MARC Formats in China: Local or International?

Ben Gu

1 Introduction

The use of MARC tapes was the first step for Chinese librarians to use MARC records. It can be traced back to 1980, when National Library of China (NLC) imported LC MARC tapes for test retrieval of books in Western languages. In July 1981, the Party School of the Central Committee of the Communist Party of China (CPC) imported LC MARC tapes containing 413,000 records during April 1979-April 1981. The tapes were used to retrieve books on international communist movement, and there were about 1000 hit records. Because Chinese libraries never used such new technology before, it should even be approved by the CPC Secretariat, the top level management at that time (An, "Yin jin MARC ci dai de shou ci ying yong"). The test work was done by MARC Cooperation Group jointly participated by the National Library of China, Peking University Library, Tsinghua University Library, Renmin University Library and the China National Publications Import & Export Corporation ("MARC ci dai yu wo guo xi wen bian mu"; Chen, "Beijing di qu xi wen tu shu ji du mu lu yan zhi jin zhan"). Actually, Chinese librarians began to pay attention to LC MARC as early as 1975, but they could not do anything but research, because the lack

of international exchange during the "Cultural Revolution" period (1966-1976) (Liu).

In 1985 Chinese libraries began to create bibliographical records of monographs in Chinese languages, and had accumulated about 1 million records by the end of 1990.

2 Decision of CNMARC Based on UNIMARC

Since the first publication of UNIMARC in 1977 and the second edition in 1980, Chinese library cataloging community, which was just considering catching up with international developments, decided to develop a Chinese local format based on UNIMARC, because it is an IFLA-developed international standard and even Library of Congress announced to implement UNIMARC (This what Chinese librarians knew about LC's decision; actually, LC's implementation was to provide conversion between USMARC and UNIMARC)(IFLA). There were other possible reasons for Chinese libraries to use UNIMARC as the basis of CNMARC:

- Main Entry Issue: Chinese libraries do not use main entry in the cataloging of Chinese materials, therefore USMARC, which was suitable for the cataloging with the concept of main entry, seemed strange to Chinese catalogers.

- MARC Expert Suggestion: It is said that Henriette Avram, the first developer of MARC format, visited China in late 1980s and expressed the feasibility for Chinese libraries to use UNIMARC.

- Japanese Example: Japanese libraries were already using MARC format earlier and basing its JapanMARC on UNIMARC.

- USMARC Shortcomings: There are some shortcomings with USMARC, and even other countries using USMARC made modifications and had their own versions.

Because there is a lack of historical records, the author can just provide the above reasons from the interviews with Mr. Yan ZHU, Ms. Beixin SUN and Mr. Yuanzheng CHEN, who were members participating in drafting of the first edition of CNMARC. People can hardly imagine the conditions of the 1970s and 1980s, when international exchange of information was extremely insufficient.

3 CNMARC and CMARC

The process of compiling the Chinese national format was the following: the National Library of China began to compile *Chinese MARC Communication Format* (later called CNMARC) in 1986, it completed a draft by the end of 1986 (Sun), sent copies to other libraries for national review in January 1987, and revised the draft according to the latest edition of UNIMARC (1987) in early 1988; the format was second-time reviewed during a special workshop hosted by the Library Society of China in 1989 ("Zhongguo ji du mu lu ge shi xue shu yan tao hui ji yao"), and was finally published by the Bibliography and Documentation Publishing House in February 1991. At the same time, the National Library began to distribute CNMARC records to libraries in China and around the world ("Beijing tu shu guan xiang guo nei wai fa xing ji du mu lu"). More and more libraries have adopted CNMARC for cataloging of Chinese materials (Chen, "Zhong wen tu shu ji du mu lu zhu ti biao yin cai yong hou kong gui fan de she xiang").

In Taiwan Province, a local MARC format (later called CMARC) for monographs, which was strictly based on UNIMARC, was published early in 1981 and was revised almost at the same year, while

a MARC format for all types of publications was published in July 1984. This format was earlier than the CNMARC in the Chinese mainland, and was used as an example for drafting of the CNMARC (Chen, "Da lu he Taiwan de Zhong wen MARC bi jiao"). Although both are based on UNIMARC, CNMARC and CMARC are slightly different, especially in the use of character sets and terminologies (Zhu and Mengjie).

In January 2002, *China MARC Format / Authorities* was released as an industrial standard of the Ministry of Culture (WH/T15-2002). It was based on UNIMARC / Authorities published by IFLA in 1991. Since then, there has been no revision of the format (*China MARC Format / Authorities*).

4   UNIMARC Translations

UNIMARC Bibliographic Format was translated into Chinese in 1986. However, it was just used for internal use for the drafting of CNMARC, and was not published and distributed through commercial channel (Chen, "Zhongguo ji du mu lu tong xun ge shi cun zai de wen ti yu xiu gai jian yi"). There have been translations of some later editions, but none of them was published.

Translation is very important for correct understanding of the original UNIMARC text. In *Manual of the New Edition of CNMARC* in March 2004 (Guo jia tu shu guan, *Manual of the New Edition of CNMARC*), for example, EAN (073 - International Article Number) was misunderstood as a kind of numbers assigned to articles in published journals.

By the end of 2013, the Chinese translation of *UNIMARC Manual: Authorities Format* (3rd Edition, 2009) will be published by the National Library of China Publishing House. The translation was done by the Working Group for the Drafting of *CNMARC Authorities*

Format, a national standard proposal supervised by the National Library Standardization Technical Committee in 2009 and approved by the Standardization Administration of the People's Republic of China.

5 CNMARC Revisions

After the publication of CNMARC in 1991, there were some critics on it, pointing that:

- There are too many fields and subfields.

- There are some duplicated fields or subfields.

- The definitions of some data elements are not clear.

In March 1993, the Ministry of Culture approved the proposal of CNMARC Format as a library industrial standard (Qi). The standard (WH/T0503-96) was published in 1996, and began to be implemented in July 1, 1997 (Zhu). This standard was based on the 1994 edition of *UNIMARC Manual*. At the same time, main contents of the standard were included in the published *China MARC Format Manual* (1995) ("Zhongguo ji du mu lu ge shi ji shi yong shou ce nian nei chu ban"), and the revised edition of the manual was published in 2001 (Pan).

The industrial standard (WH/T0503-96) was revised according to the latest edition (2002) of UNIMARC, and was intended to become a national standard. The draft for the national standard was completed and approved by an expert committee organized by the Ministry of Culture in 2003, and its major contents were published as *Manual of the New Edition of CNMARC* in March 2004. Because of some procedural problems, the national standard was not finally released.

In December 2009, the Standardization Administration of the People's Republic of China approved the proposal of two national standards, *CNMARC Bibliographic Format* and *CNMARC Authorities Format*.[1] The two national standards were planned to be released in 2011, but they are delayed due to personnel and technical problems.

6 Special Characteristics in CNMARC

In addition to some minor changes to UNIMARC format, the major differences of CNMARC are reflected in the definition of some local fields 9–, -9-, –9, and subfields, $9 (and $A, $B, etc.).

For example (Bibliographic format):

- 091 Union Books and Serials Number: numbers assigned by Chinese Administrative Agencies for books and serials, especially before the implementation of ISBN and ISSN in China.

- 092 Order Number: Numbers assigned by distributers.

- 094 Standard Publication Number: For the numbers of international, national, industrial or enterprise standards.

- 191 Coded Data Field: Rubbings

- 192 Coded Data Field: Ethnic Music of China

- 193 Coded Data Field: Chinese Antiquarian – General

- 194 Coded Data Field: Chinese Antiquarian – Copy Specific Attributes

- 393 Outsystem Chinese Character Note: For the description of Chinese characters not defined in character set.

---

[1] Standardization Administration of the People's Republic of China, December 14, 2009: http://www.sac.gov.cn/gjbzjh/201012/t20101213_56788.htm.

- 690 Chinese Library Classification

- 692 Classification for Library of Chinese Academy of Sciences

- 696 Other Local Class Numbers

- 905 Holding Information: Used in the second edition of the *China MARC Format Manual* (2001), when Chinese libraries did not use separate MARC format for holding records.

As to Pinyin Romanization, NLC and public libraries use subfield $9 to include all Latin scripts corresponding to Chinese characters in major subfields of the field, while CALIS (China Academic Library & Information System) and university libraries use $A, $B etc. for Latin scripts corresponding to Chinese characters in $a, $b etc.

Besides, CNMARC doesn't include some UNMARC fields not applicable to Chinese library cataloging, e.g. 012 (Fingerprint Identifier) and 670 (PRECIS).

As libraries in China do not use the concept of main entry for the cataloging of monographs in Chinese language, fields 700, 710, 720 are usually not used. Besides, name/title access points are not used in NLC records. Therefore, there are difficulties in adding uniform titles in bibliographic records, and there are very few records with uniform titles.

In NLC records, 701$b and 701$g are not used, while in CALIS records, 701$b and 701$g are used for foreign names (Xie).

Punctuations are not used in CNMARC records, and they are created automatically by computer systems in OPAC display. However, there are some problems with the punctuations. For example, 200$d is for Parallel Title Proper and 200$e is for Other Title Information, and we are hard to define automatic generation of punctuations for parallel other title information, because there are many possibilities for 200$e. However, CALIS solved this problem by adding punctuations in these particular subfields.

7 Co-Existence of Two MARC Formats

In China, libraries use two MARC formats, MARC21 and CNMARC. CNMARC has some special fields for Chinese publications in addition to UNIMARC fields. Most small-sized libraries use UNIMARC-based CNMARC for all materials, with the consideration of easy management. For large-sized libraries that have sizeable collections in foreign languages, they prefer USMARC/MARC21 to CNMARC with the consideration of international compatibility and easy copy cataloging.

In the National Library of China, we use CNMARC, Chinese Library Classification and Classified Chinese Thesaurus for Chinese publications, and we use MARC21 for foreign publications, including those in Western languages, Japanese and Russian. For publications in Western languages, we use AACR2 (later RDA), LC Subject Headings, LC Name Authority File and Chinese Library Classification. The Aleph500 system allows us to maintain two separate databases respectively in CNMARC and MARC21 formats. There is no relationship between the two databases now. We were considering the possibility of establishing relationships between the two authority databases respectively in CNMARC and MARC21, but the cost was estimated very high.

In the Chinese mainland, most libraries use UNIMARC-based CNMARC for Chinese publications. In Hong Kong Special Administrative Region, most libraries follow the practices of the Western countries and use MARC21 for all publications. In Macau Special Administrative Region, some libraries use CNMARC and some use MARC21. Large libraries in Taiwan Province use CMARC (also based on UNIMARC and similar to CNMARC) or MARC21; mid-sized libraries favor CMARC. For libraries using UNIMARC-based formats, they also have different rules, especially for name headings and authority records.

As to Romanization, most libraries use Pinyin now, but they sometimes use different ways of segmentation, for example "mao ze dong" or "Mao Zedong" (Gu).

8 Holdings Format

At the early stage of library automation in China, there was no local library automation system using the technology of relational database, and there was no need to use a separate holdings format. Libraries used tag 905 in the bibliographic format to record holdings data.

Since the beginning of the 21th century, more and more libraries use international systems, such as Aleph 500, and people have to consider the use of holdings format. Because there was no UNI-MARC holdings format, a research group of the National Library of China considered the localization of MARC21 Format for Holdings Data. In 2003, a Chinese translation of the MARC21 Format for Holdings Data was published (Guo jia tu shu guan, *MARC21 Format for Holdings Data*), and the format began to be applied in NLC and other libraries.

9 Problems in International Exchange

CNMARC is used mainly for the distribution of MARC records in China. Sometimes, we have information exchange with East Asian libraries, and CNMARC records are easy to be converted into JapanMARC (Japan) and CMARC (Taiwan Province) records, which are based on UNIMARC. However, we have difficulties in sharing bibliographic information with libraries in Hong Kong, a special administrative region of China, which use MARC21, and

those in Republic of Korea, which use KorMARC based on MARC21. Further, we have more difficulties in sharing CNMARC records with foreign libraries, most of which use MARC21 format.

After almost ten-year discussions and negotiations, NLC and OCLC signed an agreement for batch uploading of NLC Chinese bibliographic records to OCLC WorldCat database in 2008,[2] and had uploaded 2.3 million records by the end of 2008.[3] This is a great step forward for Chinese MARC records to be used by foreign libraries. However, because exact CNMARC-MARC21 automatic conversion is almost impossible, and manual conversion requires huge costs of human resources, NLC and OCLC had to choose the option of batch conversion, leaving details to be upgraded by OCLC users. The batch uploading process also adds NLC holding information to WorldCat.

10   Future Perspectives

Although there have been talks about the death of MARC, CNMARC is still a major format for the cataloging of Chinese materials and exchange of bibliographic records in China. If there is a need for records with other formats, such as XML, we just convert CNMARC into them. There is not any cataloging agency distributing XML records yet.

Since the bibliographic year 2009, the German National Library has been delivering its data in the format MARC 21.[4] Some libraries

---

[2] National Library of China to Add Its Records to OCLC WorldCat: http://newsbreaks.infotoday.com/Digest/National-Library-of-China-to-Add-Its-Records-to-OCLC-WorldCat-41153.asp.

[3] http://www.nlc.gov.cn/dsb_zx/gtxw/201201/t20120116_58355.htm.

[4] MARC21, Deutschen Nationalbibliothek: http://www.dnb.de/EN/Standardisierung/Formate/MARC21/marc21_node.html.

in Taiwan Province also switched its bibliographic format from UNMARC-based CMARC to MARC21 in recent years.[5]

Chinese librarians have to consider the following questions:

- What is the original purpose for us to choose UNIMARC for the basis of CNMARC twenty years ago?

- Shall we still use UNIMARC for the basis of our cataloging?

- Shall we even abandon MARC formats and use a "modern" format?

The author has answered the first question in this article, but he cannot answer the other two questions now. The author doesn't think there will be any answers in the next few years. CNMARC will still be our major format for the cataloging of printed resources. Some new fields and subfields related to FRBR in the updates 2012 will be good tools for the application of FRBR in Chinese libraries.

---

[5]http://catwizard.net/posts/20120125095904.html.

References

An, Shulan. "MARC ci dai yu wo guo xi wen bian mu". *Computers and Libraries* 4. (1983): 7–10. (Cit. on p. 265). Print.

——. "Yin jin MARC ci dai de shou ci ying yong". *Computers and Libraries* 4. (1981): 58. (Cit. on p. 265). Print.

"Beijing tu shu guan xiang guo nei wai fa xing ji du mu lu". *Science and Technology Daily* February 21. (1991). (Cit. on p. 267). Print.

Chen, Fuliang. "Da lu he Taiwan de Zhong wen MARC bi jiao". *Library Work and Research* 1. (1994): 28–30. (Cit. on p. 268). Print.

Chen, Yuanzheng. "Beijing di qu xi wen tu shu ji du mu lu yan zhi jin zhan". *Library and Information Service* 5. (1983): 1–6, 41. (Cit. on p. 265). Print.

——. "Zhong wen tu shu ji du mu lu zhu ti biao yin cai yong hou kong gui fan de she xiang". *Journal of Library Science in China* 4. (1991): 69–71. (Cit. on p. 267). Print.

——. "Zhongguo ji du mu lu tong xun ge shi cun zai de wen ti yu xiu gai jian yi". *Library Development* 5. (1992): 53–57. (Cit. on p. 268). Print.

China MARC Format / Authorities. Beijing: Zhonghua ren min gong he guo wen hua bu fa bu, 2002. (Cit. on p. 268). Print.

Gu, Ben. "National Bibliographies: the Chinese experience". *Alexandria* 18.3. (2006): 173–178. (Cit. on p. 273). Print.

Guo jia tu shu guan. *Manual of the New Edition of CNMARC*. Pechino: Beijing tu shu guan chu ban she, 2004. (Cit. on p. 268). Print.

——. *MARC21 Format for Holdings Data*. Pechino: Ke xue ji shu wen xian chu ban she, 2003. (Cit. on p. 273). Print.

IFLA. "Library of Congress Implements UNIMARC". *IFLA Journal* 11.3. (1985): 273–274. (Cit. on p. 266). Print.

Liu, Rong. "Tan LC-MARC zai wo guo de shi yong". *Library and Information Service* 4. (1982): 22–23. (Cit. on p. 266). Print.

Pan, Taiming. *China MARC Format Manual*. Pechino: Ke xue ji shu wen xian chu ban she, 2001. (Cit. on p. 269). Print.

Qi, Siyan. "CN-MARC de wen shi ji she hui xiao ying". *Library Theory and Practice* 4. (1997): 29–30. (Cit. on p. 269). Print.

Sun, Beixin. "The role of the National Library of China in standardization activities". *Scholarly Information and Standardization: Proceedings of the Twelfth Open Forum on the Study of the International Exchange of Japanese Information and Scholarly Databases in East Asian Scripts 1992/1993, November 20*. Tokio, 1992. (Cit. on p. 267). Print.

Xie, Qinfang zhu bian. *Manual for CALIS Online Cooperative Cataloging, Part I*. Pechino: Beijing da xue chu ban she, 2000. (Cit. on p. 271). Print.

"Zhongguo ji du mu lu ge shi ji shi yong shou ce nian nei chu ban". *New Technology of Library and Information Service* 6. (1995): 64. (Cit. on p. 269). Print.

"Zhongguo ji du mu lu ge shi xue shu yan tao hui ji yao". *New Technology of Library and Information Service* Suppl. (1989): 10–55. (Cit. on p. 267). Print.

Zhu, Qingqing and Huang Mengjie. *A Comparative Study of Chinese Cataloging Practices between Libraries across the Taiwan Straits*. Pechino: Shi jie tu shu chu ban gong si Beijing gong si, 2011. 15–47. (Cit. on p. 268). Print.

Zhu, Yan deng qi cao. *CNMARC Format*. WH/T0503-96 1996.2.6 fa bu 1997.7.1 shi shi. 1996. (Cit. on p. 269). Print.

BEN GU, Director of Foreign Acquisitions & Cataloging Department, National Library of China, Beijing, China.
bgu@nlc.gov.cn

ABSTRACT: The application of MARC formats in Chinese libraries has a history of about 20 years. At the age of internationalization and digitization, people ask the question: why we chose UNIMARC for the basis of CNMARC? Is it the best choice for China? Are there any other options? In this paper, the author reviews many historical documents and analyzes the present status of MARC formats in China.

KEYWORDS: Bibliographic formats; China; CNMARC; MARC; MARC21; UNIMARC.