼音，因此並不考慮輸入多個羅馬拼音。

三、外文人名的問題

許多外國著者的作品被譯爲中文，圖書館可能同時有其原文作品及中文譯本，例如，海明威 (Hemingway, Ernest)，此時，該著者應建立一筆或二筆權威記錄？此種情形其實也適用在一個中文著者同時也有外文作品的情形，例如，劉紹銘 (Lau, Joseph S. M.)。

理論上，一個著者應該只有一筆權威記錄，而將所有其他名字，包括羅馬拼音、外文名字，都著錄於同一記錄，有了此一權威記錄，則不論從任一名字均可查到該著者的所有作品。USMARC 的欄號 100 Headings — Personal Name 是不可重複的，而《中國機讀權威記錄格式》的欄號 200 權威標目----人名是可重複的，用以記載不同文字標目，因此，可以用重複的欄號來記錄不同語文的標目。但我們認爲，一筆權威記錄同時有二個主標目，固可顯示二個標目同樣重要，但從檢索的角度來看，無論是主標目 (Heading) 或是反見標目 (See From Tracing)，都同樣能達到檢索該著者所有作品的目的。而且，由於國圖與臺大使用的自動化系統並不接受一筆權威記錄可以有二個主標目，所以我們決定採用主標目不可重複，亦即只能有一個主標目的作法。

筆者以爲，主標目或許可以視著者使用的語文來訂定。非中文著者，以其原文名字立爲主標目，中文名字或譯名及羅馬拼音則著錄於反見標目。中文著者，則以其中文名立爲主標目，其外文名字及羅馬拼音則著錄於反見標目。若採用將中外文名字分立二筆權威記錄的作法，亦即同時將中文名字或是外文名字著錄於反參見標目(See Also From Tracing)的作法，則無法達到以一個名字一次即可查詢到該著者所有中外文作品的目的。

四、簡體字與繁體字的問題

臺灣地區的圖書館在編目大陸出版品時自然會遭遇簡體字與繁體字的問題。並不是所有的圖書館自動化編目系統都能處理簡體字，有些圖書館一律將簡體字改爲繁體字著錄，有些則採依樣著錄的原則，書上用簡體字則著錄簡體字，書上用繁體字則著錄繁體字。

建立名稱權威記錄時也會遭遇同樣的問題。由於同一個作品可能會有臺灣的版本與大陸的版本，該著者的名字就會有簡體與繁體之分。我們決定以繁體字名稱作爲主標目，而將簡體字名稱著錄於反見標目。但是，有些簡體字名稱實在不易轉換成繁體名稱，例如，「庄」字，有時無法得知其爲「莊」或即爲「庄」，「后」字無法確定其爲「後」或即爲「后」，無法確知者，則依樣著錄，另將各種可能之繁體名稱著錄於反見標目。

肆、與國外圖書館合作建立中文名稱權威檔的問題

兩岸三地及歐美東方圖書館界長久以來即有合作建立中文名稱權威資料庫之構想，現今網路之普及當可使此構想早日實現。國家圖書館與臺大圖書館合作建立中文權威檔，除供臺灣地區圖書館使用之外，亦極願與國外圖書館界共享。

一般而言，若要有效地合作建立名稱權威資料庫，則參與之圖書館必須要採用同樣的編目規則及機讀權威編目格式。除此之外，中文名稱權威檔的建立尚有字碼、羅馬拼音、外文人名、繁簡字體之問題。臺灣圖書館界使用《中國編目規則》，與大陸、港澳、歐美地區圖書館界使用的編目規則不同，但有關標目選擇之條文大抵參考 AACR2，彼此並無太大的歧

異，因此不會構成合作的問題。至於 CCCII 碼與 GB 碼的轉換，大概會有少數無法對應的字需藉助人工轉換。另外，機讀權威記錄格式方面，從臺灣的經驗看來，CMARC 與 USMARC 的轉換也不是問題。

進行合作會遭遇的主要問題在於各地區圖書館對於羅馬拼音、外文人名、繁體字與簡體字採取不同之處理方式而引起。目前臺灣地區所建的權威記錄雖然採用 Wade-Giles 系統，未來藉助圖書館自動化系統的自動轉換成漢語拼音，應該不會是問題，而漢語拼音是中文羅馬拼音的主流，臺灣地區圖書館將來應會採用。

外文人名、繁體字與簡體字的問題似乎可歸納為一個問題，即是以何者為主標目、何者為反見標目的問題。一個外國著者，無論是以其外文名字作為主標目，或是以中文譯名為主標目，或是二者都作為主標目，應該都是可被接受的，只要能利用任一標目都能找到該著者的所有作品即可，但各地區圖書館對於主標目的選擇恐怕很難取得共識。同樣的問題也發生在繁體字與簡體字的選擇，臺灣地區的圖書館不會以著者的羅馬拼音、簡體字名稱作為主標目，其他地區圖書館卻會如此選擇，各地區的作法大概無法改變。

筆者以為，為了合作建立中文名稱權威檔而要求各地區圖書館改變作法恐怕不太實際，因為各地區圖書館都有其遵守的編目規則、編目政策等等，將舊有書目記錄或權威記錄加以變更，除了耗費人力及經費，也會造成目錄使用者的不便。因此，恐怕無法仿效美國 NACO 之作法，共同建立一個實體的權威檔。而且，如果由某一圖書館彙整，建立一個線上實體的權威檔，則所耗費於核對、增刪資料的人力必甚為可觀。若要求參與之圖書館在提供權威記錄時，增加輸入其他地區圖書館會使用之資料，恐怕也會造成圖書館的額外負擔。

因此，未來合作機制如何建立，實有賴各地圖書館界同道進一步討論細節。謹提出可能需討論之議題略述於下，就教於方家：

應建立一個實體的共用權威檔？或是透過 Z39.50 連結的虛擬權威檔？或是僅進行單純的權威記錄交換？

為了便於彙集或交換各地區圖書館之權威記錄，目前使用之機讀權威記錄格式應如何配合修改？例如，是否在主標目、反見標目、反參見標目等處，以分欄註記語文、羅馬拼音系統、或是編目地區等等，以便於各地圖書館權威記錄之轉換？或是使用重複欄號？

爲了避免重複投入人力及經費，各地區權威記錄轉換程式如何互享？由於權威記錄之建立及運作與圖書館所使用之自動化系統有絕對的關聯，是否應結合各地區使用相同自動化系統之圖書館，共同向廠商提出需求？

雖然圖書館界都能瞭解權威控制的重要性，不可否認，在合作編目計畫中，多以彙整各館書目記錄建立聯合目錄爲第一優先，權威記錄的合作常被忽略，難免影響了合作編目的成效，也造成讀者檢索的困擾。兩岸三地及歐美東方圖書館界都期待建立中文善本書聯合目錄，或是華文書目資料庫，如果沒有互通的權威檔作爲基礎，無法克盡其功，權威檔的合作機制如何建立，實有賴各地圖書館界同道進一步的討論。

Chinese Name Authority Cooperative Program : Report of Taiwan's Experiences

Ming-der Wu, Heng-hsing Cheng

National Central Library and National Taiwan University Library started a name authority cooperative program in July 1998. It was intended to create authority files of Chinese personal and corporate names. Those files would be uploaded to NBINET and shared with other member libraries. Chinese Cataloging Rules and Chinese MARC Format for Authority Data were the standards used. Software was designed for the purposes of input, inquiry, and data transfer. There were several problems being discussed, such as, conversion between USMARC and Chinese MARC authority records, romanization, foreign names, and traditional Chinese characters and simplified Chinese characters. This paper also raised some questions about the cooperation among libraries in Taiwan, China, U.S., Hongkong, and other regions. Are we going to create a real union name authority database or a virtual one? How should the respective MARC format for authority data be revised? How should the conversion software be shared? How should libraries work together to negotiate with library automation system vendors?