



SCSL Newsletter

中国研究图书馆员学会通讯

学会主席新年致辞

新年伊始，本是祝之祷之、歌之舞之的喜庆时刻。可是，今年因新冠病毒笼罩全球，奥米克戎肆虐北美，令人心情沉重。新年同时带来新希望，新冠口服药问世，来自部分专家的分析，新冠病例已过高峰时期。期待今年是个吉祥好年，新冠疫情逐步缓和，最终成功控制。

回顾往昔，中国研究图书馆员学会(SCSL)走过了艰巨而不平凡的一年。2021年3月，在陈晰主席的主持下，成功召开了中国研究图书馆员学会网上年会。这是SCSL成立以来的第一次网上年会。总务委员会为这次会议成功做出了细致精心的安排。年会有一个重头节目是SCSL的十周年庆典。学会的五位前主席做了精彩的回顾和感言。学术委员会组织了精彩的学术报告和交流。



去年7月，每两年举行一次的第六届中美高校图书馆合作发展论坛以网上为主、实体为辅的形式顺利召开。在会议筹备及召开期间，本人亲历亲见了专业委员会所做的大量的重要的工作。这一论坛在疫情期间为中国图书馆界和北美中国研究图书馆员在专业和学术交流方面提供了一个重要的窗口。会议的两个主旨报告内容丰富，探讨了令人深思的课题。

本会信息委员会和CEAL中国资料委员会于2020年联系数据库供应商在疫情期间为北美图书馆免费开放了165种中国研究数据库。部分数据库延长至2021年6月结束。借此机会，感谢所有数据库供应商在疫情期间提供的重要而宝贵的支持。

有一个好消息和大家分享。去年4月底，执委会审阅并投票表决有关向国税局申请免税的议案。执委会全票通过。蒋树勇秘书代表学会递交了申请表。最近获悉，SCSL的申请已经获得IRS相关部门的批准。希望这一批准为本学会今后的发展在财务方面提供多种优势。

年会即将召开。这次年会和去年一样依然网上进行。目前，总务委员会正在统筹会议，学术委员会正在组织学术交流活动。今年也是学会的选举年，希望新老会员积极投入选举。

2022会员通讯即将出版。通讯的主题，庆祝图书馆界的一件大事：汉语拼音转换20周年。

在我们即将谢幕的时候，我想借此机会感谢执委会在去年一年做出的贡献；感谢所有会员对我的宝贵信任，感谢执委会对我的鼎力支持。

回顾历史，这十二年，我们的学会，筚路蓝缕，发展壮大，至今已成为一个具有相当规模的专业学会。我们的学会有很多特点。其中，我的体会，十二年来在学术专业领域，走过了一条求真求实的道路，贯穿着求真求实的精神。求真，是追求真理；求实，是尊重事实。祝愿我们的学会，继续发扬光大优良传统，取得更大的瞩目的成就！并祝愿新一届执委会，取得巨大的成功！

吴嘉勋

再会韦氏拼音，你好汉语拼音：纪念汉语拼音转换 20 周年

我们早已习惯使用汉语拼音进行编目、查找、和服务等一系列工作，可能已经忘记或者并不了解汉语拼音在北美图书馆推行和使用只有短短二十年时间。二十世纪五十年代末至千禧龙年之间，美国图书馆普遍使用韦氏音标对中文资料进行编目。自 2001 年 10 月 1 日北美各图书馆基本完成中文书目的汉语拼音转换以来，这套汉字罗马化系统极大地方便了用户，也一次次顺利地通过图书馆自动化管理系统更新的考验。在全球范围内，北美图书馆采用汉语拼音相对较晚。但是，美国国会图书馆 (LC)、北美东亚图书馆协会 (CEAL)、研究图书馆组织 (RLG)、联机计算机图书馆中心 (OCLC) 四方通力合作，协同各个图书馆在短时间内完成了上千万条数据的转换。今值汉语拼音转换 20 周年之际，我们采访了一些曾经参与转换工作的同仁，希望通过他们的回顾带我们重温二十年前那些重要的记忆。

SCSL 信息委员会

陈垚、甘露、李想、沈莉君、王成志

拼音转换大事记

19 世纪后期	剑桥大学汉学教授威妥玛 (Thomas Wade) 创建了威妥玛拼音。剑桥大学汉学教授翟理斯 (Herbert Giles) 将其扩充和修改，最终形成了韦氏音标体系。
1956 年	中国文字改革协会发表了《汉语拼音方案》。同年，国务院批准成立了汉语拼音方案审订委员会进行审议。
1957 年	美国图书馆开始使用韦氏音标作为汉字罗马化标准。
1958 年	全国人民代表大会通过了《汉语拼音方案》作为拼写汉语普通话的国家标准。
1977 年	联合国通过了《汉语拼音方案》作为中国地理名称罗马字母拼法的决议。
1979 年	LC 首次讨论用汉语拼音取代韦氏音标；联合国秘书处从 6 月 15 日起采用汉语拼音作为中国人名和地名罗马字母拼法标准。
1980 年	LC 宣布继续使用韦氏音标。
1982 年	《汉语拼音方案》成为拼写汉语的国际标准 (ISO-7098)。
1986 年	联合国正式开始使用汉语拼音。
1990-91 年	LC 再次探讨将韦氏音标转换成汉语拼音的可能性。
1996 年	澳大利亚国家图书馆开始利用计算机将韦氏音转换为汉语拼音。
1997 年	5 月，CEAL 成立了 Task Force on Pinyin Conversion，经过调研支持 LC 采用汉语拼音编目中文资料；11 月，LC 公布了《汉字罗马化指导方案》及时间表。
1998 年 5 月	CEAL 成立了 Pinyin Liaison Group，代表 CEAL 与 LC、RLG、OCLC 共同商讨实施拼音转换的相关事宜。
2000 年	4 月 1 日，北美图书馆开启汉语拼音编目准备工作； 10 月 1 日，北美图书馆正式开始使用汉语拼音进行编目工作。
2001 年	4 月 1 日，OCLC 和 RLG 完成了中文书目的汉语拼音转换； 10 月 1 日，北美图书馆完成中文书目的汉语拼音转换。

Philip Melzer

Then: Chief of the Asian and Middle Eastern Division, Library of Congress

Now: Retired

Q: Could you talk about the LC Romanization Guidelines? How did it come into the current form?

A: There was a group that worked on the guidelines and there was agreement from the beginning. We couldn't include tones. That was unrealistic and it was something very few library users would probably make use of. At the time, as I recall, guidelines for recognizing place names were kind of goofy. You romanize Chicago, chi ca go, and we couldn't do something like that. The big debate was over whether to aggregate syllables or to separate. We asked the library community for input. And, as with almost everything else involving romanization, there was disagreement. And we ended up having a big meeting involving people from different parts of the library, who were stakeholders. The discussion went around and around in the room. Finally, my boss, John Byrum, who was in charge of the meeting, said, "Wait, let's be clear on what we're doing here - What's the point of romanization?" The discussion turned to the purpose and use of romanization. It describes a language. It isn't the language itself. It isn't meant to be read, it's a tool for describing the written form or sound language or in most cases, both, a combination of the two. It's used in libraries to facilitate filing and retrieval. Then, the discussion turned to how we might build on the National Library of Australia's success converting via a computer program. How would we make use of existing Wade Giles romanization to do this? The idea of moving from separated syllables (except for proper names which were hyphenated at that time) to joined syllables was basically a nonstarter. We reflected on the fact that LC, and in some other libraries, there was an enormous range of Chinese materials in their catalogs, things going all the way back, as well as materials with technical terminology, and the technical terminology was changing frequently. The idea of putting together data dictionaries to represent from existing romanization what should be joined and what should be separated was simply unrealistic. Going from single syllable to single syllable was doable. Now, there were exceptions that we had to deal with -- we knew that there would still be complications, but that was the thrust of the discussion. And at the end, people around the room were nodding their heads, agreeing that that was what we were going to have to do.

In either maybe 1998, or early 1999, I had an occasion to meet with a representative from the National Library of China. I thought it would be a good opportunity to discuss romanization. The lady was very gracious and agreed to talk. I said we were moving toward converting to pinyin and we thought we could only do so by basically separating syllables except for certain proper names. I asked about the National Library of China's practice. And she said "Yes, we use romanization in our records. We separate all the syllables, all of them." I asked "Why do you do that?" and she said that romanization was just a tool for filing and retrieval. I thought that their conclusion was much like ours. There were going to be exceptions, of course, but they had the same understanding of what romanization was to be used for in a library setting.

Q: After LC decided to undertake the project in 1997, LC did a lot of work to plan and implement the project. Are there challenges or difficulties that particularly stand out for you?

A: Challenges and difficulties, I got a list of them! First, test files. One of the reasons the conversion was

successful was due to our coming up with two really good test files: one for authorities and one for bibliographic records. Those files really tested and strengthened the conversion programs. The RLG system and the OCLC system were completely different, so that they had to have two completely different approaches to how they were going to write and test their program. We at LC had to coordinate the writing and adjustment of the specifications, providing the Utilities with the same specifications to follow, so that they came up with same results. We worked with OCLC programmers to figure out a way to recognize Wade Giles romanization in the authority file, which are not identified by language, like bib records are. Bib records coded Chinese were converted. But in the name and subject authority files, identifying romanized Chinese was akin to finding a needle in a haystack. OCLC put together test files, I think we had nine of them. Beatrice Ohta and I then reviewed about 20,000 authority records in different batches, identifying Wade Giles text. Then once we identified Wade Giles, the challenge was how to identify the headings for people that should not be converted - those that were based on transcription and not romanization. We looked at thousands of 670 fields in authority records and evaluated how headings were chosen. Our work produced what we called the exclusion list -- something in the neighborhood of 100 headings -- which were not to be converted. The first step in all of the conversion process of heading employed the exclusion list to block conversion of those headings. The exclusion list proved to be successful, with only a couple of headings that were incorrectly converted.

Then, markers. There was agreement that markers were necessary in authority and bib records, so that one could tell that a record had been converted, and to prevent its being converted more than once. The markers also indicated text that was not converted, and why it was not converted. This made it possible to organize cleanup by manually converting unconverted headings, and then other access points, with less significant areas like notes left to the end. MARBI decided which fields should be used for the markers.

Something we debated at LC about was how to identify and not convert text that involved Wade Giles syllables and non-Wade Giles syllables, which we referred to as mixed text. We compiled a chart of syllables that were Wade Giles and corresponding pinyin, and then identified how each was to convert. Strings of text were divided by the conversion program by punctuation marks. If a string of text could convert cleanly, the program converted it. If a string included Wade Giles and non-Wade Giles syllables, it was not converted, and a notation was made in the marker.

We tried to accomplish some things that needed to be done with it, things that we couldn't get at in any other way at any other time. The name authority file had more than 100 of what were called conventional Chinese place names. They predated Wade Giles romanization. Almost all the provinces, and many of the major cities were represented by these old and disused headings, such as Canton and Kweichow province. The authority records were converted ahead of the machine conversion. Then, all of the headings were included in a data dictionary in the conversion program, so that headings for these names would be correctly converted during the machine conversion itself.

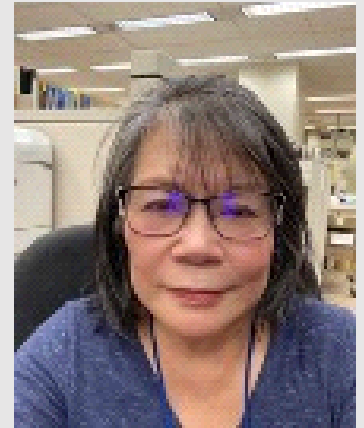
We also took the opportunity to make major changes for the classification of Chinese place names. A new class DS797 was established for places smaller than provinces. That portion of the schedule was then subdivided by province. The intention was to spread things out both in the catalog and on the shelves. After the conversion, catalogers were instructed to Cutter place names as they appeared. This led to unavoidable discrepancies on the shelf. There was also a decision that at least three numbers would be employed for Cutters to help alleviate crowding on the shelves and in the catalog.

Finally, following the conversion, we tried to identify and then correct various cataloging inconsistencies that had taken root over the years -- for example, with terms of address and titles attached to personal names, and confusion between jurisdictions and populated places. We also described how to adjust headings for jurisdictions and populated places in response to widespread changes taking place in China at that time. We drafted documents describing how to correct these inconsistencies, with helpful input from several CEAL members.

黃傳嘉 Cathy Yang

Then: Senior Cataloging Librarian, Library of Congress; Senior Cooperative Cataloging Librarian, Name Authority Cooperative Program (NACO), Library of Congress

Now: Retired



Q: 您在國會圖書館工作的經歷如何？

A: 1976 年我到國會圖書館任聯邦研究部的高級語言專員 (Senior Information Specialist), 1978 成為權部編目員, 1980 轉到採購部 (Acquisitions and Bibliographic Access Directorate), 從事與中國、台灣、香港、澳門的出版物交換工作, 並任國會館第一位中文採購館員 (Chinese Acquisition Librarian)。1983 我轉任該部中國組編目館員, 一直做到 2021 退休。其中 1999 到 2002 年, 我調到 PCC 任合作編目館員, 被任命為名稱權威合作編目項目的拼音協調員 (Pinyin Coordinator) 和中日韓文協調員 (CJK Coordinator)。

Q: 在國會圖書館, 您如何開始參與拼音轉換工作？

A: 我記得 1997 年 5 月 29 日, 國會圖書館與各大圖書館、電腦界、RLIN 和 OCLC 各方代表, 在國會圖書館開了一個 7 小時的會, 我當時做大會記錄。那個會叫 “Pinyin Project Meeting”。大家聚在一起, 探討拼音在圖書館界的應用、開發方針、實際操作評估、電腦相關配套, 提出方案及探討可行性。Philip Melzer 負責國會圖書館的拼音轉換項目(明確開始日期, 或如何開始)。我所略知, 無法細說或確認; 我只能就記憶所及, 略述我曾參與的事件。在這之前, 可能是 1991, 我花了兩個禮拜為國會圖書館起草共 18 頁的《拼音分詞指導方案》(Pinyin Word Division Guideline) 草案, 我研究中國的拼音法案和實際使用情況而寫出這個草案。在一般單音節詞間, 包括專業名詞中連字符的利用等, 不能照抄中國大陸的。簡言之, 就是如何在中文編目中規範漢字拼音分合標準, 有的字要合 (aggregate), 有的字要分。

Q: 1996 年澳大利亞國家圖書館開始利用計算機將韋氏拼音轉換為漢語拼音。您覺得這一事件的意義如何？

A: Philip Meltzer 邀請我參加審核澳大利亞國家圖書館的計算機轉換的主題詞部分 (Subject Headings) 其他的描述性質的 (descriptive) 部分我就不用管。我在每個澳大利亞的轉換後的 record 上, 主題詞部分, 分門別類, 草擬修正大綱。給他們提交總結和報告, 做了三頁總結意見, Philip 轉交澳方。此項工作, 我花了四五個月的時間來審核。

Q: 1997 年 CEAL 成立了拼音轉換工作組, 從 1999 年開始採用漢語拼音編目。您如何參與和推動？

A: 開始我與 CEAL 沒有直接接觸, 後來因為鄧石和 Sarah Elman 曾參與了我主持的國會圖書館的 “無差別人名的清理項目” (Clean-up of Undifferentiated Personal Names), 他們邀請我參加一些 CEAL

的拼音活動，我大部分都婉拒沒參與。後來我也成為 CEAL 的 RDA 小組成員。我們找包括哥倫比亞大學、康乃爾大學、普林斯頓大學和 UCLA 等 11 所東亞圖書館，確定 15 位編目員，一起做清理。這個項目就是將同名同音的不同人，區分出來。為每個人建一個獨立的個人權威檔：用生日或職業去分他們。否則中國有幾萬個人叫張三，如何區分他們？2000 年開始時，很多人開始做人名清理時學拼音，大家互相學習和共同研究。之後，國會圖書館編目政策和協作辦公室 (Cataloging Policy and Support Office) 的“地理名稱項目”(Geographical Name Project) 邀請我做“中國傳統地名拼音轉換項目”(Chinese Conventional Place Name Conversion to Pinyin Project) 我一個人做七、八個月，我改了 3000 多中國大陸地名。以前威妥瑪拼音用美國地名協會 (BGN, U.S. Board on Geographic Names), 是用傳統的形式 (conventional form)。拼音實行後，2003 年國會圖書館交給我的從威妥瑪拼音電腦自動轉換到拼音的記錄，但每個地名括號內還是英文 conventional form。這樣國會圖書館的編目數據庫裡不一致。我改了 3500 個地名，又得改書目記錄 (bibliographic records)。改了 Name Authority Records 加上相關的 bibliographic headings 到上千個 records 中，這是非常大的工程。

Q: 您參與 1997 年國會圖書館公佈的《漢字羅馬化指導方案》和時間表的制訂?

A: 時間表由 Philip 等人做的，但我參與方案的很多修正。CEAL 發表的文件中的聯繫人信息，NACO 拼音有問題的，請聯繫 Cathy Yang。但國會內部，有問題的問 Philip。Philip 參與多；CEAL 的拼音工作由他參與。與國會有關的，也邀請我開會。我負責的人名、地名兩大項目，我們沒有開會，但明確分工，效果很好。

Q: 現在您如何看拼音轉換項目的意義?

A: 第一，因多數人會用拼音，影響是正面的。以前有些反對的因素，現在不存在了。第二，威妥瑪拼音、注音等，有的圖書館沒有人力資源去轉換，這可能是個問題。圖書館人要多學其他系統，對增加知識和幫助讀者都有好處。Word division 是個潛在的問題。我們那時摸石頭過河。但拼音轉換項目很有歷史意義。也是一個必須做的工作，非做不可的。這是一代人花的心血，完成得很好！

蔣吳慶芬 Amy Tsiang



Then: Head, Richard C. Rudolph East Asian Library, University of California-Los Angeles

Now: Retired

Q: 1990 年美國國會圖書館再次探討將韋氏音標轉換成漢語拼音，當時反應如何？

A: 當時是反對的，沒有得到支持，主要是因為大館圖書多，那時沒有

online，都是手工操作，需要轉換的工作量太大。那時老師和學生到圖書館都需要用韋氏音標的卡片目錄查找圖書資料。Wade-Giles 拼音比較難懂，比如把 D 音當 T 音，高撇號、元音變音等讓他們容易糊塗，他們常常需要我們館員幫助才能找到資料；那時藏書不是特別多，很多時候他們乾脆直接到書架瀏覽查找。

Q: 1997 年 CEAL 成立拼音轉換工作組，您是成員，你們如何討論和準備轉換工作？

A: 因為 1996 年澳大利亞一下子就轉換 50 萬個記錄，引起國會圖書館和東亞圖書館界的特別重視。國會圖書館就發布了很重要的拼音分字的建議。那時我是 Committee on Technical Processing 成員，委員會主席要我協調並起草委員會對轉換的重要部分即拼音分字建議的意見。我於 1997 年 1 月郵寄出 80 份的調查問卷給 CEAL 成員，2 月 20 號收回 67 份。有的甚至包含一些圖書館員轉的教授的回復。我對回覆做綜合統計和分析。1997 年 3 月 CEAL 年會上，我做了報告 “Summary of the Survey on Pinyin Romanization”，6 月又在 JEAL 第 112 號上發表，可以看出贊成、反對、擔心及其原因等，比較具體。這也影響了國會圖書館的一些決定。

1997 年 5 月 CEAL 任命一個拼音轉換工作組，有八位成員。主要是要研究國會圖書館的轉換決定對東亞館的影響以及代表東亞館的觀點和利益，我代表 Committee on Technical Processing 成為八人小組的一員。我們主要是調查把拼音作為北美的羅馬化的標準，看到底能不能做到。工作組 1997 年 8 月為 CEAL Executive Committee 準備和提交了 Executive Summary，1998 年 3 月年會上做了正式報告。JEAL 發表了這些報告，還發表我們的 “Summary Report of the CEAL Task Force to Review a Possible Change from the Wade-Giles to the Pinyin Romanization System” (具體請看 JEAL 第 113 和 115 號)。1998 年 3 月 25 日在 CEAL 年會期間工作組還和國會圖書館、OCLC 以及 RLG 開了聯合會議，這個會我參加了。但 6 月 27 日在 ALA 年會期間又在華盛頓召開的拼音轉換的會議我沒能參加。

Q: 1998 年 CEAL 成立了拼音聯絡組，您如何參與和工作？

A: 我是聯絡組的六位成員之一，聯絡組代表 CEAL 與國會圖書館、OCLC 以及 RLG 聯絡和討論有關拼音的事宜。1999 年 1 月，我代表聯絡組就 CEAL 如何應對國會圖書館 2000 年進行拼音轉換的決定、如何計劃以及其他關心的問題做了另一項問卷調查。這次在 eastlib 上發布問卷內容，到 1999 年 3 月 5 日得到 48 份回覆，20 份代表 CEAL 成員館，27 份代表個人意見，1 份不明確。我在 1999 年 3 月的 CEAL 技術處理委員會年度會議上報告了調查結果。這次問卷調查的總結 “Summary of the Survey on Pinyin Conversion” 發表於 1999 年 6 月 JEAL 第 118 號。我對 CEAL 拼音轉換工作積極參與，也算是做了一點的事情吧。

Q: 這都是很重要的工作，也涉及和國會圖書館、OCLC 和 RLG 的合作，合作情況如何？是否有一些培訓？您作為館長如何帶領同事克服困難、按時間表完成轉換工作？

A: 可以說是合作無間。CEAL 成員館是由這幾個機構幫忙協助。因為 1990 年代的後期已有 online。RLG 最先轉換其數據庫裡的一些編目記錄；然後 OCLC 晚幾個月開始轉換。所以 CEAL 和這幾個機構一起開過數次會議和論壇，並確定轉換時間表。確定 Day 1 是 2000 年 10 月 1 日，之後原始編目必須用拼音；Day 2 是 2001 年 4 月 1 日，此後套錄編目必須使用拼音。有很多細節，各方一起討論和確定。至於培訓，主要還是國會圖書館領頭。我們是 RLG 或 OCLC 成員，實際做轉換時，他們與成員館有聯繫、有培訓。就我所知，CEAL 沒有舉行綜合性的集中培訓。

我們沒有新招工作人員。反正我們中國人韌性很大，該做什麼就去做。有什麼問題，就想方設法解決。有時可能會感到有些受挫，比如清理工作改正之後因 online 和機器或別的原因，感到又有不妥之處又得再做，很花時間。我們館轉換了 101,462 個目錄記錄，2002 完成。然後就開始清理工作到 2003 年的 6 月份完成，大概清理了 45,000 個記錄，我們手工一個個完成。那時我已不負責編目，我從 1998 年 5 月開始做代理館長、後來做館長；Sarah Elman 負責編目，很辛苦。她工作出色，帶領編目同事、學生工完成了這項工作。

Q: 您對拼音轉換項目的評價如何？下一個重大事件是如有同事預測東亞圖書館關門嗎？

A: 我覺得拼音轉換總體上來說是好的；是東亞圖書館發展史上大事件；是歷史的趨勢和技術發展以及館員同事努力的自然結果。拼音轉換之後不僅資料與總館整合一起，而且東亞館的人員和服務也從以前的分離和獨立到與總館密切聯繫和融合一起。做拼音轉換那個時候我很開心，其中幾個重要的事情，我都有幸參與。雖然當時工作比較辛苦，但我很高興做了貢獻。但我堅決不同意東亞圖書館會關門。流通部門和技術部門與總館整合一起是趨勢，而且或許有好處。可是中日韓等東亞圖書館的館藏有其特殊性，館員的參考諮詢工作和作用誰也代替不了。所以我相信目前的東亞圖書館永遠不會關門，你不用擔心。

Martin Heijdra



Then: Chinese Studies Librarian, Princeton University

Now: Director of East Asian Library, Princeton University

Coming from Europe, Martin was rather surprised to learn that the US was still using the Wade-Giles romanization system for Chinese. Multilingualism is such a founding principle in Europe that adopting Pinyin became very natural because it was the national standard used in China. In contrast, the adoption of the Pinyin system was very much a political issue in America. Back then, most librarians of East Asian descent were from Taiwan, who strongly believed that the library should not use the standard from Communist China. The same was true for the use of simplified and traditional Chinese. Martin reflected it would have been a great opportunity to switch to the Pinyin system when the first Chinese online cataloging started in 1983.

What puzzled Martin was that LC didn't plan to use Pinyin the way scholars were using it. Over the years, Martin has learned to understand some of the reasons, but still questions the reasons behind creating a localized version of the Pinyin system. "It's like creating a US version of the Chinese Pinyin standard. This would not help scholars," said Martin. Many times the reason users can't find what they're looking for is that they are not using the LC Pinyin rules in their searches. Word division has always been an issue.

Should it be Suiyuan or Sui yuan when searching 随园食单? Catalogers don't always catalog consistently, which adds another layer of barriers.



BOOK

随园食单 / 袁枚著; [王英中标点; 王英志校订]; [Sui yuan shi dan](#) / [Yuan Mei zhu](#); [Wang Yingzhong biao dian; Wang Yingzhi jiao ding].

[Yuan, Mei](#), 1716-1798.; [袁枚](#), 1716-1798.; [Wang, Yingzhong.](#); [Wang, Yingzhi.](#); [王英中.](#); [王英志.](#)
Nanjing [Shi](#): Jiangsu gu ji chu ban she; 南京市: 江苏古籍出版社; 2000



BOOK

随园食单 / 袁枚著; 别曦注译.; [Suiyuan shi dan](#) / [Yuan Mei zhu](#); [Bei Xi zhu yi.](#)

[Yuan, Mei](#), 1716-1798.; [袁枚](#), 1716-1798.; [Bei, Xi.](#); [别曦.](#)

[Xi'an Shi](#): San Qin chu ban she; 西安市: 三秦出版社; 2005

The conversion to the Pinyin system has made searching easier and was welcomed by users, although with limitations. The bigger question is how to reach out to users who don't know everything about searching. Such teaching opportunities have become rare. Martin believes that there are technical solutions that can solve some lasting issues, but he doesn't see systems have much willingness to develop them.

Two major purposes of library catalogs are to facilitate searching and provide authoritative records for citation. Unfortunately, users still find it challenging to search their library catalogs for Chinese language materials, let alone that they are able to directly use the bibliographic information from their library catalogs. This inconvenience unavoidably forces scholars to manually change catalog records to generate bibliographies. This is another reason why romanization remains essential; without it, scholars would need to go to other sources to generate bibliographies or find materials cited in other scholars' bibliographies. Thus, a major link between scholarship and libraries would be lost and broken. Catalogs would lose their purpose as a scholarly necessary tool.

With several updates, the current ALA-LC Chinese Romanization Table still has rules that could be further improved, such as word division, jurisdiction, the order of last names, and so on. These issues are not unique to the East Asian librarian and scholar community and apply for other area studies librarian and scholar communities. Searching in non-Roman scripts isn't an unsolvable mystery. Google successfully manages to find ways to build its search engine that works well for non-Roman languages. If this type of project were user-driven from the beginning, our catalogs will be much more useful for scholars. If integrated library system companies develop their systems with multilingual users in mind, many of the issues we are experiencing today would have been avoided.

Martin encourages public service librarians to be outspoken and fight for their users. East Asian studies librarians need to collaborate with other area studies librarians to challenge the monolingual-minded system. At the same time, it should be realized that different scripts and different languages are not all the same either: even Chinese and Japanese have different challenges. Unfortunately, library systems just think that limited use of Unicode is all that's needed. Each language and script need their specific adjustments and additional solutions for searching.

周欣平 Peter Zhou



Then: Head of East Asian Library, University of Pittsburgh

Now: Assistant University Librarian and Director of the C. V. Starr East Asian Library, University of California, Berkeley

Q: 1979 年国会图书馆 (以下简称 LC) 首次讨论用汉语拼音取代韦氏音标但并没有得到北美东亚图书馆协会 (以下简称 CEAL) 的支持。您能否谈谈这段历史?

A: 1957 年前汉字罗马化没有统一标准, 韦氏音标是西方最常用的一套系统。1958 年, 中国政府颁布了《汉语拼音方案》, 但在中国之外, 大多数人使用的还是韦氏音标。联合国 1979 年通过了使用汉语拼音的决议, 随即 LC 便提出要开始采用汉语拼音, 但受到了 CEAL 的反对。1992 年到 1996 年间, CEAL 年会上每年都有相关讨论, 但每次听到的多为一边倒的反对使用汉语拼音的意见, 理由是转换工作量太大, 没有经费和人员等等。在我看来, 当时很多反对的声音中隐隐约约也有一种政治和文化的看法, 并不是完全出于技术和经费方面的考虑。

Q: 1997 年, CEAL 成立了拼音工作组。作为成员之一, 请分享一下小组的工作。

A: 工作组成员由 CEAL 各个委员会各派一名代表, AAS 派一名代表, 还有两个普通成员组成。因为我有语言学背景, 而且代表了赞同派的意见, 我就成为了两个普通成员中的一员。工作组的任务是讨论 CEAL 如何对 LC 的决定作出反应, 如何找出问题, 以及如何启用汉语拼音。工作组的基本立场是支持采用拼音, 但也提出了保留意见和相关议题。比如, 是否可以用转换软件来进行介面转换而不放弃韦氏音标, 如果全面转换并启用汉语拼音, 是用中国政府颁布的《汉语拼音方案》标准还是由北美地区的图书馆自己定一个标准, 而这个自己定的标准能否解决图书馆的问题同时又具有科学性和合理性等等问题。我在语言学方面的训练对探讨这些问题都起到了一定的帮助。

Q: 1998 年, CEAL 又成立了拼音联络小组。作为组长, 请分享一下小组的工作。

A: 联络小组的工作重心这时已经转到如何做汉语拼音转换, 解决程序上的问题, 协调出一个时间表和路线图, 并协助各东亚图书馆做出相应的对策上来了。小组在 1998 年到 2000 年间非常忙碌, 要和 LC、OCLC、RLG 等方面合作协调。

小组的第一项任务是对国会图书馆提出的《汉字罗马化指导草案》提出修改意见。第二项任务是制定一个拼音转化工作的时间表。我们和各方一起确定了 Day 1 (即开始用拼音进行编目的时间) 和 Day 2 (即开始回溯转换韦氏音标的编目数据) 的具体日期。这两个时间不能隔太久, 否则就会出系统上的混乱。

Statistical Report on Survey on Pinyin Conversion

** The questionnaire was distributed through eastlib on February 26, 1999
** 48 responses received by March 5, 1999

This survey is completed to represent the views of (Please select one)

(20) an institution * (1) No indication
(27) an individual

I. Preparation and implementation plans for Pinyin conversion

1. Will your library start to use Pinyin for the romanization of Chinese after LC's lead in 2000?

(27) Yes
(15) Most likely
(1) Most unlikely
(0) No
(5) Not sure at this time

2. Does your library plan to convert existing files on your local system to Pinyin after LC's conversion in 2000?

(13) Yes
(14) Most likely
(2) Most unlikely
(0) No
(19) Not sure at this time

Final Report on Pinyin Conversion

By

Pinyin Liaison Group
Council on East Asian Libraries
(CEAL)

March 2000

Susie Cheng, University of Hawaii
Yu-lan Chou, University of California at Berkeley
Guo-qing Li, Ohio State University
James Lin, Harvard University
Amy Tsiang, UCLA
Peter Zhou (Chair), University of Pittsburgh



Preamble

For nearly half a century, libraries in North America have been using Wade-Giles romanization for cataloging Chinese language materials. In 1997, the Library of Congress (LC) announced a decision to switch to Pinyin for the romanization of Chinese in cataloging and authority records. Pinyin is a romanization scheme widely used by governments, educational institutions, commercial publishers and news media in the Western world for transliterating Chinese scripts. Pinyin conversion, now scheduled to occur in 2000, will bring about systematic changes in millions of records, both those in Chinese and those in other languages that have headings, notes, or title added entries in Wade-Giles. This will be the single largest conversion of romanization systems in the history of American libraries to date. This report discusses issues related to the change to Pinyin in North American libraries and recommends the necessary steps libraries should take to adopt the use of Pinyin and to convert their cataloging records, following LC's lead. This document is in the public domain, free of copyright restrictions. Therefore, we encourage CEAL members to adapt this report or change it into their own planning document in local deliberations on Pinyin conversion.

就这样拼音的启用和数据的回溯转换工作从 2000 年开始，到 2001 年结束，北美众多的图书馆齐心协力完成了上千万条编目数据的转换，非常高效。第三项任务是后期清扫工作，确保数据转换完整且不留后患。

Q: 如何看待 LC 的《汉字罗马化指导方案》和学术写作中的汉语拼音使用的不同？

A: 两者之间最重要的区别是音节是否连写。从教学和科研角度来说，一定要注重汉字音节的连写问题。比如，shouyinji (收音机) 是一个完整的词，不能写成 shou yin ji。联络小组、LC、OCLC 和 RLG 做出这个决定主要有两方面考量。一是方便用户。大多数用户，尤其是汉语初学者，其实他们并不能准确地判断词和字之间的区别。音节不连写就避免了词字错分的问题。这种差异是我们有意设定的，也是合理的。读者不具备良好的汉语拼音正词法知识也可以使用图书馆系统查找文献。二是考虑到编目人员操作起来方便。其实编目人员真正受过严格训练的也不多，在编目过程中能正确地将字词分开的人越来越少。

Q: 请您谈谈拼音转换项目的意义。您如何评价当时的工作？

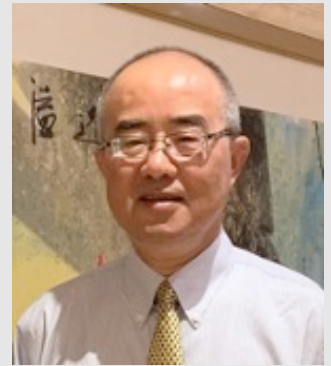
A: 2000 年的这次北美地区图书馆的汉语拼音转化工作对学界影响很大。令人欣慰的是转化工作完成后读者的反馈非常好。从图书馆的角度和以方便读者作为标准来说，《汉字罗马化指导方案》和这次韦氏音标向拼音的系统转换工作没有发生什么大的问题。我也没有听到任何批评的声音，这是大家共同努力做成的一件惠及学界的了不起的大好事。当然，从学术的角度来讲，这个方案是有不对的地方的，但是有些“错误”是有意设置成的，是美妙的“错误”。

回想起来，这次的汉语拼音转换过程计划非常详尽，思考到了所有可能出现的问题，也吸取了各方专家的意见，工作是比较完美的。回头看没有绝对遗憾的地方。在短时间内，各方共同协作完成了这么一个大项目，整个过程非常民主。参与项目各方之间互相协调的精神是非常难得的，很值得后人回顾。

李国庆 Guoqing Li

Then: Chinese Studies Librarian, Ohio State University

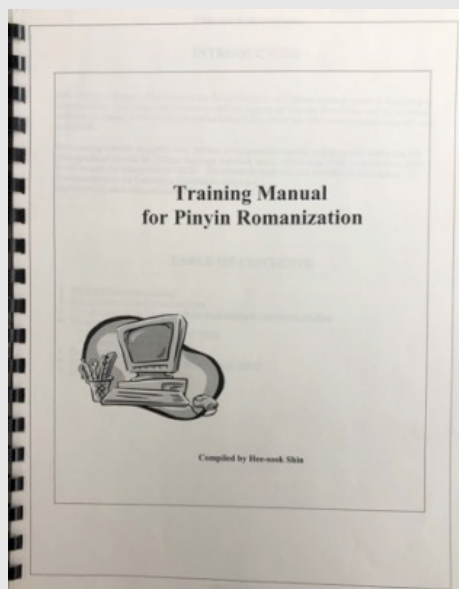
Now: Chinese Studies Librarian, Ohio State University



Q: 您提到，在 1997 年之前同行们讨论的焦点主要是要不要进行拼音转换的问题，您能讲讲当时争论的情况吗？当时大家各自所持的观点是什么？

A: 比较客观的呢，主要考虑的是技术问题。因为当时的技术还没有像现在的电子计算机技术这么发达。从具体的实践的人来说，这件事好像是很难做的。你可以想想有多少记录嘛，如果叫人手动的来改，这件事你说要改多少时候！所以说有人说，这简直是不可能啊。这是从技术上来考虑的。当然从感情上考虑，大陆来的人说，现在中国是联合国的常任理事国，中文语言是官方语言。拼音已经是除了我们图书馆系统之外都在使用的了。新闻报道啊等等全部都是拼音了。为什么美国国会图书馆坚持还要用韦氏？这是带有一些这样偏激的看法。然后呢，第一次成立那个 Task Force 的时候，里面只有 Peter 一个人是大陆来的，负责的人是马泰来那批老先生们，当然他们也并不是完全反对，他们就跟我说的那样，觉得这件事在早期的时候是非常难做的一件事。然后新来的这批人，那个时候有一个人叫温时幸，现在你们可能都不大知道了，他当时是非常活跃的，他当初是 Michigan 的编目，他那时候还是在 CALA 非常积极，做过 CALA 的主席。CALA 也同时在讨论，周幼兰、温时幸 Amy Seetoo 等都积极参与。所以温时幸代表着这一批人，他就觉得我们这些东亚图书馆的人没有考虑到在其他公共图书馆做的那批人。然后还有一方呢，比如 OCLC Users Group 也在讨论，也有他们自己的意见。慢慢到了后来的时候，像 Karl Lo 那个时候在 San Diego 做馆长，后来到国会图书馆去，已经故世的这位先生，他在技术上比较先进，他还跟人一块研究了一种怎么样来转换的小程序，那个时候算比较早的。所以这个人也是一个积极的推动者。所以等到技术成熟，觉得可以用计算机来进行转换的话，事情其实就解决了一大半了。

Q: 各个东亚馆是怎么进行转换的？这个具体的过程是怎样的呢？



A: 等到决定了以后，各个馆就是根据自己的情况来安排。我想大致的都是这样的。就是我们知道要做了，自己订一个计划，基本上每个馆都会找几个人做一个工作组，审视一下情况，觉得应该怎么做，需要多长的时间。我们馆就成立这么一个小组，由他们制定一个方案，编了一个 Training Manual。那个 manual 其实基本上就是照 Library of Congress 那个指导方案来的。然后就具体的执行。时间基本上在一年左右。其实转换基本上都是委托，像我们是 OCLC 的成员，就委托 OCLC 给我们做回溯转换，我们自己就是从哪一天开始就执行新的方案，所有的编目都用拼音。有一段时间是两种格式共存，后来的就没有了。虽然比如说到了 2001 年我们已经说全部都做完了，实际上后来又不断的在捡漏，因为他们用计算机做的那个转换还是有很多遗漏的。一种是我们自己系统的所有记录给了他们或者在他们那里保留的，它整个转换还会遗漏一部分条目，就是说并不是说我给你一百个条目，一百个你都给我转换了，有漏掉的。另外一种具体的每一个

条目里面的那个内容它也有漏掉的或者错误，所以不断地在修改。我好像觉得前两天在做古籍编目的时候，居然还发现有的记录就没有转换过。本来理想的状态就是所有的东西都已经是改换过了。

Q: 您记得您馆里大概多少条目需要做回溯，大概在什么量级吗？花费又是怎样的呢？

A: 我保存的报告里说，“OSU Libraries will use OCLC service。”然后有日期，他们说好了是从2000年10月1号开始做的，我们叫day one。从那天开始改。我们总共的记录是有八万四千四百零二个条目。八万不算多，我们只是中型的吧。我们在CEAL里面排在十七、八位。我这具体数字都有，多少是书，多少是microfiche, serial。然后记录说，“OCLC will charge 12 cents with a discount for larger file。”然后呢，“authority record with no charge”，所以总共的花费是一万零一百二十八点二四美元。但是他也说了“This cost is subject to increase based on the number of records needed to be converted。”大约你就知道，八万条一万块，他说是十二分钱一个。

Q: 拼音转换项目20年后，您现在怎么评价当时的工作、过程和结果？

A: 做当时认为是正确的事而已。我觉得当时的做法好像做的还是比较务实的，我没有觉得有什么要改的。因为那时候技术也就是那个样子，已经比他们当初讨论的时候已经顺利了，就是说这些两大数据商，OCLC和RLG都可以做machine convert。开始的时候是Karl Lo自己发明的一个小程序，提供给大家，要自己把这些东西进行转换。有些人还是用的。现在当然在网上转换很容易了。

I published a pinyin-to-Wade/Giles algorithm 計算程式 in the Information Technology and Libraries in 1991, and provided a revised version to OCLC in 1993 and still another revision to Australia in 1996. Both OCLC and Australia have implement the algorithms with great success. The new interface mentioned above is still another revision, tentatively named T42. ↵

This interface program is a dialog box run on either MS Windows95/NT or Windows 3.1. ↵

THE USER needs to input a pinyin or Wade-Giles string from the keyboard, or paste a string from another Windows program into the dialog box. ↵

THE PROGRAM will convert the string, including dividing aggregated words, and display the result in the box. Simultaneously, the converted result is ready for pasting into an OPAC, a word processor or some other Windows programs. ↵

THE PROTOTYPE, T42, that was developed recently will be demonstrated during the AAS meetings. A fuller version will be available within a few weeks. ↵

Your comments and suggestions will be most welcome. ↵

Karl Lo ↵

Q: 您刚才提到很重要的一点就是说拼音转换项目把大家都凝聚起来，现在也不见得有这样好的项目，这样好的话题了。那么您自己是不是有的时候也在思考啊观察啊，不久的将来是不是会有一个比较大的项目，一个共同的话题，像这样子再把大家凝聚在一起做事情？

A: 这个问题，如果有可能的话，我想就是东亚图书馆还要不要存在的问题，这个是这个group以后会要讨论的问题。因为这个趋势是现在越来越明显了，比如说Karen Wei, 蒋树勇那个地方原来是Asian Library现在没有啦，匹兹堡是有个East Asian Library，人也越来越少。原来有那个郭先生跟他们争取，他还写过一本回忆录，叫《学府鏖战录》，真的是鏖战，要保留这个东亚图书馆。越来越多的东亚图书馆在名义上已经取消掉了，所以说下一个大的话题很可能就是，在这样的形势下东亚图书馆是不是要作为一个单独的unit存在，这恐怕是下一个在CEAL引起轰动的大事件。已经有端倪了，我们已经有好几次已经在争论那个关于排架的问题，东亚的书是不是要单独的排架。这是这些年曾经讨论过的，这个是跟东亚图书馆是不是单独存在相关的一个事情。在讲排架的时候也会引起大家很多的意见。所以我觉得这是下一个热点。从取消单独排架开始，是不是要取消单独的东亚图书馆？我个人觉得以后除了几个大的以外，我们都要消亡了。

周幼蘭 Yu-Lan Chou



Then: Head, Technical Services, East Asian Library, University of California, Berkeley

Now: Program Coordinator – Technical Services, Santa Clara City Library; Chair, Library and Education Advisory Commission, City of Milpitas; Co-Chair, ALA Public Library Technical Services Interest Group

Q: 1979 年美國國會圖書館首次討論用漢語拼音取代韋氏音標，當時為什麼 CEAL 不支持轉換？

A: 我對 1979 年的討論不瞭解，但是一直到 1997 年美國國會圖書館正式發佈新聞稿要改用漢語拼音時，美國東亞圖書館界面臨的挑戰是一樣的。那就是舊資料的轉換。舉例來說，1979 年的時候書目資料沒有上綫，可能每本書平均要有五張卡片排序 (Title 題名, Author 作者, 2-3 張 Subject Heading 主題標目, 還有可能有 Series 叢書項)。也就是說如果要從韋氏拼音改成漢語拼音，每一本書可能要重做 Title, Author, Series 重新排序。工作量太大了，費用也是問題。

Q: 1990 年國會圖書館再次探討將韋氏音標轉換成漢語拼音，您怎麼看圖書館仍使用韋氏音標？

A: 韋氏音標為美國學界使用多年，學界裏用那一種羅馬拼音都可行。只要熟悉一種羅馬拼音，參照對照表很快就可以上手。圖書館使用羅馬拼音是因為要排序，當時還沒有 Unicode Non-Roman 字符尤其是在中日韓文中共同使用的參與排序有一定困難。漢語拼音作為國際標準，聯合國和其他機構已經行之有年。繼續韋氏拼音系統在當時圖書館造成一名多表，參見 (see), 互見 (see also) 不一致的情況。對於以漢語為母語的圖書館使用者尤為不便。有些人說在台灣使用韋氏音標，也是誤傳。那時雖然申請護照時要提供羅馬拼音，但只是形似而不是全面運用。例如韋氏音標有送氣和不送氣之分，護照上的人名並不標明。一般人申請護照時經常是被代辦的旅行社給取名了，錯誤是很常見的。後來馬英九執政時期，台灣人的護照名採取漢語拼音。另外在中國普遍使用的漢語拼音也有很多不同的使用方法。有的在母音上加上調號，有的以詞為單位，有的以字為單位。不一而足。這些不方便，對於受過學術訓練的人是很容易解決的。但是漢語母語使用者反而覺得困難重重。

Q: 1996 年澳大利亞國家圖書館利用計算機將韋氏音轉換為漢語拼音，意義如何？

A: 當年澳洲國家圖書館的 Pinyin 轉換計劃，使用了 Karl Lo 盧國邦館長（當時在 UC San Diego 任職）用 C 程式語言所創建的轉換方法。約五十萬筆書目資料從韋氏拼音轉成漢語拼音。盧先生是 CEAL 先驅，他從 70 年代東亞圖書館界普遍討論自動化時就在研究這個問題了。RLIN 的 CJK 系統是 1983 年上綫的，OCLC CJK 是 1991 年。澳洲在 1995-1996 要將 CJK 資料上綫，使用了自動化手段為主，手動清錯為輔的辦法，完成漢語拼音轉換。這個計劃給了北美東亞圖書館很多啟發，證諸可行性。

Q: 您如何看 1997 年美國國會圖書館公佈了《漢字羅馬化指導方案》和時間表？

A: 我當時負責 UC Berkeley 東亞圖書館技術部門，漢字羅馬化對編目和期刊的工作流程影響甚鉅。國會圖書館在 1979-1980 年 RLIN CJK 還沒出來的時候就已經認識到漢語拼音將漸會是主流，所以轉換只是技術層面和時間問題。對於指導方案，字詞區隔比較有爭議。因為字詞區隔會影響到系統排序和檢索。RLIN CJK 系統原來就以詞為本，而不是以字為本。好處是使用這套系統從羅馬拼音多半就可以理解書目題名短句，壞處是漢語字詞分界不明，斷詞處不一，會影響到排序。指導方案最終

決定的是人名和地名採用以詞為本。其他以字為本。

Q: 請您談談您為成員的 CEAL 拼音聯絡組如何開展工作的?

A: CEAL 在 1998 年成立拼音聯絡組，東亞圖書館的兩大編目平臺 RLIN 和 OCLC 也積極參與其中。在國會圖書館漢語拼音轉換計劃的大框架之下，徵詢各館意見，制定轉換的規則。那時沒有視訊會議，小組之間聯絡主要靠 Email。1997 之後一直到轉換開始，工作小組的主要工作是調研和與國會圖書館討論轉換規則。這中間包含向澳洲取經。除了澳洲圖書館，香港的學術圖書館例如香港科技大學，從 1990 年底開始使用 Innovative Interfaces 的 Innopac 也早有內建的字符對普通話拼音的機器自動轉換，以解決編目館員日常使用粵語發音的問題。這些機器轉換而不是人工處理，確認了計劃的可行性。國會圖書館主導這個計劃，確定轉換規則。2000 年 10 月 1 日起，國會圖書館正式開始使用漢語拼音編目中文書籍。一個新的時代就此開始。我在 2001 年 7 月從東亞圖書館離開，OCLC CJK 資料庫轉換那時剛完成；cleanup 和人工清錯都還沒有開始，未能躬逢其盛，應該算是遺憾了。

Q: LC 漢字羅馬化和學術寫作中的漢字羅馬化不一樣？怎麼處理這種不同，如何幫助讀者？

A: 這也是目前還存在的最大挑戰。LC 的非羅馬拼音語言的羅馬化，有些採行的不是母語使用者熟悉的系統，這不是只有中文的問題。雖然採用的系統有學理支持，在高等教育機構圖書館不是問題但是要廣大的母語使用者了解這些規則，有一定困難。我目前在公共圖書館服務，社區群眾很難明白像“匪我思存”這種作家，人家 wiki 都是分開寫了，為什麼圖書館要把作家的筆名連在一起，大小寫的規範也不一樣。中學和社區學院的漢語課程讀本，也和圖書館使用的字詞分類規則不同。令人無所適從。所幸的是，這些漢語拼音的困難處，多在字詞分類。如果需要，在 added entry 加註就可以，雖然有點畫蛇添足就是。

Q: 您對目前的漢字羅馬化有什麼建議？

A: 幾十年前為了排序，羅馬化很重要。但是現在任何元數據都可以解決字符的問題。圖書館也可以直接用 Unicode 字符來做檢索。繼續手動輸入漢語拼音的平行項完全沒有必要，不僅如此，還常常看見各式各樣的錯誤，影響到檢索的正確性。建議把編輯中文資料書目漢字羅馬化的工作由機器自動產生。如果可能，也不用儲存在書目裏。On the fly 即可。Google Translate 已經實現這些很多年了：Kong zi bu xiang zhi dao 英語是 Confucius doesn't want to know。

邓石 Shi Deng

Then: Head, CJK Acquisitions & Cataloging Unit, University of California-San Diego Libraries

Now: Head, CJK Metadata Unit, University of California-San Diego Library, & Head, Shared Cataloging Program, California Digital Library



Q: 加州大学圣地亚哥分校 (UCSD) 拼音转换涉及哪些部门和人员？

A: 由于 UCSD 图书馆决定使用 OCLC 拼音转换服务，工作涉及从系统输出编目数据，再将转换后的数据输入并替换原有数据，因此参与拼音转换的部门和同事挺多的，包括编目部主管、编目部系统协调员、总馆负责系统的技术人员。编目部的副主任负责和 OCLC 联络及馆内沟通工作。我们中日

韩采访编目组更是所有人都作出了贡献。前期要做许多具体准备，如测试本地数据、遴选样例、准备转换数据等，整体工作量还是很大的。等 OCLC 的转换数据送回来以后，我馆要输入替换原有数据，然后开始清理本地数据。另外，总馆编目部负责规范档的同事协助我们完成规范档的清理；编目部系统协调员和总馆负责系统管理的技术人员这两位同事一直参与这个项目，帮助我们完成了所有数据输入和输出。

Q: UCSD 图书馆本地拼音转换的大致过程如何？

A: 本馆的转换基本按照拼音转换计划时间表推进。2000 年 10 月 1 日是拼音转换的 Day 1，从这天起，所有新添的中文编目记录开始使用拼音，同时既存的韦氏编目记录等待转换。当时新添的拼音记录按要求都要加 MARC 987 字段标识拼音，以此区别系统内的旧记录。

2000 年 10 月 1 日以后，UCSD 图书馆开始进入韦氏到拼音的过渡期，我们着手准备转换既存的韦氏拼音记录。UCSD 共有大约五万条既存记录，其中近四万是中文资源目录，其他是非中文资源的记录。我 2001 年 7 月入馆时，本馆的项目筹备已经展开。我们用的是 OCLC 拼音转换服务，按计划指定的那一天把本地系统所有记录做了数据截图送交 OCLC，OCLC 再用它的自动系统识别、转换含韦氏拼音的记录——其中既有中文记录，也有非中文记录。为避免本地数据出现前后不一致，UCSD 图书馆于 2002 年 4 至 6 月期间约两个月时间完全暂停中文编目，即从数据截图送至 OCLC 之日起，直到 OCLC 的转换数据送回来匹配并覆盖本系统内的原有数据为止。这之后，我们开始根据 OCLC 提供的报告对转换数据进行人工清理，UCSD 大约在 2003 年 7 月才基本完成人工清理。

Q: 人工数据清理具体是怎么运作的呢？

A: OCLC 送返数据时给我们提供了一份报告，列出已经发现的一些需要人工复查的问题，如多音字和地名的问题。这个报告是我们后续清理的主要依据。此外，OCLC 系统在转换过程中如果碰到不确定的字段，会在 987 的 f 子字段中给出提示，这也是后续人工清理的参考信息。经过了人工复核或修改的记录，我们会按要求删除 987 的 f 子字段，再加另一个本地字段标识。我们临时雇佣的一位同事（当时我部门刚毕业的学生工）承担了本地后续清理的主要工作，记得她修改了八千多条的书目记录和三千多条丛书规范记录，大约占整体机器转换记录的 17%。还有一些“漏网之鱼”后来也被我们陆续逮到并修正了。总体上，我觉得拼音转换还是很成功的，结果让人满意。

Q: 拼音转换中遇到什么特别的挑战吗？

A: 印象最深刻的挑战是 UCSD 艺术图书馆一个有 30 万条记录的幻灯片特藏，这个特藏的记录完全是 UCSD 本地记录，里面关于中国的信息非常丰富，专有名词既有用韦氏拼音的，也有用汉语拼音的，所以形成一个异常“混杂”的本地资源。我们非常担心这批记录在转换时被搞混，甚至发生本地数据丢失。拼音转换的难题之一就是此类“混杂”资源可能引发所谓的“双重转换”（double conversion）——就是，本来已经使用拼音的条目因系统局限而无法被有效识别，从而导致被再次转换，结果，数据可能乱套，甚至牛头不对马嘴。因为这批幻灯片记录，UCSD 很早就和 OCLC 沟通。OCLC 还破例为我们做了样例测试，结果发现很多问题，导致我们最后决定放弃转换这批记录。现在回想，我觉得这个处理还是很“英明”的。除此，其他馆藏资料，像乐谱和 CD 等，样例转换都没有特别的问题。本馆的拼音转换总体还是挺顺利的。

Q: 能评述一下您对北美图书馆中文编目的未来展望和挑战吗？

A: 信息技术发展是我们的机遇，也是挑战。现在，关联数据（linked data）已经发展起来。图书馆可利用关联数据、人工智能等技术来预测、疏通、连接不同的罗马化标准和不同语言的编目，这可以

大大提高编目的灵活性和可用性，让检索和书目发现变得越来越容易。所以随着信息技术进步，我觉得，标准的殊异尽管还会是个问题，但将不再像过去那样影响巨大了。

如何利用信息技术也是对图书馆的挑战。关联数据、人工智能带来了许多可能性，这些技术一出现其他行业往往能够迅速跟进，比如 Google 就大量利用关联数据，让搜索变得简单容易。相比之下，我觉得我们这个行业可能比较受限于自己的行事模式，尽管图书馆在努力地采纳新技术，但是进展却很缓慢。所以，我们需要自我反省并寻找突破口，借鉴其它行业的先进经验，并且寻求跨界合作。近一两年北美合作编目项目 (Program for Cooperative Cataloging) 的“维基数据试点项目”(WikiProject PCC Wikidata Pilot) 令人兴奋，希望图书馆行业有更多类似的项目出现，促进了书目信息关联数据化和网络化。

蔡素娥 Sarah Elman

Then: Head of Cataloging (1999-2000), Head of Cataloging and Public Services (2000-2002), East Asian Library, University of California-Los Angeles

Now: Retired



Q: 當時在您的館裏，大概有多少人參與進來呢？

A: 拼音轉換之前的準備期，Amy Tsiang 代表 UCLA 先後參與 CEAL 的拼音轉換工作組及聯絡組。Amy 升任館長後，我接管了編目部門。那時專門做中文編目的 professional librarian 只有我。此外，還有 library assistants。我們還有學生助理，做一些比較簡單的工作。為節省人力，我盡量把一些可以做 batch processing 的問題找出來，再跟總館系統部門 (Library Systems Office) 的同事討論是否可以用程式更改。我們也請總館的 Authority Control Team 幫忙做一些事情。比如「的」這個字，Wade-Giles 都用 ti (di) 這個發音。其實這個字發 di 的時候比較少，例如「目的」、「標的」等，其他時候都是 de。因為記錄太多了，最後決定把有中文「的」字的記錄全找出來，batch-processed 之後，再把「目的」、「標的」等手工改為 di。諸如此類的 cleanup 就跟總館系統部門的人配合。其它他們沒辦法幫忙的就手工做。

Q: 那個時候總館對這個項目是支持的嗎？他們是否新招了工作人員？有沒有什麼財政上的支持？

A: 因為這是全國性的一個重要項目，而且能改善圖書館的服務品質，所以總館還蠻重視的，有財政支持以支付轉換費用，但沒有聘用新的工作人員。我那時會跟系統部門同事，以及總館和其他分館的編目主任討論相關的事情，他們都很支持。其他分館，如法律圖書館，醫學圖書館，科技圖書館等，都有些中文藏書，他們也很高興我們解決了這件他們自己無法處理的事。因此我們得到不少 moral support。

Q: 那您館裏進行拼音轉換，需不需要有什麼培訓？

A: 我們內部有自己做培訓。針對圖書館工作人員的，除了開會說明這項計劃外，也發給學習材料，

之後有問題隨時解答。針對讀者的，就用 *library instruction* 的方式。在圖書館的網站上也提供相關資訊及韋氏與漢語拼音的轉換表，供他們參考。CEAL 裡，我不記得有正式的、特別為拼音轉換做的培訓。但一般的信息都會在 *eastlib* 上發表。CEAL 開年會時，也都會安排一段時間討論這件事情。

Q: 當時 UCLA 完成拼音轉換的情況是怎麼樣的呢？有困難或挑戰嗎？

A: UCLA 的相關部門討論之後，決定把有中文的書目記錄下載，送去 OCLC 做轉換。我們盡量把一些已知的問題，在下載記錄之前就先改正，原來的記錄也做了備份。OCLC 把轉換好的記錄及轉換報告送回來後，我們先初步檢查並清理部分問題，系統部門再把舊的記錄取代。遭遇的困難就是人手比較少，又在記錄的清理上面花了很多時間，所以也不可避免的影響到其他的工作。因為我 2002 年的秋天就離開 UCLA 了，可能有一些事情我來不及做完，留給別人去處理，這是我比較遺憾的事情。我去了耶魯大學以後，就專注在耶魯東亞圖書館的工作了。

Q: 您談到在耶魯做紀錄清理的工作，您還記得有其他的和拼音轉換有關的工作是您在耶魯或哥倫比亞做的嗎？

A: 在耶魯大學時，我們也與總館的同事合作，做了一些 *batch processing* 的清理工作。項目結束以後還有一些斷斷續續的 *cleanup*。後來 LC 在拼音羅馬化指導方案上陸續做了一些改變，尤其在法人機構和地名方面，有些更新要做。我們就要跟總館的同事配合，有些就在 *authority records* 上直接改，但有些就需要手工做。還有一個問題跟拼音轉換有點關係。在耶魯、哥倫比亞及許多大學都有這樣的情況，就是特藏部門舊的 *finding aids* 上的羅馬拼音大都是用韋氏拼音。那些是否要全改，必須慎重考慮。因為早期很多華人的正式英文名字都是用韋氏拼音，或者其他不是很規範的羅馬拼音。如果硬把它們改成漢語拼音，說不定會找不到資料，或者會有張冠李戴的問題。可能在將來要用 *cross-references* 或其他方法來解決。哥大的王成志做了一些工作，提供拼音的輔助。可是讀者如果去檔案館看資料的話，*finding aids* 上還是韋氏或其他拼音。成志就偶爾要幫助讀者去使用這些檔案。因為管理特藏的人大都不懂中文，他們可能沒有意識到，或者說不覺得問題有那麼嚴重。但是長遠來說，對讀者還是不方便。這是一個有待處理的問題。

Q: 拼音轉換項目 20 年後，您現在怎麼評價當時的工作，過程和結果？如果可以穿梭回過去，您覺得會有不同的做法嗎？

A: 現在科技這麼發達，有很多事情都可以比較容易做，可是回到二十年前的話，選擇就不一樣。我覺得參與計劃的這些前輩都盡了最大的努力去籌備、安排這些事情。他們也徵求很多人的意見及想法。所以我認為他們在那時的環境下已經做得很好了。

Wade-Giles to Pinyin Conversion: Australian Experience and Local Issues

Prepared by *OCLC CJK Users Group Pinyin Conversion Task Force*
June 21, 1999

This report attempts to address technical issues relating to the Wade-Giles to Pinyin conversion for the purpose of helping OCLC CJK member libraries understand the conversion process and assisting librarians prepare for the conversion. This report includes two parts: I) Summary of the Australian conversion project undertaken in 1996; II) A list of suggestions related to the local conversion issues, for which the task force recommends libraries start dialoguing with related local library staff as early as possible.

Part I. Summary of the Australian Conversion Project

Conversion Process:

In 1996, the National Library of Australia (NLA) developed a conversion program that automatically converted 500,000 CJK (Chinese, Japanese, and Korean) records from Wade-Giles romanization to Pinyin.

The conversion program NLA developed was written to perform major functions in sequences as the following:

Q: 您覺得這個拼音轉換項目它的意義在哪裏？

A: 我覺得意義是在於它讓圖書館界和外面的世界接軌。對讀者或學者而言，他們不需要再去學多種拼音的方案，也能比較快的找到需要的資料。其實如果是利用現在的科技的話，轉換與否或許已不是那麼重要，因為電腦可以在幕後做很多事情。只要有精密的 **mapping tables**，漢語拼音和韋氏拼音可以自動轉換、鏈接多音字或簡體字和繁體字的鏈接也可以用程式自動處理。但在當時還沒有這樣的技術，所以這是一個非常重要的項目。

王晓彤 Xiaotong Wang



Then: Chinese Cataloger, University of Oregon Libraries

Now: East Asian Cataloging Team Leader/Chinese Studies Specialist, University of Oregon Libraries

Q: 拼音转换之前，大学用户使用韦氏音标检索中文资源是什么情况？

A: 不管采用哪种汉字罗马化方案，目的都是为了把汉字发音系统地转化成罗马音标。我个人对于官方选择哪个方案并没有任何偏好，但从用户的角度看，图书馆使用韦氏拼音对正在学习中文的用户的确造成困扰。研究中国的学者一般已经具备相当的语言功底，困扰可能不大。但对初学中文的人，他们上课学的是汉语拼音，若希望进一步了解中国、甚至迈入中国研究的门槛，那么一到图书馆会被要求使用韦氏拼音，而后者是截然不同的罗马化方式—这种转换对他们几乎肯定不会顺畅

Q: 您如何参与拼音计划？主要职责是什么？

A: 我当时在俄勒冈大学图书馆东亚编目组任中文编目员，记得北美图书馆决定转用拼音后，总馆和部门领导都挺支持的，我们提出的工作要求，图书馆都尽力支持并配合，这也是转换得以顺利进行的一个原因。项目开始后，主要是我的主管 **Daphne Wang** 负责与总馆及其他部门沟通协调，我和同事们则专注于本地转换的具体事务。我们的转换步骤和安排大致根据已经划定的时间线进行，我个人基本就是配合本校的计划处理拼音转换的具体问题。

Q: 大学本地转换工作的协作何如？

A: 项目一开始，参与方不仅有我们东亚编目组—其中以 **Daphne** 和我两个人为主，还有图书馆其他部门，尤其负责系统管理的技术人员，他们对甄别和系统更新条目数据贡献良多。**Daphne** 一直负责与校内外各方的联络和协调，在整个项目实施时起了关键作用。

Q: 当时用拼音编目的焦点议题有哪些？

A: 我和 **Daphne** 在实际转换中发现了一些问题，还写了一篇文章探讨拼音转换的挑战。转换并非像很多人想象那么简单、一蹴而就，很多问题都会牵一发而动全身。我记得中文资料的目录还相对容易辨识和处理基本可以通过系统自动转换，但非中文资料就复杂多了，并不是当时的技术可以轻易处理的。我们得把这些记录先甄别、提取出来，然后经常要靠人工一条一条地改。而且，包含韦氏拼音的非中文资源条目往往并不提供对应的中文信息，所以很难能被直接转成拼音，然而查对非常耗时费力，几乎是不可能的事—这是我印象深刻的一个难题。

另一个经常被提到的问题就是分词 (word division)。拼音转换在分词问题上和中国汉语拼音方案的做法完全不同，我觉得这多少也会影响学者用拼音进行检索的体验，并可能造成一些困扰。我对北美图书馆当初做出这个决定的原因和背景不甚了解。但是对比韩文资源的罗马化标准，北美图书馆就直接套用了和韩国罗马化方案基本一致的分词法，所以我猜想面对中文资源时，北美图书馆界可能有不同的考虑。

Q: 对北美拼音转换项目有何评价?

A: 由于一些历史或政治原因，北美图书馆接受汉语拼音显然经历了相当多的波折。当时有一种倾向把拼音和一些政治问题进行挂钩。如果大家一开始就能够“去政治化”地看待这个问题，事情也许会单纯并顺利很多。到2000年和2001年执行拼音转化，我觉得基本是顺应潮流迈出的积极的一步。另一方面，我认为处理非拉丁语类的资料—包括它们的规范权威档 (authority records)，应该搭配和罗马注音平行的原语 (vernaculars) 信息，才能比较准确地传达非拉丁语资料的信息。返观北美图书馆处理包括中、日、韩在内的非拉丁语原语资料时，过去因为政治或技术的原因，并没有采集重要的原文信息，而只呈现原语经过罗马化后的注音表达，我觉得这是远远不够的，对用户查找、利用资源造成不小的问题。所以拼音转换肯定是积极的一步，但也应该看到它只是朝正确方向前进的一步而已。

Q: 对目前的汉字罗马化文件有什么意见和建议?

A: 把作为象形文字的汉字进行罗马拼音转化，这项工作本身就不会有一个完美的解决方案，即使大陆的汉语拼音方案也不完美，但是相比其他罗马化方案，中国汉语拼音方案还是基本清楚的，并能处理大部分问题。因为北美用户的特点和需要，汉字罗马化势在必行；我个人的意见是希望图书馆沿用大陆的汉语拼音方案，这样编目员就无需自己再去作决定，同时也降低用户有效检索的门槛。现在我们使用的方案在处理专有名词时还有很多困难，希望情况以后能有所改善。

王维 Wei Wang



Then: CJK cataloger, University of Virginia Library

Now: Librarian for East Asian Studies, University of Virginia Library

At that time I was a cataloger. CEAL did a great job of the publicity campaign for the project, especially to the upper management of all the libraries. I was pleasantly surprised that our university librarian knew about the project. It might sound counterintuitive that my Wade-Giles proficiency was greatly improved through the Pinyin conversion project. One of the lessons learned was the importance of diacritic marks. Wade-Giles uses diacritic marks to differentiate. However, there were quite some bib records that did not have proper diacritic. Although the search results were not affected by diacritic, with or without, the auto-conversion results were. For example, Kao = Gao, but K'ao = Kao. Therefore, the conversion results of the Wade-Giles records without proper diacritic had to be manually corrected. It was nice to receive the certificate of appreciation from OCLC for my contribution to the cleanup project.

余梁戴光 Anna U



Then: Director, Cheng Yu Tung East Asian Library, University of Toronto

Now: Retired

Q: Concerning cataloging and processing Chinese language materials, what was the practice in your library? How about your collaboration with RLG or OCLC, as a Canadian library?

A: The catalogued records at UTL, including the East Asian studies records, were in machine readable format at that time when I took over the management of the EAL. So RLG negotiated with us to use our records to provide the base for the RLIN union catalogue database in the building and expanding of East Asian records for the CEAL community. The Chinese records were romanized in Wade-Giles at that time. Later and eventually OCLC took over RLG.

Q: How did your library handle the Pinyin conversion project, and your timeline?

A: I think the first real strong interest among libraries in CEAL to change over to Pinyin was after OCLC began importing Pinyin records from the National Library of China. The CEAL community was obviously interested in expanding access to a lot more Chinese materials not held in North America.

In 2000-2001, EAL/UT was one of the selective nine-member RLG North American pilot group for the Pinyin Conversion project. Within that, EAL/UT was among the first five members chosen for the priority conversion of our Chinese records in RLIN to Pinyin romanization. There were roughly 130,000 of our records converted by RLG into Pinyin. And about 16,000 records were not fully converted and had to be manually done. EAL spent a lot of time reviewing and editing those records.

As a reward for our priority participation, we were given a 40% discount when we purchased the RLIN conversion tape. That was in November, 2002. At the same time, around October, 2000, all the new Chinese records for new acquisitions were to be catalogued in Pinyin in both RLIN and UT Cat. (Our records were in both the UT Library database as well as in RLIN). While all that was ongoing, library users were asked to use both Pinyin and Wade-Giles when they searched in the UT catalogue and to use Pinyin and Chinese scripts when searching in RLIN union catalogue. In mid-January 2005, ITS of the main Library ran the tape load for Pinyin conversion in SIRSI (the UT catalogue; the public catalogue) by overlaying the existing Wade-Giles records in the system. Some 7,000 records in the system were not picked up because of various problems such as typos, serials that had changed titles and microform titles, etc. Another 30,000 converted records lacked parallel script entries. Those problematic ones had to be converted manually. But overall, about 99% of the Chinese records in the main catalogue were successfully overlaid. The problematic records had to be converted by staff manually.

Q: Ensuing the conversion, did your library provide any training sessions for your community?

A: In helping the users, the EA library offered instruction sessions to faculty and students to teach catalogue searching skills in Pinyin; and also introduced two union catalogues, the RLG union catalogue and the OCLC WorldCat. The instructional sessions were offered weekly in a hands-on drop-in format or by appointments. The feedback was very positive.