

# EFFECTIVENESS OF NATIONAL CHINESE-JAPANESE-KOREAN SERVICE CATALOGUING

**Bick-har Yeung**

This article is a case study report which was submitted to the University on behalf of Melbourne University Library for a Staff Research/Problem Grant in January 1997. I would like to share with the East Asian library colleagues in Australia and elsewhere the outcomes of this study.

In October 1996, I was granted by the University of Melbourne Library a Staff Research/Problem Grant to conduct a research study on the effectiveness of National Chinese-Japanese-Korean Service (NCJKS) cataloguing. The project was conducted in November 1996. A total of 109 Chinese titles was selected by subject randomly from the Chinese backlog to catalogue onto the NCJKS. A time study of each cataloguing step was carried out and the hit rate for searching was recorded. Analysis of the time study and hit rates for CJK searching provided detailed information, such as performance indicators, for CJK cataloguing. During the cataloguing process, I also discovered that there were some ways in which the cataloguing process could be speeded up so as to save staff time. Furthermore, I revised the CJK cataloguing procedures of the University of Melbourne Library.

## **The National CJK Service and database**

The National CJK Service adopted the Innopac library system's cataloguing and OPAC modules to provide CJK cataloguing services for Australian and New Zealand libraries. The system went live in July 1996, and the University of Melbourne Library's East Asian Collection staff have been using the CJK Service for cataloguing Chinese and Japanese language materials since September 1996. The system enables participating libraries to use CJK scripts in searches and displays. It provides a copy cataloguing service containing CJK scripts, enabling libraries to create CJK bibliographic records which could be shared by other member libraries. Moreover, it provides links to the Research Libraries Information Network (RLIN) CJK database and allowed the importing of CJK records from RLIN into the NCJK database. It enables member libraries to import CJK bibliographic records with holdings information into the local system as well as serving as an online national union catalogue for CJK materials.

At the time of the study, the NCJK database contained 500,000 bibliographic records and 220,000 holdings statements. Ninety percent of the database contained CJK characters and romanisation, while only ten percent comprised romanised-only records. The initial load of the NCJK database was extracted from the Australian National Bibliography (ABN) database with all its CJK romanised bibliographic records and holding statements. Library of Congress (LC) full CJK bibliographic records were then loaded onto the NCJKS to overlay the romanised records. Besides LC CJK records,

there were records from a number of other sources, such as OCLC, RLIN and the National Diet Library in Tokyo.

### CJK access and software

The NCJKS was accessed through an Internet connection. To view romanised-only records, a PC was required with VT100 emulation. To view CJK characters under the Windows environment, the East Asian Collection used a CJK software package called WinMASS. However, to input and view CJK characters onto the CJK Service, we used two CJK software packages, namely, JOIN and MASS.

JOIN runs in a DOS environment. The software has a CJK character set of over 52,000 characters. It is good for Chinese character input but has poor support for Japanese character input. The East Asian Collection thus decided to use MASS to input Japanese characters onto the NCJKS. MASS runs on a Unix platform and requires Xterminal emulation software. It is user-friendly and supports both Chinese and Japanese character input.

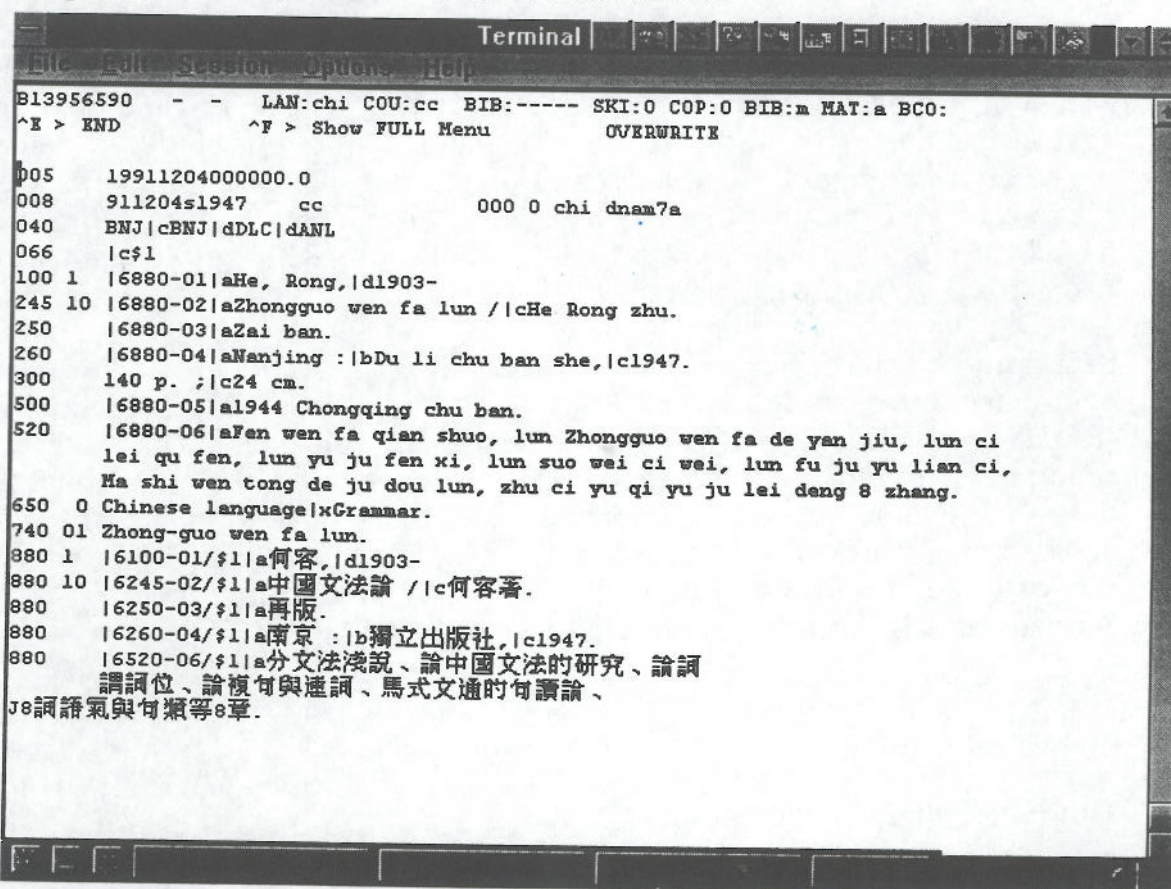
Samples of a romanised record and a full CJK records:

Sample 1: A romanised CJK record in US MARC format

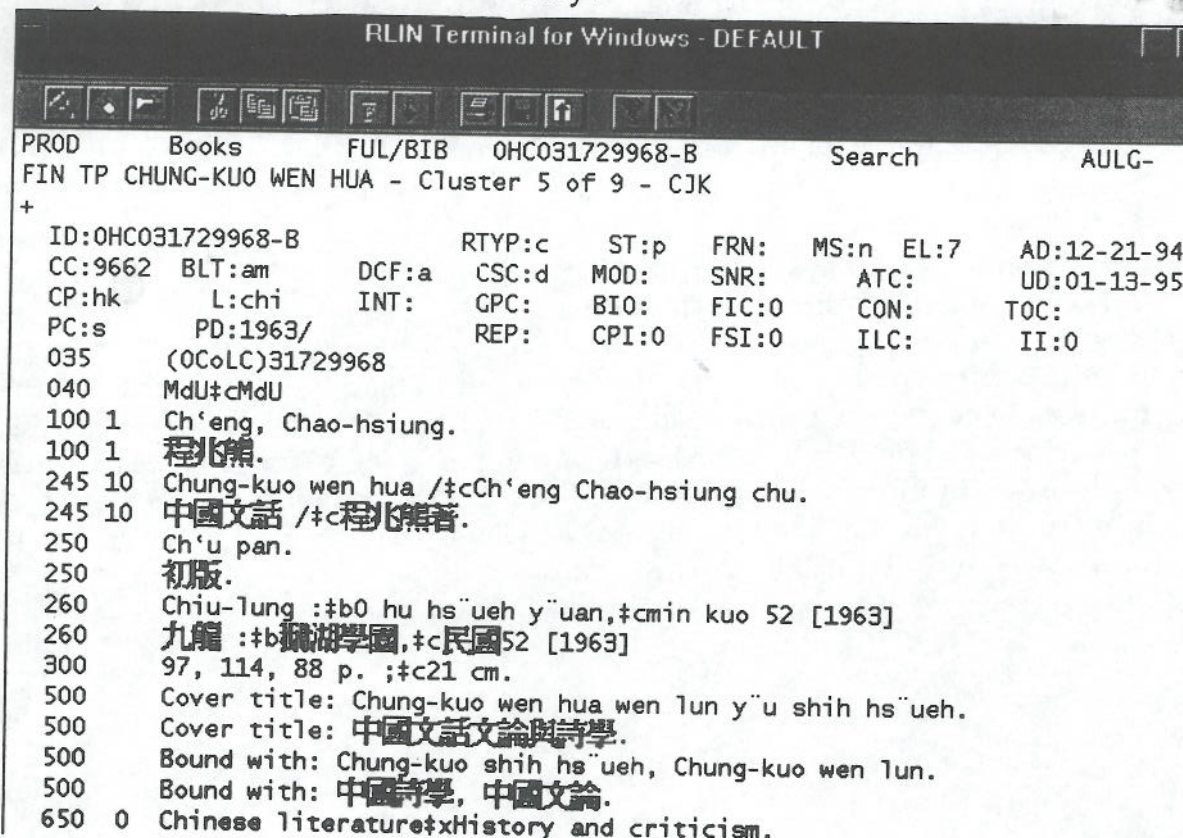
```
Terminal
B1501809x - - LAN:chi COU:cc BIB:anl SKI:0 COP:1 BIB:m MAT:a BCO:
^X > END ^F > Show FULL Menu OVERWRITE

020 7805697418 (set)
019 1 12232128
008 960508r19921935cc a 00000 chi dnam a
040 (atABN)2001|c(atABN)2001|dANL
245 10 Zhongguo yu wen xue yan jiu /|cGuang hua da xue Zhongguo yu wen xue hui
zhu.
250 Di 1 ban.
260 Shanghai :|bShanghai shu dian,|c1992.
300 1 v. (various pagings) :|bill. ;|c19 cm.
440 0 Min guo cong shu. |ndi 4 bian ;|v50.
500 Reprint.
500 "Ben shu ju Zhonghua shu ju 1935 nian ban ying yin".
501 With: Guo yu wen ti tao lun ji / Zhu Lingong bian ji. Shanghai :
Zhongguo shu ju, 1921 --Zhongguo wen fa ge xin tao lun ji / Wang Fuquan
bian ji. Shanghai : Xue shu she, 1940.
650 0 Philosophy, Chinese.
650 0 Chinese language.
650 0 Chinese language|xGrammar.
700 1 Zhu, Lingong.
700 1 Wang, Fuquan.
710 2 Guang hua da xue. |bZhongguo yu wen xue hui.
```

Sample 2: A full CJK record in US MARC format



The NCJKS provides a link for member libraries to access the RLIN database. RLIN contains 1.5 million CJK records and is the largest CJK database in the world. Records which are not found on the NCJKS can be searched from RLIN and then imported into to the NCJK database. The hit rates of CJK records in RLIN have been satisfactory.



### **Hit rates for National CJK Service and RLIN searches**

The first study of hit rates was conducted by Bick-har Yeung. One hundred and nine Chinese titles were selected randomly by subject and searched against the National CJK (NCJK) Service and RLIN. In this study, 75 percent of the books were published before 1994. On the NCJK Service there was a hit rate of 19.5 percent, while books which were not found on the NCJKS had a hit rate of 80.4 percent on RLIN. Both the NCJK Service and RLIN reached a hit rate of 82.35 percent.

Searching for the second batch of records of 108 Chinese records was conducted by a student from Royal Melbourne Institute of Technology (RMIT) who was on fieldwork at the East Asian Collection. This batch was randomly selected from recently received materials. Although publication dates of items searched were not recorded, there were known to be more books published during 1995 and 1996. We found that the NCJKS had a hit rate of 14 percent while RLIN had a hit rate of 66 percent. Together, NCJK Service and RLIN reached a hit rate of 68 percent.

The average hit rate for both the NCJKS and RLIN was 75 percent. NCJKS's performance was better than that of the ABN CJK.

### **Time study of CJK cataloguing**

A time study of CJK cataloguing was carried out by recording the time spent on every cataloguing step. During the time study periods, the study was affected by slow response time which was very frustrating. Statistics were as follows:

Average time to copy cataloguing one CJK record	= 5 minutes
Average time to assign one call number	= 3.8 minutes
Average time to create one CJK interim record using JOIN	= 14 minutes
Average time to create one Innopac item record	= 1.43 minutes
Average time to do end processing form one book	= 1.43 minutes

### **Minimum cataloguing standards for inputting CJK records to the National CJK Service**

In addition to copy cataloguing, we have been creating CJK records on the NCJKS. Because of the workload in the East Asian Collection, we are unable to perform original full-record CJK cataloguing for every East Asian title without a bibliographic record on the NCJK database. Therefore, a bibliographic level 5 (interim) record is created. This record will eventually be replaced by a full LC/RLIN CJK record.

We have been applying the following standard guidelines to create interim CJK records:

Field		Name
008		Fixed field and leader
040	VU cVU	
066	c\$1	
020/022	a	(ISBN/ISSN)
100s	??  6880-01 a	(author, if there is any)
245	10  6880-02 abc	(title)
250	6880-03 a	(edition)
260	6880-04 abc	(imprint)
300	abc	(collation)
400s	??  6880-05 av	(series, if there is any)
700s	??  6880-06 a	(joint author, if there is any)
880	??  6100s-01/\$1 a	(CJK author, if there is any)
880	10  6245-02/\$1 abc	(CJK title)
880	6250-03/\$1 a	(CJK edition)
880	6260-04/\$1 abc	(CJK imprint)
880	?  6400s-05/\$1 av	(CJK series, if there is any)
880	??  6700s-06/\$1 a	(CJK joint author, if there is any)

### Improving CJK cataloguing speed

A CJK record contains romanised data, linkages between romanised and CJK fields, as well as CJK script(s) (see record sample 2). In comparison with an ABN romanised-only CJK record (see record sample 1), creating a CJK records takes twice as long. Every year, the East Asian Collection handles approximately 3,000 volumes of Chinese and Japanese language materials. Over past years, the backlog has accumulated even with ABN copy-cataloguing. The cataloguing process needs to be speeded up. There are several ways of doing so.

#### Hints for searching

When conducting CJK searches, log into both the NCJKS and RLIN. Search on the NCJKS first; when a record is not found in the NCJK Service, copy the ISBN, toggle to RLIN and do the same search again. If the record is found, import it to the NCJK Service from RLIN. This method can save five minutes over ten searches.

#### Cloning

East Asian Collection staff have been using cloning to generate a duplicate CJK record with the same ISBN publisher code. The cloned record thus provides CJK links and imprint information for both romanised and CJK fields. We therefore only need to modify such fields as the author, title, date of publication to create one CJK interim record. We discovered that cloning a CJK record can save 15 minutes to 20 minutes per CJK interim.

### Using MASS rather than JOIN to input Chinese records

JOIN software is one option for inputting Chinese records. It has a Chinese character set of 52,000 CJK characters and provides Phonetic, Tsang-chieh, Wade-Giles, and Pinyin input methods. Of these, Tsang-chieh is the quickest input method but most difficult to learn. East Asian Staff members favour the romanisation (Wade-Giles and Pinyin) input methods. Because of the size of the Chinese character set, it is time consuming to find a character, in particular, a simplified Chinese character.

MASS is much easier for Chinese character input because MASS provides separate character sets for traditional and simplified Chinese characters. MASS also provides phrase input, which saves time. Using MASS to input Chinese records can save approximately 5 minutes per record. The East Asian Collection so far has 2 MASS licences. When the NCJKS and the local Innopac are fully implemented, there will be a need for an extra MASS licence to cope with the workload.

### **CJK cataloguing workflow**

The CJK cataloguing workflow is described below:

- Receive materials
- Perform searches on the NCJK Service
- Check VU holding if a full CJK record is found
- Add VU holding if not contained in record; record NCJK .b and .i numbers and date of searching on order slip and on book
- If a record is not found on the NCJK Service, or a romanised record is found, conduct an RLIN search. If a full CJK record is found, key in a PUT command to transfer the record to the NCJK database
- Add VU holding to the 'RLIN PUT record' the following day (it takes 24 hours for a RLIN record to be downloaded onto the NCJK Service), record NCJK .b and .i numbers on the order slip and on the book
- If a romanised record is found in RLIN or the record is not found, use cloning to create an interim CJK record onto the NCJK Service following the interim standard
- If a romanised record is found on the NCJK Service, but not upgraded CJK record exists on RLIN, then upgrade this romanised record to a full CJK record
- FTP CJK records to the VU server ( performed daily by East Asian (EA) staff members)
- Load CJK records to local Innopac Service (performed daily by EA staff members)
- Assign EA call number for the book
- Place barcode on the book and the shelflist slip
- Check duplicate EA call numbers
- Place temporary shelflist slip into the shelflist drawer
- Create Item record onto Innopac
- End processing (Book)
- File shelflist slip back into shelflist drawers and pull out temporary shelflist slip
- End

## **Lindenderry CJK backlog cataloguing plan**

The East Asian Collection was granted \$44,000 in 1997 by the Library's Lindenderry allocation to process the collection's CJK backlog. The data collected from this research study, such as hit rates and time study, were thus used to plan the CJK backlog cataloguing. A report relating to the Lindenderry CJK Backlog Cataloguing Plan was submitted to the Library in January 1997. A similar approach will also be taken in the future for the proposal for an East Asian Collection Retrospective Conversion Project.

## **Conclusion**

I would like to take this opportunity to thank the University of Melbourne Library for awarding me the Problem Solving Grant to complete this research. I would also like to thank my supervisor, Tim Meredith, who encouraged me to apply for the grant and assisted me in working out the proposal plan.

