The 5th WCILCOS International Conference of Institutes and Libraries for Chinese Overseas Studies
Chinese through the Americas
May 16-19, 2012 Vancouver, B.C. Canada

Bilingual Metadata and Discovery
- Don't Let Your Chinese Archives Get Lost In Translation

Lilly Li
CHRP Librarian/Archivist of University of British Columbia Library, Canada

Allan Cho
Program Services Librarian of University of British Columbia Library, Canada

In recent years, with the rise of China, the study of overseas Chinese has developed rapidly, become a field that attracts mainstream historians and scholars in both overseas and China. To meet their information needs, libraries and archives of higher education institutes in North America are dedicated in collecting and digitizing the overseas Chinese archives, from historical photographs to genealogical materials to census reports, on and by overseas Chinese.

However, how to develop user-oriented metadata and make the digital resources accessible to both western and Chinese scholars is a big challenge to the libraries of North America. English has been a dominant language for cataloguing, classification and searching system in the libraries of North America. In the past a couple of decades, the libraries with Chinese language skills have made a lot of efforts to catalogue and index the Chinese-language archival materials. However, the quality of the work is largely dependent on individual libraries' capabilities. Translation mistakes and irrelevant subject headings make many valuable archival materials irretrievable.

Nevertheless, inconsistent Romanization system makes the situation worse. Chinese from different parts of the world do not share a unified Romanization system. Whereas Hanyu Pinyin is now universal for China and also a widely used method of Romanization outside of China, Wade-Giles and other forms of transliteration are still far more popular in the communities of early Chinese migrants where the most Chinese archival materials are collected from. Many scholars hope, by standardizing Romanization system, to improve the accessibility of the Chinese-language archival materials. But the flaw in that approach is, not like English, Chinese language has a strong distinction between Chinese written language (文, Pinyin: *wén*) and spoken language (语/語 *yǔ*). It will be very hard, if not impossible, to standardize the many Chinese spoken languages. For more than 2,000 years, Chinese written language has linked people who have locally different spoken languages. Therefore, Chinese written language appears a hopeful solution to the messy problem.

In this paper, we discuss the importance of Chinese metadata in cataloguing and retrieving Chinese-language archival materials, and analyze the challenges and opportunities of creating bilingual metadata for the overseas Chinese archives. We also present our efforts to create the bilingual database for the "Chinese Canadian Stories" portal to improve the discovery of the Chinese Canadian historical materials.