

### CJK Steering Committee meets 22 April 1994

At its meeting in Canberra on 22nd April, the National CJK (Chinese, Japanese and Korean) Steering Committee made some landmark decisions. The Committee represents the eight libraries that are members of the CJK consortium: the Australian National University, Griffith University, Monash University, Murdoch University, the University of Queensland, University of Melbourne, University of Sydney and the National Library of Australia. The Committee selected a preferred tenderer for supply of the library software on which the National CJK system will be based. If contract negotiations with the tenderer are successful, the system should be installed around July or August of this year and fully operational in the last quarter.

The Committee also faced that most difficult of all problems — money. Although grants from the Department of Employment, Education and Training, through the Australian Research Council, had provided a large sum for the purchase of the system software and hardware, the grants did not fully cover the cost of buying the system. For the CJK Project to proceed, it was necessary for each of the member libraries to commit funds for the purchase of the system. After some negotiation, the libraries were able to agree on a formula for sharing the shortfall between the grant funds and the purchase cost: half of the cost would be divided equally among the user libraries and half divided according to the number of concurrent sessions each library expected to use.

The Steering Committee then turned to an even more difficult funding problem - how to pay for the operational costs of the system. Although the National Library had agreed to carry the NBD (National Bibliographic Database) costs, and all initial overhead, the remaining operational costs were significant. They included not only the monthly maintenance charges from the software and hardware suppliers but also the costs of providing staff for an ABN-style Help Desk, and for database and network support. After lengthy negotiations a compromise was reached. The member libraries agreed to pay a significant part of the operational costs for 1994/5 and to request a contribution from ABN for the remaining operational costs. It is hoped that eventually the governance of the CJK and the ABN systems will be combined, and in the much longer term, it is intended that the CJK System's data will migrate to NDIS, the redeveloped ABN.

At present the project members are all drawn from the research library sector. It is expected, however, that other libraries, including public libraries serving Chinese, Japanese or Korean communities will join, once the system is up and running. The system will be administered on behalf of the member libraries by the National Library. Although a program to market

the system to Australian libraries has not yet begun in earnest, it is planned to demonstrate the system at Australian library conferences as soon as there is something demonstrate. In the meantime, enquiries are most welcome, and should directed to:

Linda Groom  
National CJK Project Manager  
National Library of Australia  
Canberra ACT 2600

Ph           +61 6 262 1246  
Fax          +61 6 273 4327  
Email       lgroom@nla.gov.au

**The Unicode® Consortium welcomes increased Asian Support  
for the Unicode Standard™**

Press release, CUPERTINO, CA. April 27, 1994

The Unicode Standard™ is ready to address real-world issues from multilingual computing systems to distributed applications on global networks. That was the conclusion of the recent Unicode/ISO 10646 Implementers' Workshop held in Tokyo. The fifth of a series of worldwide workshops organised by The Unicode® Consortium, this Asian-base workshop was co-sponsored by Novell Japan and Microsoft Japan. Participants from nine different countries saw demonstrations of the first commercially available implementations supporting The Unicode Standard and were convinced that the handling of Asian languages is quite natural. They learned how even existing applications and data can be supported in a manner transparent to the user.

The Unicode Consortium was founded in 1991 and is dedicated to the development, maintenance, and promotion of The Unicode Standard, a sixteen-bit, fixed-width, worldwide character encoding. The Unicode Standard encodes the characters from the world's principal scripts and languages, allowing the conversion to and from all national, international, and vendor code set standards.

Commenting on increased ease of localisation, Mr. Shun-Ichi Kajisa, Manager of Windows NT™ Development for Microsoft Japan said: "Microsoft Japan has had very good experiences porting Windows NT, which is based on the Unicode standard, to the Japanese market. The system is robust and supports Japanese applications well. By avoiding the traditional method of Shift-JIS enabling in favour of Unicode, we were able to achieve significant reductions in manpower and time to market."

Companies represented at the workshop included Adobe Japan, Apple Japan, Digital Japan, Fujitsu, Hitachi, IBM Japan, Just Systems, NEC, NTT, Oracle Japan, Toshiba, and Yokogawa Hewlett-Packard.

### Catalogue of Implementations released:

In conjunction with the Tokyo Workshop, The Unicode Consortium released the first edition of its "Catalogue of Implementations" that documents an ever-growing list of products that are based on The Unicode Standard. Eighteen products are currently listed, including Windows NT for Intel and DEC Alpha processors, Netware 4.01 Directory Service™, Typographic International series Postscript Fonts™, and Galaxy Application Environment™ from Visix.

"The Unicode Standard provides easier support for international client/server software, specifically for Asian countries", said Dr. F. Avery Bishop, Program Manager for Japanese Windows NT development for Digital Equipment Corporation. "Therefore, DEC is planning on implementing Unicode on all supported platforms."

### Acceptance in China and Japan:

The Unicode standard treats all scripts equally and its suitability for the encoding needs of Asia is demonstrated by two significant events. In December 1993, the Chinese national standards organisation adopted International Standard ISO 10646 (which is code-for-code equivalent to the Unicode standard) as Chinese National Standard GB 13000. The Japanese national standards organisation is in the process of doing the same, and is also creating a set of authorised transcoding tables between the new standard and the existing Japanese Industrial Standard character encoding.

As Professor Kohji Shibano, Chair of the JISC Character Coding Committee charged with developing the new standard noted, "Japan has translated ISO 10646 to Japanese and is in the process of adopting it as a Japanese standard, JIS X-0221. In addition, [the Japanese Standards organisation] JIS is revising its existing standards to be consistent with the new ISO 10646/Unicode standard."

Mr. Yasushi Nakahara, a senior specialist in operating systems architecture at Toshiba Corporation, commented on this development. "Unicode/ISO 10646 is a promising standard for Japan as well. It will help to bring together international computing. Beyond universal character set standardisation, we hope to see further unification of international and Japanese operating systems' character handling as well as common APIs for text and internationalisation. This will be a great benefit to application writers and users by supporting worldwide portability and interoperability."

### The Unicode Consortium:

In addition to cooperating with ISO on the International Standard ISO 10646, the Consortium is responsible for providing technical information to implementers and promoting the Unicode Standard and ISO 10646.

The Unicode Consortium is comprised of major computer companies. Members of the Unicode Consortium are: Apple Computer, Digital Equipment Corporation, Hewlett-Packard Company, I.B.M. Corporation, Lotus Development Corporation, Microsoft Corporation, NeXT Computer, Novell, The Research Libraries Group, Symantec Corporation, Taligent, Unisys Corporation, and WordPerfect Corporation.

The Unicode Consortium can be reached on the Internet at [Unicode-inc@unicode.org](mailto:Unicode-inc@unicode.org), or by phone at (408) 777-5870.

The Unicode Standard and Consortium and the Unicode design mark are trademarks of Unicode Inc. All other trademarks are owned by their respective companies.

Mike Kernaghan  
Vice President, Unicode Inc.

### **Core Bibliographic Record for Non-Roman Scripts**

The Program for Cooperative Cataloging (PCC) recently posted a survey on the Internet aimed at determining the core bibliographic record necessary for non-Roman scripts. The following paragraphs explain the background to the survey, which was targeted at North American libraries.

PCC was established this past year to foster expanded cooperative cataloguing at the national level. The PCC was the outgrowth of the work of the Cooperative Cataloging Council, and has endorsed the core record standard that was developed under the aegis of that group.

A core bibliographic record is meant to be a record that can be accepted *as is* -- without any changes to *existing* elements -- and can still be enhanced depending on local priorities and resources. Core records will include more access points than minimal-level cataloguing although not necessarily all access points mandated in full level catalogue records. PCC core records will be supported by national level authority work. In descriptive elements the core record standard will require only the minimum information needed for identification of the bibliographic item. Cataloguing is done according to the standards and rule interpretations at the time the item is processed.

To date a core record has been defined for monographs, and a proposed core record for music is out for public discussion. This spring a task group was also formed by the PCC to define a core program record for cataloguing using non-Roman scripts. (That is, Chinese, Japanese, Korean (CJK), Hebrew, Yiddish, Arabic, Persian, Cyrillic, and any languages using these scripts]. The requirements for non-Roman scripts set forth by the task group are *only* for cataloguers who create original records with non-Roman scripts -- cataloguers who do not have non-Roman script capabilities for the languages they catalogue are expected to follow the core records established by the PCC, using romanisation only. The non-Roman script requirements

thus *supplement*, rather than replace, any of the PCC requirements for each bibliographic format.

#### PCC GENERIC CORE RECORD ELEMENTS

(Specifics may differ among formats)

Fixed Length Data Elements - mandatory - code fully

ISBN - Mandatory if present on item

Cataloging Source - Mandatory

Authentication Code - Mandatory

Classifications Numbers - Mandatory at least one

Main Entry - Mandatory if applicable

Description - Mandatory

Series Statement - Mandatory if present

Notes - Optional, except:

    General note - Required for source of title if not from title page

    Contents note - Required for multi-part items with separate titles

    Reproduction note - Required for reproductions

Subject Access - Required if appropriate

Added Entries - Required if applicable

#### NON-ROMAN SCRIPT CORE RECORD REQUIREMENTS

Note: PCC non-Roman script records must be contributed from systems that comply with the USMARC Format for Bibliographic Description which defines a 066 field (Character Sets Present), multiple occurrences of 880 fields (Alternate Graphic Representation) and a \$6 non-Roman linking subfield, with non-Roman unlinked fields having an occurrence number of "00" in 880 \$6.

##### 1. Non-Roman Script Description:

A paired non-Roman script field, \*transcribed\* from the item, is mandatory if applicable for the following elements and all their subfields, if present:

245 \$a,\$b,\$c,\$h,\$n,\$p	Title page transcription
250	Edition
260 \$a, \$b, \$c	Imprint
4XX	Series area

##### 2. Exception to PCC-Core Level romanized fields if non-Roman scripts used:

- a. If the full title paragraph is present in the non-Roman script, the romanized 245 may be abbreviated to the subfields a, n, and p.
- b. Contents note (for multi-part items with separate titles, field 505) may appear in \*only\* romanization \*or\* non-Roman script. (Both romanization and non-Roman script is optional.)

### 3. Non-Roman Script Access Requirements:

A paired non-Roman script field is mandatory if applicable for access points and all their subfields if present for:

- |                          |                      |
|--------------------------|----------------------|
| a. 1XX                   | Main entry           |
| b. 6XX (600,610,611,630) | Subject name entries |
| c. 7XX                   | Added entries        |
| d. 8XX                   | Added entries        |

Further information can be obtained from John Eilts.

John Eilts  
E-mail: bl.jae@rlg.stanford.edu  
The Research Libraries Group  
1200 Villa Street  
Mountain View, CA 94041-1100

### **Information Technology and Libraries Special East Asian issue**

The December 1993 issue of *Information Technology and Libraries* contains a special section dedicated to *A Decade of East Asian Scripts on RILIN* (with thanks to Hideyuki Morimoto for bringing this to our attention over Eastlib):

Aliprand, Joan M. "Introduction." *Information Technology and Libraries* (December 1993) 12 (4): 423.

Kaneko, Hideo. "RILIN CJK and the East Asian Library Community." *Information Technology and Libraries* (December 1993) 12 (4): 423-426.

Wu, Ai-Hwa. "With Characters: Retrospective Conversion of East Asian Cataloging Records." *Information Technology and Libraries* (December 1993) 12 (4): 427-431.