2-1-1993

Experiences and Problems in Retrospective Conversion of East Asian Language Materials

Ai-Hwa Wu

Follow this and additional works at: https://scholarsarchive.byu.edu/jeal

BYU ScholarsArchive Citation
Available at: https://scholarsarchive.byu.edu/jeal/vol1993/iss98/3

This Article is brought to you for free and open access by the All Journals at BYU ScholarsArchive. It has been accepted for inclusion in Journal of East Asian Libraries by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
ARTICLES

EXPERIENCES AND PROBLEMS IN RETROSPECTIVE CONVERSION
OF EAST ASIAN LANGUAGE MATERIALS*

Ai-Hwa Wu Arizona State University

I. Introduction

In this computer age, either by choice or necessity, more and more libraries in North America have been engaged in having the bibliographic records of their retrospective collections converted from catalog card into machine-readable format. Unlike the automated cataloging of current materials in the libraries, the conversion of these records does not always begin with those for Roman language materials and end with those for East Asian language materials. For instance, libraries such as the University of Iowa Library are converting both their retrospective Roman and East Asian language records concurrently; yet, institutions, like Arizona State University (ASU), have started converting their East Asian language records only after they converted the ones in other languages.

Retrospective conversion (or recon, for short), initially with the purpose of facilitating the library's automated circulation system, has now become an integral part of the library's plan toward having a local online system that not only encompasses online public access catalog, acquisitions, circulation, and catalog maintenance, but also enables the users to access distant data bases and collections as far away as the other side of the world. Due to the ever increasing budget constraints and inflation, access to, rather than ownership of, material is becoming the only alternative for many libraries. In order to know about the holdings of other libraries and for other libraries to know about ours, it is imperative that all retrospective library collections be converted and be electronically as searchable as the current collections.

II. Conversion Models

Similar to that for other language materials, the retrospective conversion of the East Asian language records takes on two main patterns: contract-out and in-house. Each pattern has its set of problems and provides some unique challenges. The advantages and disadvantages

---

*This paper was originally presented in abridged form at the Western Conference of the Association for Asian Studies, Tucson, Arizona, October 23, 1992.
for these patterns are described and analyzed below. The advantages of one pattern are most often the disadvantages of the other.

Contract-Out Conversion

1. Scarcity of Vendors

By the very nature of the East Asian language materials which requires the data in vernacular characters to be a necessary part of the bibliographic record, there are fewer vendors who are equipped to be in the market for conversion of these records. In 1989, there were only two companies in North America that provided recon service for East Asian language materials. They were the Asian Shared Information & Access in Alhambra, California and the OCLC Online Computer Library Center, Inc. (OCLC) in Dublin, Ohio. Today, four years later, there are two additional organizations who have come on the scene. These are Retro Link Associates, Inc. in Provo, Utah and Fastcat-Asia of the Pacific Automation Systems Services in Beverly Hills, California.

2. Higher Cost

Also by the very nature of the materials and as agreed to by the Research Libraries Group East Asian studies librarians, each full-level, standard East Asian recon record needs to have its bibliographic information in both romanized and vernacular forms in at least three areas of the record: title and responsibility, publication and distribution, and series statement.

To hire people with both suitable cataloging expertise and language background to do the job, the vendors will have to charge a fee higher than that for a regular non-East Asian language project. For example, around 1987, OCLC provided Arizona State University with an estimate of $6.00 to $6.50 per record to convert ASU's East Asian collection while it charged $1.85 for each non-East Asian language record. In 1989, OCLC charged the University of California at Los Angeles (UCLA) $6.00 per East Asian record conversion while it charged $2.20 per non-East Asian language book record. In 1991, Asian Shared Information & Access quoted $17.50 per record to be upgraded with East Asian vernacular characters and $33.00 per record to be input into the Research Libraries Information Network (RLIN) data base.

3. Specifications

The specifications for recon have to be written in a very precise, specific, and step-by-step manner, listing all the points concerned and providing ample examples that will illustrate succinctly the romanized as well as the vernacular fields. For instance, what constitutes a match from your library's point of view and what are the criteria for choice of a record? What sort of data needs to be edited and what is to be left unchanged? Which areas of the record need to have the data in both romanization and vernacular? In the case of a record that needs to be input, how much authority checking needs to be done by the vendor?
Last year, 500 East Asian language cards were shipped by ASU to a potential vendor. The cards were sent to the vendor for a trial run of the conversion in order to provide ASU with a more accurate price quote for doing the whole project by the company. The company converted roughly 50 percent of the cards. Problems, ranging from wrong MARC contents designation, mix-up in application of AACR1 and AACR2, to supplying English terms in lieu of transcribing the East Asian data as given on the cards, were detected. For example, ASU’s local Optical Character Recognition (OCR) number and item level data were lumped and keyed into a 590 field instead of a Local Data Element A (LDA) field (i.e., an RLIN local data element field equivalent to an OCLC 94x field) as specified in our specifications; LC’s location appeared in the ASU call number area; headings were coded and placed in wrong fields with incorrect subfield designations; duplicate records were not caught; and wrong records in the RLIN data base were used. Granted that these problems occurred due to the lack of both cataloging expertise and RLIN-specific US-MARC knowledge on the part of the vendor, our specifications would have been more helpful if they had been more illustrative.

4. Communication

Continuous communication with and monitoring of the vendor are extremely important for the success of the project. Assurance of having the specifications interpreted as they are intended is essential. Things not clearly stated in the specifications are subject to a free interpretation by the vendor.

In-House Conversion

1. Recruitment, Training, and Supervision of Project Personnel and Equipment and Furnishings

There are basically three versions of staffing pattern for an in-house project: (a) hiring of a small group of new personnel having different levels of expertise to work solely on recon, (b) add-on of clerical assistants to support the existing staff members who have been designated to do recon part-time, and (c) a complete reassignment of the existing staff to do only recon. For versions (a) and (b), it takes time to do recruitment, training, and supervision. But time spent in hiring and training is short-termed. Time taken for supervision compensates for the time taken for long-distance communication with the vendor and for monitoring of the vendor’s products. However, additional equipment and furniture for the project and its staff could be a heavy investment.

2. Lower Cost

On the average, it costs ASU $3.04 to convert an East Asian language record using a combination of RLIN’s Batch Retrospective Conversion (BRCON) software program and direct online input methods. Compared to the $6.00 average per record done by OCLC for
UCLA in 1989, it is much less expensive for ASU to do the project in-house. However, the support and commitment of the library administration is essential since an in-house conversion will inevitably disrupt the regular activities and productivity of the existing staff to a certain degree, depending on the level of their involvement in the project.

The average cost of $3.04 per record for ASU's project was calculated in this way. As of last August, ASU had approximately 13,000 Chinese and Japanese catalog cards to be converted. From our past experience in using RLIN's BRCON to convert more than 1,000 East Asian records, we know it took us 4 minutes on the average to create a BRCON search record consisting of the search fields and the local data that we considered necessary for a correct match and for creation of our records in the RLIN data base. Therefore, it would take 867 hours to create 13,000 BRCON search records. Using a student assistant being paid $5.00 an hour plus the BRCON startup fee of $50.00 for a current RLIN user, $4,385 would be needed to create the 13,000 BRCON search records. We also learned from experience that we could expect a 42.81 percent hit rate using BRCON and that, on the average, it took 5 minutes to review a record created through BRCON. This review process is essential for detecting matches done on wrong RLIN records, resolving word division problems, and adding information. The review should be done by a cataloger who knows the language of the record. The median hourly wage of a Library Specialist at ASU is $10.83. Review done by a Library Specialist for the 5,565 records generated from the 42.81 percent hit rate would require 463.75 hours which would cost $5,022.41. Also, each positive record derived by BRCON costs $0.65; the total cost for the 5,565 derived records would then be $3,617.25. We estimated that 22.74 percent of the 13,000 cards would not be derived by BRCON but would have romanized-only records in the RLIN data base, 33.78 percent would need to be input as new records into RLIN, and 0.67 percent would already have ASU records in RLIN. Our experience also indicated that it took 20 minutes on the average to input a new record into RLIN and that it took the same amount of time to upgrade a romanized-only record. The total of 22.74 percent and 33.78 percent of the 13,000 cards would then require 2,449.20 hours of a Library Specialist's time; this would cost $26,524.84. Hence, the total project would need $39,549.50 (= $4,385 + $5,022.41 + $3,617.25 + $26,524.84) which equals $3.04 per record for a total of 13,000 cards.

3. Specifications

The specifications are still necessary even for an in-house project in order to provide a consistency and accuracy in converting the records. But they would not be subject to an arbitrary interpretation as they otherwise would be by the vendor.

4. Communication

It is easier and more convenient to maintain effective communication when the project personnel are physically located within the same building. Questions and problems can be readily answered and resolved to prevent further complication of the problems. Communication costs can also be waived. Frustration level would be alleviated. With the project
under immediate and much better control, the work can be accomplished more efficiently and effectively.

5. Disruption of Regular Activities of the Staff Members

If existing staff members are involved in the project, a disruption to their regular duties is bound to happen. The scale of the disruption varies according to the degree and nature of participation of the members. With the present East Asian recon project undertaken by ASU which has the East Asian staff's time totally committed to complete the project in one year as mandated by the library administration, cataloging of new titles practically comes to a standstill.

III. Conclusion

Although the wholesale contract-out and in-house jobs are two basic models for retrospective conversion for East Asian language material, some libraries are adopting variations by doing the project in different phases and fashions. For instance, several years ago, the University of Michigan contracted with UTLAS to convert the University's retrospective Asian holdings into romanized form and planned later to upgrade their romanized-only records with East Asian vernacular characters added by the existing staff. The University of Minnesota started early this year the retroconversion of its East Asian collection by doing it informally as an "on the side" project and using student help to convert the catalog cards for those titles that had their records in the RLIN data base whether the records had the CJK vernacular or not, and passed them immediately into the University's Online Patron Access Catalog. The Nelson-Atkins Museum of Art in Kansas City, Missouri is doing both its recon project and reclassification of its collection at the same time. At ASU, we are using a combination of RLIN Batch Retrospective Conversion software and direct online input methods to convert our collection. We have hired two students for a total of thirty hours weekly and have been given thirty to forty hours of additional help weekly from existing staff members in other units of the department who have East Asian language backgrounds. We expect to finish the project in one year.

In preparing and refining our specifications for recon at ASU, quality control raised its head many times. Although the administrator's and the librarian's perspectives and preferences may differ, compromise always prevails.