

**AUSTRALIAN NATIONAL CJK SYSTEM USERS
PRESENTATION TO
THE KINETICA ANNUAL USERS MEETING
AUGUST 28 & 29, 2000**

Presenter : Magdalen Lee,
East Asian Collection Librarian
University of Sydney Library

BACKGROUND

Since the early 1980s the Australian Bibliographic Network (ABN), and more recently its successor Kinetica, has provided Australia with a national database that will support online cataloguing, copy cataloguing and resource sharing through union catalogue data. However, this national system is not yet able to support the use of vernacular script or characters in records for non-Roman and non-alphabetic materials.

Libraries holding material in non-Roman scripts, particularly in Asian scripts, wanted a system that would allow the storage and display of characters. The search for a library system that could accommodate Chinese Japanese and Korean characters can be traced back as early as 1991. It was the users, mainly researchers on China, Japan and Korea, who put forward the request. The Australian East Asian library community, including the National Library and the seven universities with the biggest Chinese-Japanese-Korean (CJK) collections in Australia, took up the challenge in 1993 and worked for more than three years to turn this idea into reality. On the 3rd June 1996, the Australian National Chinese Japanese Korean (ANCJK) Service database finally became operational.

AIM OF THE ANCJK SERVICE

To support Australia's closer relationship with Asia by providing Australian libraries and their clients with improved access to the distributed national collection of Chinese, Japanese and Korean materials.

Services

1. Allow libraries to create catalogue and authority data for CJK material they hold, to include CJK characters in the data, and to share the data.
2. Provide a copy cataloguing service for CJK items i.e. allow libraries to transfer to their local library systems cataloguing created by other member libraries and cataloguing loaded from sources such as the Library of Congress, RLIN and NDL. Even Kinetica users can also use ANCJK as a source of MARC records. Cataloguers are aware of the efficiencies of copy cataloguing - in the case of the much more labour-intensive CJK cataloguing, these efficiencies are even greater.

3. For member libraries whose own library systems do not support CJK characters, allow library users to search for CJK material held in their library, using CJK characters in their search terms and seeing CJK characters in their search results.
4. Allow library users and staff to search for CJK material held by other libraries, i.e., to provide a national union catalogue function.

THE ANCJK DATABASE

The ANCJK Service uses the INNOPAC system software, with special CJK software providing the necessary support for character input, storage and display. Until quite recently, the ANCJK service also supported two parallel CJK databases. The Wade-Giles database went live in June 1996, with a parallel Pinyin database introduced three months later. Historically, this was to support the two different Romanisation schemes used by participating libraries.

The ANCJK database is growing fast. In July 2000, the database contained more than 1,000,000 bibliographic records and 346,573 location records compared with 520,000 bibliographical records and 220,000 location records in June 1996.

In addition to original records being contributed online by member libraries, CJK records are loaded on a regular basis from the Library of Congress (4,500 per month), from the National Diet Library (12,000 Japanese records per month), and records selected from the RLIN database by members (about 80 per day).

Who can use it?

Member libraries:

The number of members has grown from the original 8 in 1993 to 22 members today. They include the National Library, 15 university libraries, 6 public and special libraries and 1 library consultant. Kinetica users do have access to the ANCJK databases for searching and obtaining MARC records via KineticaWeb. In addition, Kinetica users can actually see most CJK characters via KineticaWeb if they have the appropriate CJK viewing software.

Modes of access :

The ANCJK database can be accessed by using either:

- telnet - for cataloguing and searching
- Web (WebOPAC) - for searching
- Z39.50 - for searching and capturing records
- KineticaWeb - for searching, capturing records

The first three modes of access are only available to ANCJK member libraries.

ISSUES

1. The future of ANCJK (relationship between ANCJK & Kinetica)

The ANCJK database exists as a subset of the National Bibliographic Database (NBD), on a completely separate system. (ie ANCJK resides on INNOPAC, NBD on AMICUS)

One of the key issues for the future of the ANCJK Service is whether the ANCJK database is integrated with the NBD, and if so, what level of integration can be achieved. There are basically 3 options, as outlined below.

Option 1. Total separation

This is what we have at the present time. Different platforms, different software, different service model.

ANCJK users favour keeping the National CJK Service on a separate system for the following reasons:

- The current system functions well -- for cataloguing CJK materials; checking bibliographic details for reference queries and checking holdings for interlibrary loans.
- The system is user-friendly and easy to use.
- The Romanisation is uniform, i.e., in Pinyin only for Chinese records in the Pinyin database and in Wade-Giles only for Chinese records in the Wade-Giles database, not a mixture of Pinyin and Wade-Giles data as exists in the NBD.

However, there are some disadvantages to maintaining the ANCJK database as a separate system. For example,

- confusion for readers and librarians in selecting which database to search. CJK materials are reflected in both Kinetica & the ANCJK systems, and this can be confusing and misleading. The gap between the ANCJK database and the NBD is very large in term of bibliographic and holdings data. Libraries with large East Asian collections have not catalogued in Kinetica for a number of years. However some libraries still catalogue and add holdings for their CJK materials in Kinetica.

- confusion for library staff as they need to decide from which database to catalogue or to download records. Sometimes records are downloaded from the wrong database and this can cause unnecessary clean up work in local systems.
- a related issue is that some non-CJK items may contain CJK scripts in their record for various reasons (e.g. uniform title for translations, parallel titles, etc.). Some libraries add CJK scripts to the record in their local system, but this data can not be input into AMICUS. If the record is downloaded again, or in some case, comes down as part of a bibliographic product, i.e., a bibliographic change file, then the local record containing CJK scripts is replaced by the record with no CJK scripts from Kinetica.

Option 2. Full integration

In the future, CJK users would prefer to see ANCJK data integrated into one National Bibliographic Database for the convenience of users and librarians alike. However, this is still not possible and won't be possible until

- A system can be found that has full support for Unicode - the software must support CJK characters and encoding
- The ANCJK character data can be reliably mapped to Unicode (ANCJK uses the East Asian Character Code - EACC)
- The ANCJK character data can be correctly converted to Unicode.

Any plans to combine the ANCJK data and NBD data into one database must address the following issues:

- converting or deleting Wade-Giles romanisation in NBD data
- duplicate bibliographic records in the ANCJK and Kinetica.

Option 3. Partial integration

Partial integration already exists in that the CJK database can be accessed via the same interface as other NBD data, i.e., via KineticaWeb. However, the role and usefulness of the current CJK gateway in Kinetica is limited.

The current gateway to the ANCJK database via KineticaWeb does not display CJK holdings. KineticaWeb allows users to search CJK records, but does not display holdings. This can be inconvenient for users, especially reference librarians and document delivery staff, as they have to conduct a separate search from the ANCJK Webpac to look at the holdings, and they can only do this if their library is an ANCJK member library.

The inability to view ANCJK holdings via KineticaWeb is due to an incompatibility with the implementation of the Z39.50 standard between the ANCJK and Kinetica systems.

Libraries can still use the current Kinetica to CJK gateway to support copy cataloguing, and some reference inquiries.

2. How to attract more library members to use the ANCJK

At present, libraries with large East Asian collections catalogue their CJK materials in the ANCJK System while some libraries still catalogue and add holdings for their CJK materials in Kinetica.

How can other libraries be encouraged to catalogue CJK materials in the ANCJK database? Problems include:

- having staff with CJK language skills (option could be to outsource)
- not enough financial incentive - the current charging and membership structures (only Full Members can catalogue, but Full Membership charges may be prohibitive for some smaller institutions, or those with smaller CJK collections)

Recommendation:

- 2.1 That membership and charging structures be reviewed - to allow Kinetica users to use the ANCJK System without separate charging, especially for small/ public libraries, special libraries and university libraries with small CJK collections that would need to be able to justify to their managers the costs of supporting ANCJK System.

3. The impact of LC's conversion to Pinyin on Kinetica, especially the authority files

The ANCJK database uses Pinyin Romanisation for their Chinese headings whereas the NBD on Kinetica, predominantly still uses Wade-Giles Romanisation. The Library of Congress (LC), who for many years adhered to Wade-Giles Romanisation, is now adopting Pinyin.

When the updated bibliographic records (for materials in western languages with personal or geographic names in Pinyin form) are loaded into Kinetica from LC, the converted Pinyin headings from LC will probably clash with the old Wade-Giles headings in Kinetica. We know that Kinetica staff are making changes to NBD headings in line with the changes issued by LC via the LC Weekly lists. Kinetica staff will also attempt to change well-known Wade-Giles headings.

Recommendations:

- 3.1 That adequate staff resources and priority be allocated to clean up the old Wade-Giles headings in Kinetica which will clash with the converted Pinyin headings from LC.
- 3.2 That NBD staff consider the feasibility of converting all Wade-Giles headings in the NBD to Pinyin.
- 3.3 That Pinyin should be the only romanisation scheme for Chinese in Kinetica from October 1 to keep in line with LC's practice.

4. Authority file for the ANCJK

There are minimal authority records in the ANCJK at present (approximately 400 authority records on Chinese geographic names and personal names, created by the ANCJK Team and a few member libraries).

Recommendation:

- 4.1 That ANCJK staff approach LC and investigate the opportunity to acquire the LC CJK authority file for the ANCJK System when the conversion to Pinyin is finished.

5. Extra sources for copy-cataloguing

The present procedure of 'putting' RLIN records in the ANCJK is working well. It would be beneficial to users if extra sources are available for copy cataloguing, such as the OCLC database, which is a particularly rich source of CJK data.

Recommendations:

- 5.1 That ANCJK staff negotiate with OCLC for the OCLC CJK records to be downloaded into the ANCJK database, similar to 'putting' RLIN CJK records to ANCJK System.
- 5.2 That ANCJK staff contact the Library of Congress to see whether they have plans to make original records from China, Taiwan and Korea acceptable to libraries outside these regions in the same way that they have processed the National Diet Library (NDL) records.

6. New compatible CJK software

CJK users, particularly data creators, need software that will work better in both the Windows and Web environment.

The current CJK software packages are:

- JOIN software — DOS-based and not suitable for most member libraries' computer systems.
- XMASS — is excellent software for all CJK data but needs X-vision, is difficult to install, and quite expensive.
- WinMASS — not very user-friendly for inputting scripts (other than Chinese) and lacks sufficient characters for cataloguing.

Recommendation:

6.1 That the ANCJK staff improve the system software by working with Innopac and software system vendors :

- to increase support for the input of CJK data through Windows and Web based applications;
- to develop a EACC/Unicode interface which would allow users to use software that has Unicode support to input and display CJK characters.

(Innopac does offer some support for Unicode via the ANCJK WebPAC interface. However, the mapping from EACC to Unicode characters is problematic.)

7. More support/better resourcing for the ANCJK Team

Better resourcing for the ANCJK Team is required to work on issues that will improve the ANCJK System, both in system and software, in addition to the ANCJK database maintenance. Some CJK users feel it would be better to have more support from the NCJKS Team in solving software problems and installation problem. There is no reflection on the present staff in the ANCJK Team, who are most helpful and hardworking. If the Team has other non-CJK tasks, this reduces their available time to work on CJK issues.

Recommendation:

7.1 That the present staffing support for the ANCJK System be reviewed to ensure adequate quality service for CJK users.

We understand that the National Library will undertake a review of the ANCJK Service this year. We hope that the issues raised and the recommendations proposed will be addressed within this review process.

CONCLUSION

Australian libraries, together with the National Library of Australia, have ultimately created the world-first national online bibliographic database which can handle multiple non-Roman and non-alphabetic scripts - the Australian National Chinese, Japanese and Korean System.

The successes of the ANCJK System have been noted internationally and have raised the profile of Australian librarians. For example, Australian libraries' decision to change the Chinese romanisation standard from Wade-Giles to Pinyin has provided strong impetus for other libraries to follow. Also, the conversion program (called 'W2P'), which the ANCJK System used to convert Wade-Giles romanisation to Pinyin romanisation, has attracted worldwide interest from libraries that are contemplating Wade-Giles to Pinyin conversion of their own data, in particular, the Library of Congress, libraries in Hong Kong and the United Kingdom. The National Library of Australia has released this conversion program. Interested libraries can download the program as well as the user manual, and so far, more than 50 institutions have acquired the program.

Leadership shown in this field will be continued. At present, the ANCJK System handles three difficult Asian scripts (Chinese, Japanese and Korean). Perhaps other Asian and non Asian scripts can be accommodated in the near future, for example, Hebrew and Thai. Let's work further together to provide better service to library users.