BEASL

Bulletin of the European Association of Sinological Librarians

Contents

Articles & Reports

Chinese Collections in Europe (4)

Reviews

Announcements

Ongoing Projects

BIBLIOGRAPHIC CONTROL IN THE PEOPLE'S REPUBLIC OF CHINA

by Huang Jungui
黃俊貴

Editor-in-Chief, *National Bibliography of China*. Director, Acquisition and Cataloguing Department, National Library of China

Editor's Note:

This translation of a paper delivered in Chinese by Mr Huang in London on 14 September 1989, during a visit to Britain as a guest of the Library Association, is based on an English version prepared by his staff. The latter has been revised and in large part rewritten on the basis of the Chinese original, references to which have been inserted where appropriate.   C.A.

CONTENTS

1. Cataloguing Sources

The depository system for publications in China has its origins in the 'Qing Code for Printed Matter' *Da Qing Yinshuawu Zhuanlü* 大清印刷物专律 enacted in the early years of the twentieth century by the imperial government. In 1916 the Ministry of Education of the Northern Warlords Administration issued a general order to the whole nation that all domestic publications be reported to the Ministry for the record, and designated the Metropolitan Library of Peking *Jing Shi Tushuguan* 京师图书馆 the precursor of the National Library of China) as the institution to receive deposit copies. In 1926, the National Government gave instructions that four copies of every book issued by publishers or private individuals be sent to the local department of education in the respective provinces for distribution as follows: one copy to the Ministry of Education, for deposit in the Metropolitan Library of Peking; one copy to the National Editing and Translation Office *Guoli Bianyi Guan* 国立编译馆, and one copy each to the relevant provincial and district library. 'Rules for Deposit of New Publications' *Xinchu Tushu Chengjiao Tiaoli* 新出图书呈缴条例, drafted by the Grand Academy *Da Xue Yuan* 大学院 came into force in 1927. Article 8 of the Publishing Law *Chubanfa* 出版法 promulgated by the National Government in 1937 designated the following bodies as depository receipt agencies: the Ministry of Internal Affairs, the Ministry of Propaganda, the Office of Local Administration, the National Central Library and the Library of the Legislative Yuan. Publications of political parties and institutions were included in the provisions of the above-mentioned law. The law was amended in 1947 to include the Information Department of the Administrative Yuan and the National Library of Peiping in the list of deposit libraries.

However, regulations for the deposit copy system were hardly enforceable in the years before 1949. As a result of the prolonged state of war in China, and publishers being concerned solely with profit seeking, deposit copies commonly failed to be submitted. Moreover, the *Guomindang* government attempted to use the deposit system as a form of censorship of books with

progressive ideas, and this was strongly resisted by the intelligentsia.

Since the founding of the People's Republic of China, the Chinese government has paid special attention to publishing statistics and registration. In order to ensure permanent and complete preservation of the cultural heritage of China, government departments have since 1951 frequently issued documents concerning the deposit of sample copies. The regulations for collecting sample copies of books, periodicals and newspapers issued by the State Bureau for Publishing Administration *Guojia chuban shiye guanli ju* 国家出版事业管理局 on 18 April 1979, restate the rule that three copies of the first edition of all new publications must be sent to the National Library of China; 'first edition' is defined to include revised, new, second, third and subsequent editions.

In recent years, 50,000 titles have been published in mainland China, of which the National Library of China has received about 30,000 as deposit copies, not including reprints, odd issues and volumes and those publications that the Library does not cover. The work of receiving deposit sample copies has become the focus of activity in the Chinese acquisitions department of the Library. Over the years, the deposit copies, which account for half the Library's accessions, have served as a basis for the acquisition of supplementary titles to balance the holdings, and also as an essential criterion to check the progress of acquisitions, of which deposits form the bulk.

2. Current Problems in Book Deposition

The depository system does not yet have the force of law in China, as publishing legislation is still at the initial stage of drafting. Though most publishers are cooperative, some are concerned only with their own profitability and are reluctant to submit the requisite copies, especially of expensive publications. There is no legal redress against publishers who refuse or neglect to send copies. Thus the reason why the deposit collection lacks completeness is simply that the system has no enforceable authority.

Coverage of deposits is narrow and not comprehensive. A certain choice is exercised in the selection of materials for deposit, but the problem remains that such materials are confined to printed documents, and audio-visual materials and microforms are not included. Moreover, the materials are limited to publications issued by the 500 or so publishing houses which are registered with the authorities. A large number of publications of government departments, scientific research institutes, production units, universities and colleges is omitted.

Depository collection is separate from uniform cataloguing and the compilation of the national bibliography. Since a nationwide bibliographic control programme has yet to be put into effect, publishers in general have no concept of bibliographic standards. They need not have their books registered, and so do not concern themselves with whether their books are entered in the national bibliography or not. Even the 'General Catalogue of Books Published in the People's Republic of China' *Quanguo zongshumu* 全国总书目 compiled by the Deposit Library of China *Zhongguo banben tushuguan* 中国版本图书馆, which is controlled by the publishing authorities, is incomplete, contains very simple entries and is subject to long delays in publication. The National Library of China, which is responsible for uniform cataloguing, is able to supply catalogue cards for only a portion of newly published books; it is unable to obtain all current publications, or to integrate these two aspects of its work.

The National Library, as the national bibliographic centre, should cover documentary information and describe all the publications published in a given period punctually, completely and precisely. For this reason, a total collection of deposit copies is a prerequisite for compiling a national bibliography. The National Library of China attempts to solve the problem of defining cataloguing sources by:

playing an active part in helping to formulate publishing legislation and commenting on the rules governing deposit samples

revising the acquisition rules of the National Library of China, clarifying and redefining acqui-

sition policy to include all Chinese materials published at home and abroad, including Taiwan, Hong Kong and Macao, and a broad range of informally published books, periodicals and reference materials;

reforming Chinese book acquisitioning and processing:

(a) by purchasing 'instant samples' *kuai yang shu* 快样书 from publishers and distributors so as to open up sources of new books and so quicken the pace of cataloguing, also using them to supplement the ordinary deposit copies;

(b) by combining the 'instant samples' with deposit copies and other acquisitions, using them to complement each other as an extended cataloguing source.

3. Bibliographic Techniques and Methodology

The techniques and methodology of modern bibliographic production are backward in China, despite its rich bibliographic heritage. In the 1960s and 70s scientific methods and technical means were applied to bibliography in the West, resulting in the perfection of MARC, which has exercised a great influence on Chinese library and information circles. Since 1974, as a result of China's open-door policy, progress in international bibliographic control has gradually been introduced to China. A National Technical Committee for Standardization of Bibliographic Documentation *Quanguo wenxian gongzuo biaozhunhua jishu weiyuanhui* 全国文献工作标准化技术委员会 was established in 1979, and China joined the International Standardization Organisation (ISO/TC46) in the same year. Under the Committee there are seven sub-committees, for bibliographic description, classified subject indexing, etc. Adhering to its principle of 'vigorous adoption of international standards', the Committee has made every effort to organize library and information circles to participate in international bibliographic control activities and to enhance the development of the Committee itself. Since that time, the standardization and automation of bibliographical work in China has started to move towards a new stage of development.

3.1 Description

For bibliographic description and related subjects, China has adopted the International Standard Bibliographic Description (ISBD) and the standards drawn up by the International Standardization Organisation as the chief sources and, on the basis of the characteristics of the Chinese language and Chinese documents, has formulated her own international standards. Twelve of these have been already promulgated and put into effect:

(1) 'General Rules for Document Description' *Wenxian zhulu zongze (GB3792.1-83)* 文献著录总则

(2) 'Descriptive Cataloguing Rules for Chinese General Books' *Putong tushu zhulu guize (GB3792.2-85)* 普通图书著录规则

(3) 'Descriptive Cataloguing Rules for Serials' *Lianxu chubanwu zhulu guize (GB3792.3-85)* 连续出版物著录规则

(4) 'Descriptive Cataloguing Rules for Non-book Materials' *Feishu ziliao zhulu guize (GB3792.4-85)* 非书著资料录规则

(5) 'Descriptive Cataloguing Rules for Archive Materials' *Dang'an zhulu guize (GB3792.5-85)* 档案著录规则

(6) 'Descriptive Cataloguing Rules for Maps' *Ditu ziliao zhumu guize (GB3792.6-85)* 地图资料著录规则

(7) 'Descriptive Cataloguing Rules for Ancient Books' *Guji zhulu guize (GB3792.7-85)* 古籍著录规则

(8) 'Descriptive Cataloguing Rules for Indexing Serial Articles' *Jiansuo qikan tiaomu zhulu guize (GB3793-83)* 检索期刊条目著录规则

(9) 'Descriptive Cataloguing Rules for Reference Matter' *Wenhou cankao wenxian zhumu guize (GB7714-87)* 文后参考文献著录规则

(10) 'Chinese International Standard Book Number' *Zhongguo guoji biaozhun shuhao (GB5795-86)* 中国国际标准书号

(11) 'Chinese Phonetic (Pinyin) Spelling for Titles of Chinese Books and Periodicals' *Zhongwen shukan mingcheng hanyu pinyin pinxiefa (GB3469-83)* 中文书刊名称汉语拼音拼写法

(12) 'Media Codes for Bibliographic Forms' *Wenxian leixing yu wenxian zaiti daima (GB3469-83)* 文献类型与文献栽体代码

In addition to the formulation of these standards, Descriptive Cataloguing Rules for Indexes, Descriptive Rules for Musical Works, etc. are ongoing and will soon be issued. The above standards will be revised before 1990 on the basis of the current version of ISBD and the ideas and comments which come from actual practice.

China joined the International Standard Book Number (ISBN) in 1982, when a China ISBN Centre was set up in the National News Publishing Bureau *Guojia xinwen chubanju* 国家新闻出版局. The national standard, the China Standard Book Number (GB5795-86) was formulated and issued in 1986, and has been put into effect since January 1987. So far, 491 publishers use this book number in their publications; it replaced the former National Uniform Book Number *Quanguo tongyi shuhao* 全国统一书号 in January 1988. The identification notation in pinyin of the fundamental classes of the Chinese Library Classification *Zhongguo tushuguan tushu fenleifa* 中国图书馆图书分类法 was added to the China ISBN. With the exception of the Industrial Technology class, which is identified by two letters, other books of various subjects are all denoted by one letter. Publishers themselves assign the title number, which is the serial number of books on the same subject published by a given publisher. The first part of the China ISBN is the principal part, i.e. the international current number and the second part comprises the classification number and the title number. They are arranged in two parallel lines or in a single line, separated by a slash, for example:

ISBN 7-144-00316-X/TP.340

China joined the International Serials Data System in 1986, and a China ISDS Centre was established in the National Library of China. At the 6th and 7th sessions of ISDS in 1986 and 1988, representatives of the Centre were elected to the Council of ISDS. By the end of 1988, the China ISDS Centre had assigned ISSNs to 1,200 serial titles published in China.

3.1.1 Chinese MARC Format

In 1982, on the basis of international standard ISO2709-73, China formulated its own international Information Exchange Tape Format *Wenxian mulu xinxi jiaohuan yong cidai geshi (GB2901-82)* 文献目录信息交换用磁带格式, which is now being revised according to ISO2709-81. This standard defines only the structure of bibliographic records on tape; not the contents of the record. A recommended format for records is contained in an appendix.

Two forms of MARC coexist in the world, UNIMARC as recommended by IFLA and CCF *Common Communication Format* as recommended by UNESCO. Which one China will adopt is not yet decided, but the general tendency is for libraries to use UNIMARC and for information agencies to use CCF. The two formats will probably coexist in China. Thus, in order to exchange bibliographical information between libraries and information agencies, interchange between the data units of UNIMARC and CCF must be perfected. In 1986 the National Library of China compiled the Chinese MARC Communication Format *Zhongguo jidu mulu tongxun geshi (GB?)* 中国机读目录通讯格式, based on UNIMARC, with modifications made necessary by the special features of Chinese documents and the Chinese language. This format has been used in the compilation of the National Bibliography. The general adoption of these standards of document description and registration by Chinese library and information circles has laid a firm foundation for the international compatibility of Chinese bibliographic data.

Authority files are an indispensible prerequisite for descriptive control. The personal names of Chinese authors, especially of ancient authors, are of an uncommonly complicated constitution, and institutions and societies very often have no fixed nomenclature. In these circumstances, standardisation is a necessity. Moreover, for the

same book in different editions or different translations with different titles, a uniform title must be established so as to ensure accuracy and completeness in catalogue arrangement and information retrieval. To this end, the National Library of China began in 1986 to compile name authority files for Classical Chinese Authors, Modern and Contemporary Chinese Authors, Transcribed Names of Foreign Authors, Names of Institutions, Organizations and Societies, and Uniform Titles. This is a gigantic project, to which considerable manpower and material resources are being devoted. The preliminary plan is for the project to be completed before 1995 and gradually brought into use thereafter.

3.2 Information Retrieval

3.2.1 Classification

In the tradition of Chinese catalogues and the retrieval habits of users, libraries and information agencies have prepared classified catalogues to serve as the principal retrieval approach for users, and have consistently emphasised the compilation of classified catalogues. Three main classification schemes are at present in use in China. They are all hierarchical classifications.

(a) The Chinese People's University Library Classification *Zhongguo renmin daxue tushuguan tushu fenleifa* 中国人民大学图书馆图书分类法 (CPULC), first edition 1953, fifth edition 1982, offers 17 classes mainly for books in the social sciences. From 1956 to 1987, the uniform book number *tongyi shuhao* 统一书号 printed on the back cover of each book published in China included one of the 17 class marks of this scheme. They were also employed in the years before 1974 on the centralised cataloguing cards issued by the National Library of China. Only a few universities and colleges still use this classification today.

(b) The Chinese Academy of Sciences Library Classification *Zhongguo kexueyuan tushuguan tushu fenleifa* 中国科学院书馆图书分类法 (CASLC), second edition 1958, offers 25 classes, and over 30,000 sub-classes. The natural sciences

are covered in some detail. This scheme is used by more than 1000 subordinate libraries and information units belonging to the Chinese Academy of Sciences system.

(c) The Chinese Library Classification *Zhongguo tushu fenleifa* 中国图书分类法 (CLC), first edition 1957, second edition 1980, third edition due 1990, offers 24 classes with 25,000 sub-classes. It has three supplements:

(1) the Chinese Book Materials Classification *Zhongguo tushu ziliao fenleifa* 中国图书资料分类法, second edition 1982, lists altogether 40,000 sub-classes, mainly for the use of information agencies;

(2) the Concise Chinese Library Classification *Zhongguo tushuguan tushu fenleifa jianben* 中国图书馆图书分类法简本, second edition 1982, is used by small and medium-sized libraries;

(3) the CLC Periodicals Classification *Zhongguo tushuguan tushu fenleifa qikan fenleibiao* 中国图书馆图书分类法期刊分类表, first edition 1987.

CLC is at present the most widely used scheme in China (by over 9,000 institutions, according to incomplete statistics) and is adopted for large-scale bibliographic and indexing projects such as centralised cataloguing, the 'National Bibliography of China' *Zhongguo guojia shumu* 中国国家书目 and the 'General Catalogue of Books Published in the People's Republic of China' *Quanguo zongshumu* 全国总书目.

For historical reasons, the above-mentioned classification schemes will continue to coexist. For classical works, the old 'Four Storehouses' *Si ku* 四库 classification and the Chinese Book Classification *Zhongguo tushu fenleifa* 中国图书分类法 devised by the late Prof. Liu Guojun 刘国钧 are still used.

3.2.2 Subject Indexing

Modern subject headings came in the wake of modern classification. China produced its first thesaurus *Zhuti zongbiao* 主题总表, compiled by Shen Zurong 沈祖荣, in 1934. In the 1950s and 60s several comprehensive and professional

thesauruses appeared. They either applied Library of Congress Subject Headings to both Chinese and foreign books, or directly adopted the U.S. NASA Thesaurus to catalogue foreign books, or concentrated solely on adapting them to the special requirements of Chinese materials. In spite of these efforts, however, none of them became popular, and they were only adopted in a few institutions. Since the 1970s, subject headings in Chinese have made considerable progress. According to preliminary statistics, 65 thesauruses of various kinds have been compiled by libraries and information departments, including three comprehensive thesauruses, 62 professional thesauruses (46 for natural sciences, engineering and technology, 16 for social sciences); 63 follow the general pattern of thesaurus construction, two are organic combinations of classification and subjects.

The Chinese Language Thesaurus *Hanyu zhuti cibiao* 汉语主题词表 compiled by the National Library of China and the Institute of Scientific and Technical Information of China *Zhongguo keji qingbao yanjiusuo* 中国科技情报研究所 (ISTIC) and published in 1979-80, is China's first large comprehensive thesaurus. The work is in three parts and ten volumes and includes 108,568 terms. The book was compiled in two divisions -: humanities and social-sciences, and natural sciences. In addition to the main tables, there are appendixes and three indexes.

The national standard 'Rules for Document Subject Indexing' *Wenxian zhuti biaoyin guize* (GB3860-83) 文献主题标引规则 was formulated and issued in 1983, with reference to the international standard ISO5963. So far, over two hundred institutions use this standard for indexing. Since 1985, the National Library of China has adopted it for centralised cataloguing, compiling and printing of catalogue cards and the National Bibliography. The widespread use of subject indexing enables libraries in China to provide their users with a new bibliographic tool and at the same time to promote the development of automatic document retrieval systems.

3.2.3 Integration of Classification and Subject Headings

The adoption of a unified retrieval language for the integration of classification and subject headings has paved the way for the compilation of a national Chinese language thesaurus. This enables classification and subject headings to be subjected to a single analytical process. The adoption of an integrated listing of classification and subject headings will simplify and improve the efficiency of classification and subject indexing. Due to the compatibility between classification numbers, subjects and subject headings, an organic connection between classification and subject retrieval systems can be established, which facilitates interchange between classification and subject data in computer retrieval systems. Since 1987, libraries and information agencies in China have started work on a 'Chinese Classified Subject Thesaurus' *Zhongguo fenlei zhuti biao* 中国分类主题表 which is a counterpart to the 'Chinese Library Classification' *Zhongguo tushuguan tushu fenleifa* 中国图书馆图书分类法 and the 'Chinese Language Subject Headings' *Hanyu zhuti biao* 汉语主题表. The first volume, 'Concordance of Classification Numbers and Subject Headings' *Fenleihao zhutici duiying biao* 分类号,主题词对应表 is already complete and work on the second volume 'Concordance of Subject Headings and Classification Numbers' *Zhutici fenleihao duiying biao* 主题词分类号对应表 is in progress. Publication is scheduled for 1990.

3.2.4 Remaining Problems in Information Retrieval

The existing classification schemes are imperfect. The libraries of China are divided into three main systems: the public library system, the Academy of Sciences library system and the college and university library system. Each of these stresses its own requirements in the application of classification and subject indexing, and though CLC is the dominant scheme, it has not replaced the other two schemes, partly for administrative and historical reasons, but mainly because CLC itself is not of a sufficient standard to win wide adherence. Many of its subject headings require revision, it is inconsistent and has other obvious faults which are incapable of amendment. All of these drawbacks greatly weaken CLC as a practical and authoritative scheme.

The study of foreign classification schemes is inadequate. Harmonisation of the existing Chinese classification schemes is a formidable task, which must include comparative study of foreign classifications. However, at present not only are there no Chinese translations of the Dewey Decimal Classification (DDC), Universal Decimal Classification (UDC) and U.S. Library of Congress Classification (LCC), but introductions to or studies of these seem also to have been neglected, and the question of harmonisation with international classifications has not yet been considered. As a consequence, none of the existing Chinese bibliographies use any of the classification numbers of the above schemes, each of which is internationally recognised, and this must affect the exchange of Chinese bibliographical information with other countries and hinder resource sharing.

Subject indexing also lacks coordination. Chinese subject indexing has a short history and a very meagre basis, but in recent years over sixty different schemes have been independently devised, each vying for supremacy. As for the quality of thesauruses, even the 'Chinese Language Thesaurus', in the compliation of which 1,378 specialist personnel from 500 institutions took part, has its faults; its subject headings are ill-chosen, it is badly organised and its unwieldy bulk makes it difficult to consult.

There is no consensus on the overall design of document retrieval. In China the main retrieval tool for users and stock arrangement is classification which also forms the main catalogue. Nonetheless, this is a less direct and effective method of retrieval than subject indexing. Since the rapid development of Chinese subject indexing in the 1970s, class and subject catalogues have run in parallel in catalogue systems, not only inconveniencing users, but also consuming libraries' and information departments' manpower and material resources. The question of whether these two catalogues will continue in parallel on an equal basis, or whether priority will be given to one, or one will supersede the other, will not be solved until consensus is reached, through long-term practice, on an overall design for document retrieval.

3.3 Catalogue Sequencing

Control of catalogue sequencing includes the handling of Chinese characters (character sets and Chinese character inputting) and filing (Chinese character arrangement and retrieval and filing rules).

3.3.1 Character Sets

The total number of Chinese characters is estimated at 60,000, of which 6,000-7,000 are in common use. A knowledge of 3,000-4,000 characters is adequate for reading newspapers and magazines, but for cataloguing purposes, a great number of Chinese characters is required, because variant forms of characters often occur in the names of authors. Titles of classical works contain more rare characters than those of modern works. Chinese characters for information exchange should first be divided into three levels on the basis of frequency, viz.: common characters, less frequently used characters and rarely used characters. Orthography and pronunciation of characters in the character set should conform to regular standards. China issued the national standard 'Basic Character Set with Chinese Character Codes for Information Exchange' *Xinxi jiaohuan yong hanzi bianma zifu ji jiben ji (GB2312-80)* 信息交換用汉字编码字符集基本集 in 1981. Taking into consideration the present status of the plan for simplifying Chinese characters, namely that the original complex characters continue in use although simplified characters are the legal orthographic norm, that libraries need to process classical books, and that documents in Chinese from Taiwan and Hong Kong are all published in the original complex characters, appropriate measures have been taken for encoding both forms, in the interest of easier handling and exchange of Chinese information. The basic character set contains 7,445 entries, of which 6,763 are Chinese characters, the remaining 682 including common symbols, the Roman alphabet, Japanese katakana and hiragana. The 6,763 Chinese characters are divided into the 3,755 most commonly used (arranged alphabetically according to the Chinese phonetic alphabet) and 3,008 less commonly used (arranged according to radicals and stroke sequence). These 6,763 characters cover more than 99% of those used for entering modern publications.

The basic character set has already been registered by the International Organization for Standardization (ISO Registration 58). 'Supplementary Sets Two and Four to the Chinese Character Code for Information Exchange' *Xinxi jiaohuan yong hanzi bianma zifu ji di'er fuzhu ji, disi fuzhu ji (GB?)* 信息交换用汉字编码字符集,第二辅助,第四辅助 were promulgated in 1986. Together with the basic set, these make up a complete simplified character set. Supplements Two and Four comprise 7,426 characters each, a total of 14,852. Though the frequency of these characters is lower than those of the basic set, (those in Supplement Four being even less common than those in Supplement Two), they are still useful. Supplements One, Three and Five, now in process of compilation, contain the corresponding original complex characters. It is estimated that Supplement One alone will meet 99.9% of the requirements for publications in the original complex characters from Taiwan and Hong Kong, and that rare characters which account for less than 1% of occurrences will be found in Supplements Two to Five. Two seven or eight bit codes are used to encode these character sets. Each character is represented by two bytes, in accordance with the national standards 'Character Set with Seven Bit Code for Information Handling and Exchange' *Xinxi chuli jiaohuan yong qi wei bianma zifu ji (GB1988-80)* 信息处理交换用七位编码字符集 and 'Method of Extending Character Set with Seven Bit Code for Information Handling and Exchange' *Xinxi chuli jiaohuan yong qi wei bianma zifu ji de kuochong fangfa (GB2371-80)* 信息处理交换用七位编码字符集的扩充方法. These two standards are identical with international standards ISO646-1983 and ISO2022-1983 respectively. The adoption of international standards creates favourable conditions for achieving compatibility and interchange between the Taiwanese CCCII, Japanese JIS6226 and non-Chinese character sets such as ASCII (American Standard Code for Information Interchange) and EBCDIC (Extended Binary Coded Decimal Interchange Code).

3.3.2 Orthography

The smallest unit of the Chinese script is the monosyllabic Chinese character. One Chinese character may represent one word, but most words in the Chinese language comprise two or more Chinese characters, which are written sequentially without word divisions. If word divisions cannot be determined in the course of reading a text, or if they are placed incorrectly, the reading will be erroneous or unintelligible. For Chinese to be the kind of natural language suitable for document retrieval, the word, the basic logical unit of language, must be used for recording and transmitting information. Chinese orthography must fulfil two tasks; the first is word division and the second is rules for transliteration and transcription between Chinese and non-Chinese languages. 'Basic Orthographic Rules of the Chinese Phonetic Notation (Trial Edition)' *Hanyu pinyin zhengcifa jiben guize (shiyonggao)* 汉语拼音正词法基本规则（试用稿）was issued in 1984 and finalized in 1987. It will be issued as a national standard 'Orthography for Writing Chinese Phonetic Notation' *Hanyu pinyin zhuyin shuxie zhengcifa* 汉语拼音注音书写正词法. Still under discussion at present, 'Rules for Word Division of Modern Chinese for Information Handling' *Xinxi chuli yong xiandai hanyu fenci guifan* 信息处理用现代汉语分词规范, now in its third draft, requires further revision. The difference between 'orthography' and 'word division' is that whereas the former has to solve the problem of notation and spelling of Chinese characters, taking words as the unit and writing them in sequence, the latter has to solve the problem of information handling of Chinese with the word as the unit, regardless of the question of spelling. However, the principle of word segmentation is the same in both cases.

As for the rules for transcription and transliteration between Chinese and non-Chinese languages, 'Orthography for Writing Chinese Phonetic Notation' requires geographical and personal names to be transcribed according to the original romanised forms; non-romanised scripts will be transcribed in accordance with the standard for the romanisation of Chinese. Rules for transcribing Chinese into roman script are already defined in international standard ISO7098, 'Romanisation of Chinese', but some additional remarks are required. The names of Chinese minorities are spelled on the basis of the national standard 'Romanization and Codes for Names of Chinese Minorities' *Zhongguo ge*

minzu mingcheng de luoma zimu pinxiefa he daima
(GB3259-82) 中国各民族名称的罗马字母拼写
法和代码. 'Transliteration and Transcription
into Chinese Phonetic Alphabet of Geographical
Names of Minorities' Shaoshu minzu diming hanyu
pinyin zimu yinyi zhuanxiefa (GB?) 少数民族地
名汉语拼音字母音译转写法 is taken as the
standard for transliterating geographical and
personal names in the languages of ethnic
minorities. International standards will be fully
adopted for rules of romanised transcription of
non-romanised scripts, including the Japanese
syllabaries, Cyrillic, Greek and Arabic alphabets,
which will be treated according to the relevant
ISO standards.

3.3.3 Character Input

Various encoding methods have now superseded
integral inputting of characters, and the develop-
ment of Chinese character recognition technology
has already made on-line handwritten input
increasingly effective.

Integral character inputting uses a tablet divided
into 3000-4000 small squares each containing
one character, i.e. equivalent to the first level
of the basic character set. Use of an electric
pointer rather than a keyboard makes for slow
speed, but the method has the advantage of direct
perception and is similar in use to the traditional
Chinese typewriter.

In recent years, no fewer than five hundred
different encoding devices for Chinese character
inputting have been devised, only about fifty of
which are truly practical. These fall into two
main types: encoding by character structure, or
by pronunciation. The majority of systems are
of the former type -, but the latter is in fact
more widely used.

The 'four corner' sijiao haoma 四角号码, 'stroke
order' hanzi bixing bianma 汉字笔形编码 and
'five stroke' wu bi zixing 五笔字性 systems are
examples of encoding by character structure.
These encoding methods break the structure of
a Chinese character down into its component
strokes, elements and radicals, each of which is
encoded in accordance with certain rules. At the
moment, the most popular is the 'five-stroke'
method. After a short training programme,

operators achieve input speeds of 50-60 characters
per minute, and the proportion of ambiguous
codes is small. However, the rules for structural
encoding are very involved and strict, and not
easy to master, which is why this method is
generally adopted by specialised personnel only.

The 'head and tail' shou wei ma 首尾码 system is
a simplified method which combines structural
with phonetic input. All that is required is to
input the stroke, element or radical at the top
and bottom of the character, plus the initial
letter of its pronunciation. In cases of ambiguity,
a menu of possible characters is presented on
the screen. This system resembles the Taiwanese
'Cang jie' 仓颉 system. It is suitable for
applications where the volume of inputting
required is not large, as the rules are straight-
forward and easy to learn.

Phonetic encoding of Chinese makes use of the
Chinese phonetic alphabet hanyu pinyin 汉语拼
音, or Chinese phonetic symbols zhuyin fuhao
注音符号. The initials and finals of Chinese
syllables are input directly using only the 26
letters of the roman alphabet or 35 phonetic
symbols. The operator must be able to spell
correctly, but there are many homophones in
Chinese and the ambiguity rate is high, at least
two choices being offered for selection on the
screen. Input speeds are correspondingly slow.
Despite this shortcoming, the system has achieved
some popularity and been adopted by some
institutions.

The phonetic/radical encoding method is mainly
phonetic, but with the addition of a structural
component. It is merely necessary to input the
pronunciation of the character plus its radical.

Input of combinations of characters as compound
words is a method which holds out promise of
improving Chinese data inputting. This method
requires a stock of encoded characters and words;
inputting individual characters or combinations
of characters produces on the screen a series of
words and phrases for selection, thus enabling
text to be input continuously rather than piece-
meal, with a consequent increase in speed and
accuracy.

Structural and phonetic encoding methods both have their advantages, and users also have different requirements; it is likely that several methods will exist side by side for the forseeable future. Computer terminals for Chinese data processing made in China will operate all the methods described above and handle all 7,000-odd characters in the basic character set.

As far as libraries are concerned, however, not only librarians but readers too have to use Chinese character terminals for data processing and retrieval. The latter find it very difficult to retrieve data, because of both unfamiliarity with the rules of structural analysis and encoding, and inability to transcribe characters into the Chinese phonetic alphabet *hanyu pinyin* 汉语拼音. On-line handwritten Chinese character recognition offers a possible solution to this problem. This method merely requires users to write the characters to be input on a tablet with a special pen; the computer recognises what has been written and inputs it automatically. This device is already on sale in China.

The 'On-line Handwritten Chinese Character Recognition System' *Lianji shouxie hanzi shibie xitong* 联机手写汉字识别系统 can recognize the 7,000-odd characters in the basic character set, at a rate of less than half a second per character. As long as the handwriting is correct, the hit rate is 100%. The size of the handwritten characters is about 15x15mm, and provided the operator writes carefully and neatly and uses standard orthography, the desired character will appear on the screen. If a mistake is made, it can be deleted and input once again. No training is required to learn this inputting method, which is therefore equally suited to readers and those librarians who have not received training in Chinese character inputting. The method holds out great promise for the future.

3.3.4 Chinese Character Sequencing and Interchange of Character Sets

Catalogues of Chinese materials, including subject, author and title catalogues, must adopt some filing sequence for sorting by automated data-handling. Filing sequences used at present are:

(1) the Chinese phonetic alphabet
(2) stroke sequence
(3) radical sequence
(4) four corner numbers.

The 'Attributive Dictionary of Chinese' *Hanyu shuxing zidian* 汉语属性字典 and the 'Chinese Character Information Handling Supporting Software System' *Hanzi xinxi chuli zhizhang ruanjian xitong* 汉字信息处理支撑软件系统, both compiled by the National Library of China, were published in 1986 after two years' trial. The dictionary contains the 6,763 Chinese characters of the basic character set, sorted by the following sequences: radicals and structural elements, strokes, Chinese phonetic alphabet and four corner numbers. These sequences serve as standard references for the manual sequencing of Chinese characters. By using the character attributes, the software automatically sorts the characters.

Chinese characters which can be handled by the software are not confined to the basic character set, but also include the various supplementary sets. In addition, they can interchange between the national standard character sets and the Taiwanese CCCII as well as other national standards, such as JIS6266, ASCII and EBCDIC. This includes conversion between the national standard Chinese phonetic alphabet *hanyu pinyin* and the Wade-Giles transcription, and provides favourable conditions for the exchange of machine readable information between China and other countries or regions. This software has been used on the M150H computer of the National Library of China, with good results.

Chinese character bibliographic sorting using this software is more than two hundred times more efficient than manual sorting, and highly accurate. In addition, the National Technical Committee for Documentation Standardisation *Quanguo wenxian gongzuo biaozhunhua jishu weiyuanhui* 全国文献工作标准化技术委员会 has already begun compilation of a national standard for bibliographical filing, including rules for arrangement of entries and references, on the basis of international standard ISO7154 'Rules for Bibliographical Filing', and taking into account the characteristics of the Chinese language. This will integrate the various methods of Chinese bibliographical filing used at present.

4. Bibliographic Systems

Bibliographic system control refers to various types of bibliographic work organized as an organic entity. Bibliographic work in China is divided among relevant institutions and editorial departments and forms different bibliographic systems, which constitute the national bibliographic enterprise.

4.1 National Bibliography

In order to fulfil its function of optimising national bibliographical work, in 1984 the National Library of China set up an editorial board for the 'National Bibliography of China' *Zhongguo guojia shumu* 中国国家书目 (NBC), which began publication in 1987. Compilation of NBC is closely linked with national publications depository work and centralised cataloguing. The main functions of the board are:

(1) to provide a comprehensive record of publications in China and thus a true reflection of literary development in a given historical period

(2) to compile and distribute centralised catalogue cards with uniform titles, standard author names, Chinese subject headings and two classification numbers (including a CLC number), as laid down by ISBD and the PRC national standards for document description, in order to provide a national bibliographic standard

(3) to publish and distribute NBC in book form, providing an up-to-date record of the nation's publications

(4) to produce NBC in machine readable form, set up and maintain a Chinese bibliographic data base and distribute Chinese machine readable bibliographies, making the National Library a centre for automation and networking.

NBC applies comprehensive definitions of 'nation' and 'language'. The definition of 'nation' has three components

(a) publications in various languages, of varied types and in varied media published within the territory of China (including Taiwan Province, Hong Kong and Macao)

(b) publications published jointly by China and other countries

(c) materials published in other countries by Chinese citizens or publishing agencies.

The concept of 'nation' here defines publications produced within the borders of China or written by authors of Chinese nationality, including those who retain their nationality while living abroad. In the case of authors who have renounced Chinese citizenship, works written while they were still Chinese nationals will be included in the retrospective NBC, but works written after their loss of citizenship will no longer be included. Certain difficulties of identification remain in this area, but every effort will be made to adhere to this definition.

The definition of 'language' includes all works written in the Chinese language and script regardless of place of publication. In addition to publications originating within the borders of China itself, the following are also included:

(a) publications in Chinese written by citizens of foreign nationality but Chinese descent

(b) sinological works written in Chinese published overseas

(c) other works written in Chinese published overseas.

There are three levels at which NBC can apply these comprehensive definitions of 'nation' and 'language':

(1) the collections of the National Library of China

(2) supplements to the collections of the National Library of China

(3) a rigorous application of the definitions of 'nation' and 'language', leaving out nothing.

Actual bibliographic experience shows that one hundred per cent comprehensive coverage (i.e. the third level above) is both impossible and unnecessary. The sheer quantity of materials in a country the size of China is immense, and in addition to formal publications, there is a great amount of casual publishing. Chinese publications published outside China, are also not inconsiderable in number. The only practical way is to attempt comprehensive coverage of publications within China itself and make an objective selection of unofficial publications in China and the Chinese publications from overseas. publications of under 15 pages and reprints are excluded. For historical reasons, publications from Taiwan, Hong Kong and Macao are not yet obtainable through the depository system, but in line with its principle of comprehensive coverage, the National Library of China endeavours to include all of the following:

(1) books, periodicals and reference materials which reflect past and present political, economic, social and cultural aspects of the above-mentioned province and regions

(2) works of eminent personages of various circles

(3) books, periodicals, reference materials and reference books on philosophy, economics, culture and the humanities of high academic and reference value.

As the quality of published materials varies, NBC places the emphasis on recording works with reference value for the state and the people. Inconsequential and ephemeral items are excluded. Naturally, this kind of selection process cannot be applied to each individual work, but only to certain categories of publication.

NBC attempts to supplement the collections of the National Library of China in order to achieve comprehensive coverage. The National Library, as the national repository, should have a basic, comprehensive collection of works from China of literary, artistic and reference value. This it succeeds in doing to a certain extent, but it cannot include everything. NBC is not in fact the catalogue of the holdings of the National

Library of China; it is more than that, and aims to include bibliographic data of every kind covered by its comprehensive definitions of 'nation' and 'language'. The bibliographic information sources of NBC thus comprise deposit copies (the collections of the National Library of China) plus other bibliographic data.

The specific coverage of NBC is: books, serials, doctoral dissertations, maps, photographs, materials in non-book form, materials in national minority languages and scripts and braille reading matter. The cumulative volume for 1985 contains mainly ordinary books and serials, but those for 1986-89 will include doctoral dissertations, maps, materials in languages and scripts of national minorities and foreign languages. As soon as conditions exist for further improvement, non-book materials and other bibliographic forms will be included, and monthly updates will be issued.

The 1985 edition of NBC was published in two volumes: text and index. The text is arranged by the classification codes of the CLC scheme. Each entry has a serial number, comprising the NBC code, year number and entry sequence number, e.g. CN-87-642. The index is in two parts: titles and authors. When conditions permit, a subject index will be added. Compilation of NBC has now been automated; compliers complete 'Chinese Bibliographic Record Entry Forms' *Zhongwen tushu shumu jilu shuru dan* 中文图书书目记录输入单 which conform to the 'Chinese MARC Communication Format', and operators input the data for automated cataloguing. The result is a diskette or tape of the NBC in machine-readable form. Centralised catalogue cards are derived from MARC and manual production of these cards, which began in the 1960s, will gradually be phased out.

4.2 General Retrospective Bibliography

The 'Bibliography of Chinese Classical and Rare Books' *Zhongguo guji shanben mulu* 中国古籍善本目录, compiled by an editorial committee and published by the Shanghai Classics Publishing House *Shanghai guji chubanshe* 上海古籍出版社, records classical works of historical, research, and artistic value, irrespective of age, held by over five thousand museums,

galleries, libraries, universities and colleges. The work will comprise five parts: Confucian Classics, History, Philosophy and Belles-lettres *Jing, Shi, Zi, Ji* 经、史、子、集 (the four traditional categories of Chinese literature), plus Collectanea *congshu* 丛书. Part I (Confucian Classics), in five volumes, was published in 1985.

The 'General Catalogue of Books Published in the Republic of China (1911-1949)' *Minguo shiqi zongshumu* 民国时期总书目, compiled by the National Library of China and published by the Bibliography and Documentation Publishing House *Shumu wenxian chubanshe* 书目文献出版社 in 20 volumes, records 100,000 book titles published between the Revolution of 1911 and September 1949. The arrangement is by subject. The volume on 'Language' has already been published.

The 'Union Catalogue of Chinese Serials in Libraries in China, 1833-1949 (Enlarged Edition)' *Quanguo zhongwen qikan lianhe mulu zengding ben* 全国中文期刊联合目录（增订本）compiled by the Union Catalogue Editorial Section, National Library of China and published by the Bibliography and Document Publishing House in 1981, records over 20,000 Chinese serials published before 1949.

The 'Draft Union Catalogue of Chinese Newspapers prior to 1949' *1949 nian qian zhongwen baozhi lianhe mulu* (caomu) 1949 年前中文报纸联合目录（草目）, compiled and printed by the Union Catalogue Editorial Section of the National Library of China in 1967, is now being revised and supplemented, and will be published.

In addition to the above, there are other union catalogues on special topics, such as the 'Union Catalogue of Chinese Local Histories' *Zhongguo difangzhi lianhe mulu* 中国地方志联合目录, 'Union Catalogue of Classical Works on Chinese Agronomy' *Zhongguo gu nongshu lianhe mulu* 中国古农书联合目录 and 'Union Catalogue of Books, Periodicals and Reference Materials on Library Science, Information Science and Archival Science' *Quanguo zhongwen tushuguanxue, qingbaoxue, dang'anxue shukan ziliao lianhe mulu* 全国中文图书馆学、情报学、档案学书刊资料联合目录.

4.3 Publishers' Catalogues

The most influential publishers' catalogues are the 'Current Bibliography of Social Sciences' *Sheke xinshumu* 社科新书目 and the 'Current Bibliography of Science and Technology' *Keji xinshumu* 科技新书目, compiled and printed by the Beijing and Science and Technology Distribution Centres of the main New China Book Store *Xinhua shudian zongdian beijing, keji faxingsuo* 新华书店总店北京、科技发行所. The former carries forthcoming publications of institutions at central government level or in the Beijing region, while the latter carries information on scientific and technical books published in Beijing, Shanghai, Tianjin and Chongqing. The Chinese Scientific and Technical Book Company *Zhongguo keji tushu gongsi* 中国科技图书公司 of Shanghai is responsibile for collating local editions of scientific and technical bibliographies nationwide and compiling the 'National Current Scientific and Technical Bibliography of Local Editions' *Quanguo difangban keji xinshumu* 全国地方版科技新书目. Some branches of *Xinhua shudian* 新华书店 compile comprehensive current bibliographies of local editions; for example, its Capital Distribution Centre *Shoudu faxingsuo* 首都发行所 compiles and prints the 'Beijing Book Information Bulletin' *Beijing tushu xinxi bao* 北京图书信息报, duplicating to some extent the work of the Beijing Distribution Centre. Others compile subject bibliographies of current local editions; for example, the 'Shanghai Current Bibliography' *Shanghai xinshumu* 上海新书目 carries new books in the social sciences published in that city. The Beijing Distribution Centre of Xinhua Shudian also compiles and prints the 'Standard Current Bibliography' *Biaozhun xinshumu* 标准新书目 and 'Subscription Catalogue for Restricted Books' *Neibushu zhengding mulu* 内部书征订目录, as well as the 'Universities and Colleges Union List' *Gaodeng xuexiao lianhe shumu biao* 高等学校联合书目表 which collates new books published by university presses throughout the country and is edited by Qinghua University. The above catalogues all function as prospectuses, and facilitate the placing of pre-publication orders.

The monthly 'National Current Bibliography' *Quanguo xinshumu* 全国新书目, compiled and

published by the China Deposit Library *Zhongguo banben tushuguan* 中国版本图书馆, records books deposited under the national deposit system. Its annual cumulation is the 'General Catalogue of Books Published in the People's Republic of China' *Quanguo zongshumu* 全国总书目, compiled by the China Deposit Library and published by the China Book Company *Zhonghua shuju* 中华书局, which records, in addition to books, newspapers and magazines newly published by Beijing or local publishing agencies that year. Though both of these are compiled after the publication of the books and the receipt of the deposit copies, descriptions are brief and these publications amount to little more than lists of titles, merely performing the function of recording the state of book production.

The 'Complete Chinese Newspapers and Periodicals' *Zhongguo baokan daquan* 中国报刊大全, Part I (periodicals distributed by the Post Office *youfa bufen*, （邮发部分）; Part II (periodicals not distributed by the Post Office *fei youfa bufen* （非邮发部分）, compiled and published by the People's Post and Telecommunications Publishing House *Renmin youdian chubanshe* 人民邮电出版社, records 5,300 current periodical and 1600 newspaper titles. Its basic function is as a subscription catalogue.

4.4 Library Catalogues

Library catalogues in China are usually arranged according to genre and format. The Chinese catalogue usually includes a classified catalogue, title catalogue, author catalogue and subject catalogue. Of these, priority is given to the classified catalogue and title catalogue, which are arranged separately rather than in the mixed arrangement of a dictionary catalogue. Subject catalogues are not yet widespread and are still at an initial stage of development, being found only in large or medium sized libraries. Topical catalogues are generally set up on a temporary basis to meet special demands and are subject to frequent modification. Periodical catalogues are mainly title catalogues, usually accompanied by a classified catalogue. Other types of documents (e.g. newspapers, maps, scientific and technical patents, standard specifications, audio-visual materials, etc.) are catalogued simply by

title or according to their own serial numbers. Library catalogues are limited to cards or ledger format; very few libraries provide readers with access to MARC records.

4.4.1 Archival catalogues

According to 1984 statistics, there are in China 2,924 different archives, all of which contain various kinds of archival catalogues, accessions catalogues, catalogues of epigraphy, paintings and books, as well as loan catalogues.

4.4.2 Information Bibliography

All information agencies in China compile secondary and tertiary documentation in the form of accessions newsletters, title catalogues, abstracts, indexes, and reviews, mostly in book form. Since 1985, information agencies have gradually been perfecting a system of periodical publications for information retrieval. A computer-based information retrieval service is available in Beijing, Shanghai and Wuhan.

4.5 Centralised Cataloguing

Since 1958, the National Library of China, the Library of the Chinese Academy of Sciences and the Library of the Chinese People's University have jointly conducted centralised cataloguing of Chinese, western-language and Russian books, and have distributed letter-press catalogue cards to libraries throughout the country. Centralised cataloguing of Russian books ceased after 1966. Since the 1970s, the introduction of LCMARC and other foreign cataloguing standards has greatly reduced the need for cataloguing in western languages. By 1984, subscribers to Chinese centralised cataloguing exceeded four thousand, and the number of cards distributed amounted to sixty million. At present, the National Library of China combines the 'National Bibliography of China' with centralised cataloguing, and produces multimedia catalogues with the same cataloguing data. About 25,000 titles of Chinese books are catalogued each year. However, because production and distribution of the cards is very slow and the cost of card paper has risen, many libraries now do their own cataloguing, and subscribers to centralised cataloguing have fallen to less than a thousand.

4.6 Books in Print

Since 1980, the Bibliography and Documentation Publishing House, Peking University Press and Qinghai People's Publishing House have co-operated on an experimental selective 'books in print' catalogue, but a complete version is still at the preparatory stage. In addition to national standards for document description and indexing, which have already been issued, other standards for cataloguing books in print are in preparation, such as 'China Standard Serial Numbers' *Zhongguo biaozhun kanhao* 中国标准刊号, 'Standard Title Page Format for Books' *Tushu shumingye geshi (GB?)* 图书书名页格式 'Standard Title Page Format for Serials' *Lianxu chubanwu timingye geshi (GB?)* 连续出版物题名页格式 and 'Data Format for Listing Books in Print' *Tushu zaiban bianmu shuju dan* 图书在版编目数据单. A copyright law is about to be enacted, which does not contain rules for registration of copyright, but it is intended that publishers will furnish final proofs of the title, copyright and contents pages to the National Library, which, in collaboration with regional centres, will be responsible for preparing a catalogue of books in print.

4.7 Catalogue Automation

Some twenty thousand computers are at present in use for bibliographic purposes in China, and as many as 36 kinds of imported MARC tapes are being tested (the majority are of the UNIMARC type). Chinese MARC formats, except for one developed in Taiwan, are modelled on UNIMARC, such as the formats developed by the National Library and Peking University Library. These two formats are largely identical, and an effort is currently being made to unite them into a single national format. In 1981 the Libraries of Peking University and the People's University began joint compilation by computer of a 'Union Accessions List of Western Language Books for the Beijing Region' *Beijing diqu xiwen xinshu lianhe tongbao* 北京地区西文新书联合通报, and established a bibliographic database. In 1988, a key science research project for an integrated library automation system, developed by the Shenzhen City Library with the sponsorship of the Libraries Division of the Ministry of Culture, was being tested. This system will offer seven software subsystems, including book purchasing, cataloguing and online information retrieval to small and medium-sized libraries by 1990. Shenzhen City Library is at present collaborating with Hunan Provincial Library, Nanjing Library and other institutions on a retrospective bibliographic database of books published after 1988. The integrated library automation system is on trial, using a HUN 386 minicomputer, in three small libraries: the Chongwen District Library in Beijing, the Jing'an District Library in Shanghai and the Shekou District Library in Shenzhen.

Some progress in introducing automated cataloguing systems has also been made in certain university and college libraries. These include the libraries of Jiaotong University, Shanghai, which has an automated cataloguing and retrieval system for Chinese and western books, Qinghua University (optical disc catalogue for western-language books), Shenzhen University, Fujian Normal University, Beijing Power Institute and Peking University, which all have their own minicomputer cataloguing systems. The Library and information systems of the Chinese Academy of Sciences have also set up cataloguing systems tailored to their particular needs; especially noteworthy are the online retrieval system of the Beijing Documentation Service *Beijing wenxian fuwu* 北京文献服务 and the international online retrieval service of the of the China Science and Technology Information Research Centre *Zhongguo keji qingbao yanjiusuo* 中国科技情报研究所.

Large databases & The Union Catalogue of Chinese Scientific and Technical Periodicals at ISTIC

This text has been edited from various sources.mainly supplied by mr. Lian Yachun 练亚纯 of the China Science and Technology Information Research Centre (ISTIC) *Zhongguo keji qingbao yanjiusuo* 中国科技情报研究所

L.F.

Part I Overview

Part II Technical Report

Part I General Look
1. Summary

The Union Catalogue of Chinese Scientific and Technical Periodicals System (UCCP) is a nation-wide computerized database management system featuring on-line Chinese character information retrieval. More than ten thousand titles of Chinese language periodicals held by the information agencies of the ministries, most of the commissions at the State Council, and at the provincial level were gathered in UCCP. Each title of a Chinese journal constitutes one record in the UCCP database, each record contains at most 34 bibliographic items such as title, editor, publisher, holdings information of participants etc. It involves almost all techniques and methods related to Chinese information processing e.g. Chinese data capture, input and output methods, data file transfer and medium-scale Chinese character database design, data search, update and maintenance.

UCCP is the largest Chinese computerized management system assembled in China so far in terms of the quantity and quality of its bibliographic data. It can be used not only for improving the conditions of domestic information retrieval for researchers and contributing to the control and coordination of national journal publication, but also for providing data sources and techniques of Chinese information processing for institutions who intend to build up a local union catalogue of Chinese journals.

This project was organized by the Bureau of Scientific and Technical Information (BSTI) of the State Science and Technical Commission (SSTC) and implemented by the Institute of Scientific and Technical Information of China (ISTIC). It was a result of international co-operation between ISTIC and the International Development Research Center (IDRC) of Canada. It took about five years to completion in October 1988. Meaningful efforts began only in the spring of 1986, although attempts were initiated in 1984.

The output of the project includes two publications and machine-readable data products:

• **A Guide to Chinese S&T periodicals** which gives a description of all Chinese journal titles,
• **The Union Catalogue of Chinese Periodicals** which describes participants holdings information, and
• **Magnetic data tapes and data diskettes** which contain all bibliographic items of the UCCP database.

2. Preface

Chinese S&T periodicals are an important part of domestic S&T information resources. The increasing more importance placed by China on scientific and technological development has led to a greater demand for scientific and technical information by researchers and scientists.

As early as 1979, ISTIC compiled a list of Chinese S&T journals, comprising 4551 titles of Chinese periodicals without holdings information. In 1982, ISTIC together with the Beijing Library and 103 other institutes produced a Union Catalogue of 18,900 western language titles with holdings information for the period 1962 to 1978. These union lists were made manually and consumed considerable manpower resources.

Three years ago (1985), BSII drew up a development plan for a computerized information system, to include the creation of more than ten Chinese-language data bases, containing some one million bibliographic records up to the year 1990. At that time, a few Chinese institutions started creating data bases of Chinese materials. Although the technical problems of handling Chinese characters on computers have not been completely solved, significant progress has been made, mainly on micro-computer. However mediumscale Chinese information processing systems on super-minicomputer like VAX are still rather rare.

To building up experience in Chinese character processing on such minicomputers and develop Chinese bibliographic data bases into multi-user systems, ISTIC started a project to set up a national Union Catalogue system of Chinese scientific and technical periodicals in the beginning of 1985, soon after the international workshop sponsored by IDRC on Chinese character processing for computerized bibliographic information exchange held in Hong Kong 17-20 Dec. 1984. This catalogue is a list of periodicals with an indication of the holdings of each participating institution and library.

This was an international cooperation project undertaken by ISTIC and IDRC. ISTIC was in charge of the implementation of the project and IDRC granted considerable financial assistance. It took nearly five years from the spring of 1984 to October 1988 to complete. The specific objectives of the project include:

1) To establish a Union Catalogue of Chinese Scientific and Technical Periodicals;

2) To develop and test the methodology for producing a Chinese-language Union Catalogue;

3) To produce a machine-readable database of Chinese scientific and technical periodicals which will be used as the basis for the input of a future national union catalogue;

4) To gain experience in Chinese character input and output processing and interfacing of related equipment;

5) To train experts in the handling of Chinese characters and computerized software development;

6) To test and promote Chinese bibliographic standard;

7) To integrate the Union Catalogue operation with ISTIC's general information automation activities.

In order to implement the aforementioned tasks, after one full year of preparation, a national workshop on the UCCP was held in Chengdu, the capital of Sichuan province, in the spring 1986. Sixty-one members of information institutes in ministries, most of the commissions of the State Council and in provincial level participated in this meeting. It was considered that holdings held by the participants represented the existing publication situation of Chinese S&T periodicals. During the three days' workshop, the general system scheme and its implementation schedule was drawn up and it was expected that the whole project would be completed by July 1988.

By implementing this project, significant technical results and experiences in Chinese character processing were gained. Almost all techniques and methods relating to the construction of medium-sized Chinese character machine-readable databases were involved, including:

1) Chinese data collection and collation;
2) Data input methods: batch mode and on-line mode;
3) Chinese data file transfer;
4) Design, setup, modification, update, indexing and retrieval of Chinese language databases.

3. Schedule of the project
(this part omitted)

4. Participants

It was estimated that the number of Chinese S&T periodicals published in China was more than 7,000, all of which could be found in information institutes around the country. For technical reasons and administrative convenience, the first batch of participants invited was limited to institutes within the information community at the ministry and the provincial level, sixty-one in all, of which fifty-six became members of UCCP. Their collections of S&T periodicals cover 95% of all publications.

List of participants:

No	Code	Name of participating institute
1	Q3060	ISTIC
2	Q3061	Chongqing Branch of ISTIC
3	Q3070	Beijing Documentation Services
4	Q3120	Institute of S&T Information of the Ministry of Security
5	Q3260	Institute of S&T Information of the Ministry of Agriculture
6	Q3270	Institute of S&T Information of the Ministry of Forestry
7	Q3320	Institute of S&T Information of the Ministry of Hydraulic & Power Engineering
8	Q3330	Institute of S&T Information of the Bureau of Architectural Technology
9	Q3350	Institute of S&T Information of the Ministry of Metallurgical Industry
10	Q3350	Institute of S&T Information of the General Corporation of Nonferrous Metal
11	Q3360	Institute of S&T Information of the Ministry of Mechanical Engineering
12	Q3380	Institute of S&T Information of the Ministry of Aeronautical Engineering
13	Q3390	Institute of S&T Information of the Ministry of Electronics
14	Q3400	Institute of S&T Information of the Ministry of Arms
15	Q3410	Institute of S&T Information of the Ministry of Space
16	Q3430	Institute of S&T Information of the Ministry of Oil Industry
17	Q3450	Institute of S&T Information of the Ministry of Textile Industry
18	Q3460	Institute of S&T Information of the Ministry of Light Industry
19	Q3470	Institute of S&T Information of the Ministry of Railway
20	Q3480	Institute of S&T Information of the Ministry of Transportation
21	Q1490	Institute of S&T Information of the Ministry of Posts and Telecommunications
22	Q3610	Institute of Medical Information of the Chinese Academy of Medicine
23	Q4130	Institute of S&T Information of the Bureau of Building Material Administration
24	Q4160	Institute of S&T Information of the Bureau of meteorology
25	Q4180	Institute of S&T Information of the Bureau of Ocean Administration
26	Q4620	Institute of S&T Information of the Bureau of Metrology
27	Q4640	Institute of S&T Information of the Bureau of Medical Administration
28	Q4660	Institute of S&T Information of the Bureau of Survey
29	Q5110	Institute of S&T Information of the General Corporation of Shipping Engineering
30	Q9110	Institute of S&T Information of Beijing
31	Q9120	Institute of S&T Information of Tianjin
32	Q9130	Institute of S&T Information of Hebei Province
33	Q9140	Institute of S&T Information of Shanxi Province
34	Q9150	Institute of S&T Information of Inner Mongolia Autonomous Region
35	Q9210	Institute of S&T Information of Liaoning Province
36	Q9210	Institute of S&T Information of Jilin Province
37	Q9230	Institute of S&T Information of Heilongjiang Province
38	Q9310	Institute of S&T Information of Shanghai
39	Q9320	Institute of S&T Information of Jiangsu Province
40	Q9330	Institute of S&T Information of Zhejiang Province
41	Q9340	Institute of S&T Information of Anhui Province
42	Q9350	Institute of S&T Information of Fujian Province
43	Q9360	Institute of S&T Information of Jiangxi Province
44	Q9370	Institute of S&T Information of Shandong Province
45	Q9410	Institute of S&T Information of Henan Province
46	Q9420	Institute of S&T Information of Hubei Province
47	Q9430	Institute of S&T Information of Hunan Province
48	Q9440	Institute of S&T Information of Guangdong Province
49	Q9450	Institute of S&T Information of Guangxi Province
50	Q9510	Institute of S&T Information of Sichuan Province
51	Q9530	Institute of S&T Information of Yunnan Province
52	Q9610	Institute of S&T Information of Shanxi Province
53	Q9620	Institute of S&T Information of Gansu Province
54	Q9630	Institute of S&T Information of Qinghai Province
55	Q9640	Institute of S&T Information of Ningxia Hui Autonomous Region
56	Q9650	Institute of S&T Information of The Xinjiang Uighur Autonomous Region

5. Disciplines, Period, Data and Standards

5.1 The disciplines covered by UCCP are:

- Science and Technology
- Psychology
- Library and Information Science
- Management Science

5.2 Period covered by UCCP

Holdings of a few participants, e.g. ISTIC, up to 1987; the rest up to the end of 1985. The updating of data is under preparation by the participants.

5.3 Bibliographic items

On the basis of the national standard "**Descriptive Cataloguing Rules for Serials**" *(Lianxu chubanwu zhulu guize (GB3792.3-85)* 连续出版物著录规则), the ISDS Manual and "Machine-readable Data Elements" recommended by the National Information Retrieval System Leading Group under BSTI, the following 34 bibliographic data items are included in the UCCP system:

- Title
    - Subtitle
    - Former title
    - Subsequent title
    - English translated title
    - Editor
    - Editorial address

- Serial numbering and date
- Publisher
    - Place of publication
- Dimensions
- Series title

- ISSN
    - China Serial Number
    - Post Office Serial Number (POSN) (youju kanhao, guonei daihao)
    - guowai daihao
    - Registration number (qikan dengji zhenghao)
    - Price & date

- Global note
- Brief introduction to journal
- Status of publication
- Type of publication

- Serial frequency
    - Publishing regularity
    - Type of item
    - Index indicator
    - Serial started date
    - Serial stop date

- Descriptors
    - Chinese Classification Number
    - English title indicator
    - ISTIC access number
    - Distributed class code
    - Worksheet number

5.4 Output

The output of this project includes two publications and two machine-readable documents for data interchange.

- *Guide to Chinese S&T Periodicals* . This publication covers most of the bibliographic data elements, up to 25 datafield contents of the UCCP record e.g title, editor and summary etc., pp 930 (2.75m characters) plus six indexes by::

    Subject
    Chinese Library Classification
    ISSN
    China Publication Regested Number
    Post Issue Number
    Export Issue Number

- *Union Catalogue of Chinese S&T Periodicals.* This publication covers 18 bibliographic data items and all holding information of participants. pp 1,145 (6m characters)..

- **Magnetic data tapes for Interchange**. containing machine-readable data from the base file in the UCCP database, created on the basis of the national standard GB 2901-88 (i.e. ISO 2709-81).

- **Data diskettes.** containing all machine-readable data from the base file in the UCCP database, based on a test format recommended by the Technical Committee on National Standardization of Documentation (TCNSD), also available in users' preferred format.

6. Major Features of the UCCP System

- Easy to update
- Easy to extend
- Full text search, all fields searchable, automatic Chinese & English indexing.

7. Statistics

The number of periodicals collected	10,014
Participants	56
Current journal (1985)	5390
Size of the bibliographic data	5 MB
Size of holdings information	9 MB
Size of total UCCP system (includes invented & thesaurus file)	22 MB
Total holdings of all participants	99,672
The total holdings of 29 general information institutes	87,858
The average holdings of general institutes	3,030
The total holdings of 27 professional institutes	11,814
The average holdings of professional institutes	438

The list of participants holdings

ISTIC	7,691
Insti. of S&T inf. of Hubei Province	5119
Insti. of S&T inf. of Hunan Province	4723
Chongqing Branch of ISTIC	4041
Insti. of S&T inf. of Liaoning Province	4032
Insti. of S&T inf. of Hebei Province	3806
Insti. of S&T inf. of Jilin Province	3766
Insti. of S&T inf. of Fujian Province	3346
Insti. of S&T inf. of Qinghai Province	3300
Insti. of S&T inf. of Sichuan Province	3280
Insti. of S&T inf. of Gansu Province	3250
Insti. of S&T inf. of Shanghai	3143
Insti. of S&T inf. of Shanxi Province	3079
Insti. of S&T inf. of Anhui Province	3038
Insti. of S&T inf. of Zhejiang Province	3,027
Insti. of S&T inf. of Jiangxi Province	3,014
Insti. of S&T inf. of Guangdong Province	2,970
Insti. of S&T inf. of Henan Province	2,950
Insti. of S&T inf. of Yunnan Province	2,871
Institute of S&T Information of Shanxi Province	2,719
Insti. of S&T inf. of Guangxi Province	2,718
Insti. of S&T inf. of Ningxia Hui Autonomous Region	2,664
Insti. of S&T inf. of Jiangsu Province	2,064
Insti. of S&T inf. of Tianjin	1,930
Insti. of S&T inf. of Inter Mongolia Autonomous Region	1,902
Insti. of S&T inf. of Heilongjiang Province	1,713
Insti. of S&T inf. of The Xinjiang Uighur Autonomous Region	1,559
Insti. of S&T inf. of Shandong Province	1,533
Beijing Documentation Services	1,172
Insti. of S&T inf. of the Ministry of Transportation	951
Insti. of S&T inf. of the Ministry of Electronics	891
Insti. of S&T inf. of the Ministry of Agriculture	886
Insti. of S&T inf. of the Ministry of Mechanical Engineering	727
Insti. of S&T inf. of Beijing	647
Institute of S&T Information of the Ministry of railway	600
Insti. of S&T inf. of the Ministry of Arms	577
Insti. of S&T inf. of the Ministry of Aeronautical Engineering	573
Insti. of S&T inf. of the Bureau of meteorology	567
Insti. of S&T inf. of the Bureau of Medical Administration	526
Institute of Medical Information of the Chinese Academy of Medicine	468
Insti. of S&T inf. of the Ministry of Forestry	434
Insti. of S&T inf. of the Ministry of Posts and Telecommunications	391
Insti. of S&T inf. of the Ministry of Hydraulic & Power Engineering	371
Insti. of S&T inf. of the General Corporation of Shipping Engineering	329
Insti. of S&T inf. of the Bureau of Architectural Technology	306
Insti. of S&T inf. of the Bureau of Ocean Administration	232
Insti. of S&T inf. of the Ministry of Metallurgical Industry	206
Insti. of S&T inf. of the Ministry of Oil Industry	205
Insti. of S&T inf. of the Ministry of Light Industry	170
Insti. of S&T inf. of the Bureau of Building Material Administration	165
Institute of S&T Information of the Bureau of Survey	134
Insti. of S&T inf. of the Ministry of Space	126
Institute of S&T Information of the Bureau of Metrology	110
Insti. of S&T inf. of the Ministry of Security	109
Insti. of S&T inf. of the Ministry of Textile Industry	98
Insti. of S&T inf. of the General Corporation of nonferrous Metal	64

8. The uses of the UCCP system

1) Providing a search tool through which one can locate the journals one needs in the main information institutes nationwide.

2) Providing bibliographic references for participants creating their own holdings tools, e.g. cards.

3) Providing machine-readable data source for institutions creating a local union catalogue

system and providing experience in database creation and management.

4) Providing a basis improving the distribution of journal publication as well as furthering the cooperation, interchange and interloan between various institutes.

5) Providing an important input data source for a future national union catalogue system.

6) Providing products for interchanging bibliographic data in form of magnetic tapes and diskettes.

9. The members of the project team and the workload

Project team:

Leader:
Lian Yachun (system design and software development)

Deputy leader:
Bao Guangwu (bibliographic data preparation and editor in chief)

Members:
Liu Chunke (computerized data handling)
Feng Haitao (bibliographic data preparation and checking)
Gu Qiang (bibliographic data preparation and checking)
Su Xiaolin (part-time. bibliographic data preparation and checking)

Work load:	Person-years
Organization administrative:	0.5
System design and analysis:	0.5
Worksheet preparations:	5.0
(8 per day)	
Holdings data preparation :	5.0
(average one person month per participant)	
Collating holdings	0.5
Data entry, correcting	4.0
Data proofreading	3.0
Programming	1.0

10. Technical supports

1) Hardware

VAX 11/750 Memory	8 MB
resident disk space of UCCP	150 MB

Chinese character terminals

VT82	2
JS200	2
DEC terminals VT220	2
Workstation IBM PC/XT	2
Proofreading printer DECLA84	1
Chinese character printer SP-8 (32X32 dots):	1

2) Software

CVMS V4.3

VAX BASIC

TRIP Full Text Database Management System (Chinese Version)

3) CPU Consumption

CPU time:	42 hours
Connection time with terminals:	1,500 hours

Part II Technical Report

Setting up a computerized union catalogue system featuring Chinese characters on a nation-wide scale involves a number of Chinese data processing techniques and methodologies. It includes data preparation, Chinese character entry, data transfer, database creation, updating and retrieval, as well as output and publication of a union catalogue.

1. Data preparation: collection and collation

A top level system diagram indicating the input, output and process is shown in Figure 1.

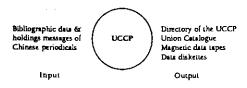| Bibliographic data & holdings messages of Chinese periodicals	UCCP	Directory of the UCCP Union Catalogue Magnetic data tapes Data diskettes
Input		Output

Figure 1. System processing diagram

The input of the UCCP consists of two parts, bibliographic information and holdings data. The disciplines concentrate on science and technology but also include a few sciences related to science and technology, e.g. psychology, management science, library and information science. Popular science periodicals are excluded. In principle, the lowest local-level editorial publishing board accepted should be at prefectural level. Basic bibliographic items are reference to the national standard and the rules specified in International Serials Data System (ISDS) Manuals.

Based on ISTIC's collections, ISTIC filled in 8,225 worksheets in advance. Each worksheet lists bibliographic items of one serial title and its holdings information of ISTIC up to 1987. Besides filling in the worksheets, ISTIC sent two volumes of holdings forms listing the worksheet number and the title of ISTIC-held serials to each of sixty-one information institutes of ministries and provinces for the collation of their holdings. Due to small holdings or other reasons, the actual number of participating institutes, including ISTIC, was fifty-six. It was estimated that all periodicals held by these institutes could basically represent the actual features of published Chinese language S&T periodicals.

For serials held by institutes other than ISTIC, which could not be found its posting in the two volumes of holdings forms, the participating institutes entered brief information such as title and publisher with their holdings information on source cards. The total number of cards received was about 55,000.

ISTIC normalized 110 volumes of holdings forms at first and then collated the cards and sorted them out manually. 2,000 journal titles were extracted from the sorted cards. ISTIC staff visited the particular institutes and filled in the corresponding bibliographic data on 2,000 worksheets. By the end of 1987, a total number of 10,350 worksheets had been made. After sorting and merging by computer, 10,013 different journal titles were recorded. This is the largest database on the Chinese S&T periodicals so far assembled in China. The worksheets were the main source of bibliographic data and the 110 volumes of holdings forms and source cards were the major source of holdings input.

2. Data entry

Data input is an important task of this project. It involves a number of operations such as Chinese character processing and Chinese database creation.

ISTIC adopted either the micro IBM PC as workstation, or J5200 Chinese terminals, (a VT200 emulated terminal) and VT82 (DEC product) on the VAX system for online data entry as well as data modification. ISTIC has modified the CDS/ISIS software provided by UNESCO to include Chinese data input by releasing the restriction on 7-bit ASCII characters to 8-bit Chinese character representation (i.e inclusion of bytes HEX(20) to HEX(7E) and HEX(AO) to HEX(FE)). The operation system widely used on IBM PC/XT and AT (and compatibles) in China is CCDOS version 2.0 to .3.0. It is a modified version of MS DOS with changes made on the display, printer and keyboard handling interrupts in the BIOS routines. CCDOS occupies around 100 KB for system and 200 KB for Chinese character patterns specified in GB-2312 (national standard basic Chinese character set). CCDOS supports a number of Chinese character utilities such as Chinese wordstar . and a large number óf input methods, of which ISTIC makes use mainly of the "five-stroke method" with an average of 2.7 key strokes per character. An average of 16 thousand Chinese characters and maximum of 22 thousand have been entered per operator each day.

Actuated by testing of the data entry, the data including all holdings were entered via four input ways into the computer:

I) 8,225 worksheets of ISTIC-held serials were entered into the computer via micro CDS/ISIS entry mode directly into its master file. CDS/ISIS database is taken as an intermediate data store. There was no need for data-entry programming. Various formats of text file were easy to output for data transfer and conversion into the database on VAX 11/750. Within a same field of CCDS/ISIS database Chinese could not mixed with English character set.

2) 2,000 worksheets of ærials held by institutes other than ISTIC were entered via wordstar into text file. There was no need to develop software for this either. Chinese characters might be mixed with English in the same field. As control characters had to be created manually, possibilities for error were obviously greater than in the previous case.

3) Holdings information of all serials was keyed in as English because it does not involve Chinese characters and could be entered in two ways: via micro PC or/and normal VAX terminals to gain text data files, and/or via online data entry operation. In the latter case the data was directly keyed in to the database management system TRIP installed on VAX, and good quality data control could be maintained.

3. Data transfer

Although the IBM PC/XT or AT microcomputer as Chinese workstation can accumulate Chinese character data, the data base of Union Catalogue had to be established on a VAX machine due to the large amount of data which can not be processed on single-user personal computers. So it was necessary to transfer the bibliographic data of all worksheets captured in micros to the VAX system via a terminal emulation file transfer program.

In 1986, DEC had already developed Chinese character software CEP on the VAX machine. Together with the VMS operating system, it is packaged to become the Chinese operating system CVMS. This software provides an input/output interface between the VAX system and the Chinese peripherals. However, the representation of the Chinese character code in both micro and VAX uses two bytes for each Chinese character with the first high-order bit (bit 8) set to one.

The data transfer program must have high reliability to ensure transfer of 8-bit ASCII code. Any bit loss from the Chinese two byte code will lead to data disruption of the subsequent Chinese characters. Three file transfer packages were used to upload the data captured in IBM PC onto VAX. Two of them resulted occasionally in lost bits and caused serious problems. The third, KERMIT package was shown to be most reliable for Chinese data telecommunication.

4. Database design of UCCP

Due to the complex nature of Chinese character processing, Chinese processing software had to be developed independently at the commencement of system design. Considering that the majority of

UCCP data (holdings information) uses western script which could be handled on many existing database mainframe systems at that time, a software package known as TRIP, which had been installed in ISTIC's VAX 11/750 machine for demonstration purposes, was applied to the collection of holdings information. This is a full text management and retrieval database system developed by Paralog AB of Sweden. Its found that its data entry facility was found easy to use to collect holdings data, especially for entry of variable length data which is the most common data type in a union catalogue system. Moreover it was able to accept Chinese characters directly without any need of modification exept for Chinese character search. Therefore, TRIP was adopted as a tool for constructing the Union Catalogue database.

TRIP is a multi-user system occupying about 200 KB memory plus 50 - 100 KB per user on the system. Its database consists of a sequence of records. A record is divided into fields, a field consists of one or more subfields or paragraphs. The paragraphs are further subdivided into sentences. They are given unique number by the system.

The size of a TRIP database is limited only by the capacity of the computer used. The size of a record and the length of fields and subfields are not restricted. Each field can be one of six different data types: text, phrase, number (with the subtype integer), date, time, and string. All of a record except its string fields can be made searchable, and in text and phrase fields Chinese may be mixed with English.

Although TRIP is a database management system, it is based on file inversion. Three files are being used, one base file and two inverted files. The base file is actually a database file where the original data are stored in sequential order. The inverted files are used for quick searching.

For text and phrase data this means that for each word in a database a sorted list of all locations where the word occurs is maintained by utility programs. Such lists are sorted in an inverted file with hashed forms of the corresponding words as entry keys. In the same file lists indicating the locations of entire phrases (the contents of a subfield of the type phrase is here regarded as a

phrase) are stored as well. For each searchable number, date, or time field there is a single entry in the inverted file. There the values for all the subfields and their locations (record, field, and subfield numbers) in the database are stored as a list sorted on location.

In addition to the above-mentioned normal inverted file there is also an inverted vocabulary file for each database. This is a small inverted file where all occurrences of a substring up to three characters are indicated. Free string searching (i.e searching for any part of a word or a phrase) is enabled by this file.

In the databases as well as the inverted files text and numeric information is stored compactly and in the inverted files the sorted lists are divided into individially compressed segments.

Within the Union Catalogue database, information on each journal is contained in one TRIP record which holds all bibliographic data as well as the holdings information of the current 56 participating institutes. For the time being, a record consists of 90 fields. For searching and output, bibliographic data holds total 34 fields and each participating holdings message holds one field. More holdings fields can be added for future new participants because the numbers of fields within a record is unlimited on TRIP.

As to the bibliographic data items their attributes of data type were assigned in following ways:

Fields and subfields, like editors and their addresses, can share the same information using pointers and large variable lenght data fields, whose maximum lenght is impossible to predict because they may contain several paragraphs of text, can easily be handled by specifying the field as being of the type 'text data'.

For the items, only whose data field or subfield might need to be repeatable, e.g. access number, title, subtitle and holdings etc., phrase data type should be chosen. The following table lists the record structure in UCCP.

Table: Record structure of UCCP

Fieldname	Fieldnumber	Datatype	Mandatory	Searchable	Name
WSN	1	phrase	y	y	worksheet number
DIC	2	phrase	y	y	distributedclass
CAN	3	phrase	n	y	ISTIC access number
COB	4	text	y	y	editorandits address
TIT	5	phrase	y	y	journal title
PRT	6	phrase	n	y	subtitle
TIP	7	text	y	y	title in English
SER	8	phrase	n	y	series title
LOP	9	phrase	n	y	place of publication
PUB	10	phrase	n	y	publisher
PSN	11	phrase	n	y	distributed number of post
XPN	12	phrase	n	y	export number of journal
RGN	13	phrase	n	y	register number
DEP	14	phrase	y	y	serial numbering and date
SIZ	15	phrase	n	y	physical description
PST	16	phrase	y	y	publishing status
TPC	17	phrase	n	y	type of publication
FRE	18	phrase	y	y	serial frequency
RGL	19	phrase	n	y	publishing regularity
TYB	20	phrase	n	y	typeof item
IDX	21	phrase	n	y	index indicator
BEY	22	phrase	n	y	beginning year
ENY	23	phrase	n	y	stopyear
PRI	24	phrase	n	y	price
FOT	25	text	n	y	former title
LAT	26	text	n	y	subsequent title
GLN	27	phrase	n	y	global note
ABS	28	text	n	y	brief description to the journal
DES	29	phrase	y	y	descriptors
CLC	30	phrase	y	y	China Library Classificaton
IDT	31	phrase	y	y	English title translated indicator
RNF	32	phrase	n	y	sorted number of title
ISN	33	phrase	n	y	ISSN number
CBN	34	phrase	n	y	Chinese issued number
Q3060	35	phrase	n	n	holdings of ISTIC
Q3061	36	phrase	n	n	holdings of Chongqing Branch
Q3070	37	phrase	n	n	holdings of BDS
Q3120	38	phrase	n	n	holdings of M. Security
Q9440	39	phrase	n	n	holdings of Guangdong Prov.
Q9140	40	phrase	n	n	holdings of Shanxi Prov.
Q9620	41	phrase	n	n	holdings of Gansu Prov.
Q9330	42	phrase	n	n	holdings of Zhejiang Prov.
Q9430	43	phrase	n	n	holdings of Hunan Prov.
Q9310	44	phrase	n	n	holdings of Shanghai
Q9340	45	phrase	n	n	holdings of Anhui Prov.
Q3320	46	phrase	n	n	holdings of M.Hydraulic &Power
Q9150	47	phrase	n	n	holdings of Inter Mongolia
Q9230	48	phrase	n	n	holdings of Heilongjiang Prov.
Q9650	49	phrase	n	n	holdings of Xingiang
Q9640	50	phrase	n	n	holdings of Ningxia
Q4180	51	phrase	n	n	holdings of Bureau of Ocean
Q9610	52	phrase	n	n	holdings of Shanxi Prov.
Q3330	53	phrase	n	n	holdings of Bureau of Archit. Tech.
Q9510	54	phrase	n	n	holdings of Sichuan Prov.
Q9530	55	phrase	n	n	holdings of Yunnan Prov.
Q9630	56	phrase	n	n	holdings of Qinghai Prov.
Q3410	57	phrase	n	n	holdings of M. Space
Q3460	58	phrase	n	n	holdings of M. Light Industry
Q4130	59	phrase	n	n	holdings of Bureau of Building Mater.
Q4660	60	phrase	n	n	holdings of Bureau of Survey
Q9210	61	phrase	n	n	holdings of Liaoning Prov.
Q9350	62	phrase	n	n	holdings of Fujian Prov.
Q3351	63	phrase	n	n	holdings of Nonfer. Metal
Q3270	64	phrase	n	n	holdings of M. Forestry
Q3260	65	phrase	n	n	holdings of M. Agricature
Q3400	66	phrase	n	n	holdings of M. Arms
Q9110	67	phrase	n	n	holdings of Beijing
Q9450	68	phrase	n	n	holdings of Guangxi Prov.
Q9420	69	phrase	n	n	holdings of Hubei Prov.
Q3350	70	phrase	n	n	holdings of M. Metallurgical
Q3360	71	phrase	n	n	holdings of M. Mechanical Eng
Q5110	72	phrase	n	n	holdings of Shipping
Q4160	73	phrase	n	n	holdings of Bureau of Meteorology
Q4640	74	phrase	n	n	holdings of Bureau of Medical Admin.
Q4620	75	phrase	n	n	holdings of Bureau of Metrology
Q9120	76	phrase	n	n	holdings of Tainjin
Q9ZO	77	phrase	n	n	holdings of Jilin Prov.
Q9320	78	phrase	n	n	holdings of Jiangsu Prov.
Q9360	79	phrase	n	n	holdings of Jiangxi Prov.
Q9370	80	phrase	n	n	holdings of Shandong Prov.
Q9410	81	phrase	n	n	holdings of Henan Prov.
Q3610	82	phrase	n	n	holdings of Academy of Medicine
Q3450	83	phrase	n	n	holdings of M. Textile
Q3430	84	phrase	n	n	holdings of M. Oil
Q3490	85	phrase	n	n	holdings of M. Post & Telecom.
Q3470	86	phrase	n	n	holdings of M. Railway
Q3380	87	phrase	n	n	holdings of M. Aeronautical Eng.
Q3390	88	phrase	n	n	holdings of M. Electronics
Q9130	89	phrase	n	n	holdings of Hebei Prov.
Q3480	90	phrase	n	n	holdings of M. Transportation

5. Creating and updating the UCCP database

Data can be entered, modified or deleted in two ways by off-line mode (batch) via a TRIP input format file, termed a TFOM file (TFOM mode) or interactively by on-line mode via the terminals (Data Entry mode).

5.1 Off-line Mode

For batch update, a text file called TFOM file containing all update data with control string identifying records, fields and subfields is used. If the record exists and the fields/ subfields do not exist, they are inserted in the record. If the record and the fields/subfields specified exist, they are updated. If no data come after the record, field or subfield identifier, the particular item is deleted from the database. As a result, participating institutes can submit to ISTIC magnetic tapes or floppy diskettes containing their updates in text files. These can be converted into TFOM files for updating the same way as the current conversion of data from C-CDS/ISIS.

5.2 On-lineMode

Alternatively, data can be updated via the terminal using data entry forms. Several entry forms can be defined for one database. Different views of a record can be restricted to different users. For example, an entry form can be designed for holdings of each institute. When input staff update holdings of a particular institute, they will not accidentally modify other data on the record. During data entry, entries can be validated against intervals, data types, specific list of

accepted values, patterns or values specified in a particular field in a specified database. Length of fields can be limited by the field length as specified on data entry forms. Records can be globally updated by commands. One can insert, delete or replace a field, a subfield, a sentence, or a paragraph in a subset of records identified either by a search result or a list of record numbers. For details, please refer to the *"User Operation Manual of UCCP"*.

6. 'Sinification' of the TRIP System

According to Chinese convention, there is no separation between words and each word can contain one or more characters. For the time being, it is no simple matter to isolate the words in a Chinese text where space is not used as a word delimiter. In order to enable Chinese character search to function on the UCCP system, ISTIC co-operates with the TRIP owner, Paralog AB to modify TRIP to deal with existing Chinese coding features on VAX VMS computers. This is what we call 'Sinification'.

The 'Sinification' consists of four major parts:

1) According to the coding rules in CVMS of VAX computer, all Chinese characters are represented in two-byte code. TRIP has to be modified to see each character in this way: if a character is read with the 7th bit value "1" then this character and the subsequent one, whether its 7th bit value is "1" or "0", will be handled as a united Chinese two-byte code.

2) Modifying TRIP scan software to regard all Chinese two-byte codes with a value of more than Hex(ACAO) as searchable. Chinese character with a Hex value less than ACCl are not indexed at all. Stop words are not operational.

3) Modifying TRIP to regard each Chinese character as a word, the indexing of mixed Chinese and English text becomes trivial. Searching for a Chinese word of two or more characters then corresponds to searching for a phrase of two or more words. As a result of this, entries in the inverted file on the database are confined to no more than seven thousands in the UCCP database. Meanwhile, the system will recognize the Chinese character for space (Hex

value 0101), period, exclamation mark and question mark as sentence separators, any one of them followed by two returns will separate paragraphs.

4) Add a Chinese menu to TRIP.

More than one year's trial and improvement based on the UCCP system with terminals handling Chinese and database of dozens of megabytes of Chinese text has shown that the 'sinification' of TRIP has been successful. In the present operation, all data fields of UCCP including the field of brief descriptions of journals (full Chinese) and holdings information have been inverted, which means that all data fields in UCCP are quickly accessible. Some interesting statistics are listed below:

- Size of the base file
  8,917,197 bytes
- Size of the inverted file (for all fields)
  11,649,024 bytes
- Average disk space of one inverted character
  1.31 bytes
- The Number of Chinese characters in the base file
  1,188,444 characters
- Size of the file for Chinese
  4,356,096 bytes
- Average disk space of one inverted Chinese char.
  3.66 bytes

As shown above, the inverted file disk space for the position of one Chinese character in a database of a few megabytes will on the average be less than 3.66 bytes. Since one Chinese character can represent one word and a normal Chinese word will on the average be less than three characters, a comparable English database (i.e one with the text translated into English) would be considerably larger. Thus it seems that a Chinese database with its inverted file implemented in the manner described here requires no more disk space than the corresponding English database and its inverted files. However, the most important advantage is that this Chinese indexing method might be a good choice for institutes which are short of professional indexing staff

7. Searching in UCCP

The VAX 11/750 computer on which UCCP is installed has been connected with a telecommuni-

cation processor, through which a remote user can access the UCCP database via a telephone line. The remote terminal has to be either DEC Chinese terminal (e.g. VT82, J5200) or IBM PC XT/AT and PC compatible. In the latter case a VT200 emulation program is needed.

Searching is done via the Common Command Language (CCL) which allows single or unlimited character truncation on both the left and the right sides. Boolean operators and comparison operators can be used. Proximity search is also available. A CCL search order such as: "Find 'software package' or 软件包" is then evaluated in the following manner:

1) Fetch the sorted lists indicating where the words *software* and *package* occur in the database from the inverted file.

2) Scan the lists in parallel and save a temporary search result, T1, indicating all positions of the word *package* where it occurs directly after the word *software*. This is the case where record, field, paragraph, and sentence numbers coincide for occurrences of *software* and *package*, and where the word number for *package* is one higher than that for *software*.

3) Fetch the sorted lists for the Chinese Characters 软 and 件 and form a temporary search result, T2 indicating all occurrences of 件 in the Chinese character string 软件 . This is done in the same manner string as in steps 1 and 2.

4) Fetch the sorted lists for the Chinese characters 包 and use this list and T2 to form a temporary search result, T3 indicating all occurrences of 包 in the Chinese character string 软件包.

5) Merge the T1 and T3 to form a search result indicating the word package in all occurrences of the phrase *software package* as well as the Chinese character 包 in all occurrences of 软件包

The TRIP-package conclusion

Since its successful adoption of the TRIP-package, ISTIC has become the licensed national distributor of this software in the PRC, and several other institutions and government bodies have applied

it to other fields. The consensus is that it is a very potent free-text database program for large databases. The ISTIC periodical file is a relatively small database in relation to the potential capacity of TRIP, which is virtually unlimited -,except by the amount of storage space available. The main problem for European clients wanting to use the Chinese language capabilities of this package is the unavailability of Chinese VAX terminals and emulation software in Europe; although both exist, maintenance could be a problem. To make the system available on the European market, the emulation software must be rewritten so that it runs properly in an 8-bit environment. Alternatively, existing software must be adapted for some of the more popular kinds of personal computer such as the IBM-PC and Macintosh.

CHINESE COLLECTIONS IN EUROPE (4)

Die chinesische Sammlung der Asien-Afrika-Abteilung der Deutschen Staatsbibliothek in Berlin.

von Helga Keller

Mit mehr als 50 000 Bänden verfügt die Asien-Afrika-Abteilung der Deutschen Staatsbibliothek über die größte und bedeutendste Sammlung chinesischer Literatur in der DDR. Diese chinesische Sammlung beruht auf den Beständen der "Libri Sinici"-Sammlung der Orientalischen Abteilung der ehemaligen Preußischen Staatsbibliothek, die infolge des 2. Weltkrieges außerordentlich schwer von Verlusten und Verlagerungen betroffen wurde. Ihr Anteil an dem heutigen Sinica-Bestand der Deutschen Staatsbibliothek beträgt 24 000 Bände. Mehr als die Hälfte der Sammlung besteht aus Neuerwerbungen aus der Zeit nach 1945, vorwiegend aus Publikationen der VR China. Die chinesische Sammlung der Asien-Afrika-Abteilung hat eine weit zurückreichende Geschichte. Der älteste chinesische Bestand stammt aus der Gründungszeit der Bibliothek, der damaligen "Churfürstlichen Bibliothek zu Cölln an der Spree, aus der zweiten Hälfte des 17 Jahrhunderts. 1665 wurden zum erstenmal chinesische Bücher in der Kurfürstlichen Bibliothek erwähnt. Spezielles wirtschaftliches Interesse für den Fernen Osten veranlaßte Friedrich Wilhelm, den Großen Kurfürsten, zur Anschaffung von chinesischen Büchern für seine Bibliothek. In seinem Auftrag und mit seiner großzügigen Unterstützung wurde dank der intensiven Bemühungen von Andreas Müller (1630-1694), Propst von Berlin, und Christian Mentzel (1622-1701), kurfürstlicher Hofrat und Leibarzt, eine zur damaligen Zeit in Europa einmalige und beachtenswerte Sinica-Sammlung zusammengetragen. Im Jahre 1683 befanden sich 25 chinesische Werke in 276 Bänden im Besitz der Berliner Bibliothek. Die Sammlung wurde durch zwei von Andreas Müller Zusammegestellte Kataloge erschlossen. Es sind die ältesten gedruckten Kataloge der Bibliothek überhaupt. Zu dieser frühen Berliner Sinica-Sammlung zählen unter anderem Schriften Jesuitischer Missionare, aber auch mehre'. Lexika, wie das Tzu-hui von Mei ying-tso, die Ssū-shu des Konfuzius, medizinische Werke sowie 2 Exemplare der chinesischen Reichsannalen, Tzu-chih-t'ung-chien. Durch den Erwerb weiterer chinesischer Bücher, darunter das chinesisch-spanische Wörterbuch des Dominikaners Francisco Diaz, besonders aber durch den Ankauf der von Mentzel hinterlassenen Sinica, überwiegend Publikationen medizinischen und botanischen Inhalts, dabei das große pharmazeutische Werk Pen-ts'ao-kang-mu, war die Sammlung bis 1702 auf insgesamt 45 Titel in 400 Bänden angewachsen. Ein Titelverzeichnis von diesem Zugang liegt nicht vor.

Nach einer Stagnation von 100 Jahren kamen durch den Orientalisten Julius Klaproth (1783-1835) in den Jahren 1810/11 zusammen mit mandschurischen und mongolischen Handschriften wieder einige chinesische Bücher in die Bibliothek, so auch das K'ang-hsi-tzu-tien in der Originalausgabe von 1716. Klaproth machte sich außerdem verdient um eine bessere Erschließung des damals vorhandenen Sinica-Bestandes. In seinem 1822 erschienenen Katalog gibt er für 57 Werke, darunter auch Mandschurica, eine ausführliche Beschreibung. Name und Titel der Bücher werden in chinesischer Schrift wiedergegeben. Unter den verzeichneten Werken finden wir auch die Romane Shui-hu-chuan, Lieh-kuo-chih und San-kuo-chih, die in illustrierten Holzblock-Editionen vom Ende des 17. Jahrhunderts vorliegen, sowie das Ming-shih in einem altem Druck von 1739.

1831 übernahm die Bibliothek von dem Münchener Gelehrten Karl Friedrich Neumann (1793-1870) einen Teil seiner in Canton mit viel Sachkenntnis angekauften Bücher, nämlich 236 Werke in 2410 Bänden. Dieser umfangreiche Zuwachs umfaßt eine ganze Anzahl wichtiger Werke aus allen Wissensgebieten. Ein handschriftliches Titelverzeichnis liegt vor. Die Neumann'sche Sammlung ging ein in die als "Codices manuscripti et libri sinici", kurz als "Libri Sinici", bezeichnete Sammlung und bildet den eigentlichen Grundstock der großen Berliner chinesischen Sammlung. Sie wurde zu einem Teil mit anderen Neuzugängen der Jahre 1822-1840 von dem Orientalisten Wilhelm Schott (1802-1889) in einem gedruckten Katalog beschrieben.

Ende des 19. und Anfang des 20. Jahrhunderts kamen folgende vier große, bedeutende Sammlungen chinesischer Literatur in die Bibliothek, die den schon vorhandenen Sinica-Beständen als geschlossene Sammlung einverleibt wurden. Durch sie konnten große Lücken im Bestand der "Libri Sinici" ergänzt werden: "Sammlung Hirth", chinesische Spezialbibliothek des Sinologen Friedrich Hirth (1845-1927), 164 Werke in 1183 Bänden und 10 Handschriften. Sie wurde 1890 für 22 500 Mark angekauft. Zu ihren wertvollsten Stücken zählen die große Enzyklopädie T'u-shu-chi-ch'eng und die 24 Reichsannalen, Nien-ssu-shih. "Sammlung Müller", von dem Orientalisten F. W. K. Müller (1863-1930) 1901 in Peking erworbene Sammlung, bestehend aus ca. 200 Werken, darunter 36 seltene chinesische Werke sowie die alte Kultausgabe des Pekinger Tripitaka in 8000 Faltbüchern, die in einem gedruckten Katalog, "Katalog des Pekinger Tripitaka", Berlin 1916, von Alfred Forke beschrieben wurde.

"Pekinger Sammlung". Sie enthält 83 zum Teil sehr kostbare chinesische und mandschurische Handschriften und Druckwerke in 1039 t'ao, die der Bibliothek 1903 "überwiesen" und 1909 in ihren Bestand eingearbeitet wurden. Offensichtlich handelt es sich um Bücher aus der Kaiserlichen Bibliothek in Peking.

"Sammlung Franke". Diese Sammlung wurde nach einer von den beiden Sinologen Otto Franke (1863-1946) und Alfred Forke (1867-1944) aufgestellten Desideratenliste in den Jahren 1907/08 von Franke in China angekauft. Sie umfaßt 114 schon lange gesuchte Einzelwerke in 774 t'ao. Hierzu gehört auch der Kyotoer Neudruck des chinesischen Tripitaka in 188 Bänden, der später gesondert aufgestellt wurde.

Die Bestände dieser Sammlungen sind im "Libri Sinici"-Katalog der Bibliothek erschlossen. Es existieren für sie, außer für die "Sammlung Müller", handschriftliche Sonderkataloge.

Die "Sammlung Franke" ist im Krieg verlorengegangen. Die anderen drei Sammlungen gehören zu den verlagerten Beständen der Deutschen Staatsbibliothek.

Die weitere systematische Komplettierung der "Libri Sinici"- Sammlung in den folgenden Jahren durch Zahlreiche wichtige Neuzugänge sowie die einheitliche und sorgfältige Erschließung des Gesamtbestandes sind das verdienst des Sinologen Hermann Hülle (1870-1940), seit 1909 verantwortlicher Bibliothekar für die Bearbeitung der Ostasiatica der Bibliothek und später Direktor der 1919 gegründeten Orientalischen Abteilung. So wurde in den Jahren 1912-14 eine weitere umfangreiche Sammlung erworben, die wegen ihrer Wichtigkeit als "Neue Sammlung" signiert wurde. Es handelt sich um 693 wertvolle und bedeutende Einzelwerke in insgesamt 11 200, Heften die nach einer von Hülle zusammengestellten Desideratenliste durch Herbert Müller vom Königlichen Museum für Völkerkunde im Auftrag der Bibliothek, in China angeschafft wurden. Darunter befanden sich eine beachtliche Anzahl von Werken zur historischen Geographie, Sammelwerke, wie das Shuo-fu, 48 Inschriften-Bücher, das wichtige Wörterbuch P'ei-wen-yün-fu in der Originalausgabe und auch ein seltenes Fragment des Tao-tsang von 1598 mit 649 Einzelschriften.

In diese "Neue Sammlung" werden seit 1913 bis zum heutigen Tag alle Neuzugänge eingearbeitet. Die einzelnen Werke werden in einem nach Numerus currens geführten Standortkatalog verzeichnet, der durch zwei Indices, einem Titel- und einem Verfasser-Index, alphabetisch erschlossen wird. Die Titelaufnahmen erfolgt seit 1901 nach der Transkription von Wade-Giles, zur Verifizierung der Autor- und Titellesungen werden die chinesischen Schriftzeichen angegeben. Der systematische Katalog, 1910 vom Hülle angelegt, ist, bis auf einige Ergänzungen, nach den vier Klassen. des Kaiserlichen Kataloges "Ssu-k'u-ch'üan-shu-tsung-mu" geordnet.

Auch in der Folgezeit ist ein ständiges Anwachsen des chinesischen Buchbestandes zu verzeichnen. So kamen unter anderem die großen Sammelwerke Ssu-pu-pei-yao und Ssu-pu-ts'ung-k'an in die Bibliothek, aber auch zeitgenössisches Schrifttum wie Schriften von Sun Yatsen und K'ang Yu-wei, Lehrbücher für den Chinesischunterricht und verschiedene Zeitschriften. Die Neuzugänge wurden laufend katalogisiert. Dank der Aktivitäten und des Arbeitseifers von Hermann Hülle wurde in Berlin eine hervorragende Sinica-Sammlung aufgebaut. Der inhaltliche Schwerpunkt der Sammlung lag auf den Gebieten der Geschichte, der Geographie, der Sprache und Literatur sowie der Inschriftenkunde,

der Archäologie und der Religion. Eingeschlossen waren auch wichtige naturwissenschaftliche, technische und medizinische Abhandlungen. Zum Bestand der "Libri Sinici"-Sammlung gehören auch einige Mandschurica.

Im Jahre 1943 besaß die Orientalische Abteilung der damaligen Preußischen Staatsbibliothek mit einem Gesamtbestand von annähernd 70 000 Bänden die wertvollste und umfangreichste Sammlung chinesischer Druck- und Handschriften in Deutschland und eine der größten und besten in Europa. Durch den 2. Weltkrieg wurde diese wertvolle Sammlung zersplittert bzw. vernichtet. Ein Drittel des Bestandes muß als Kriegsverlust betrachtet werden. Nur 24 000 Bände der "Libri Sinici"-Sammlung kehrten aus der kriegsbedingten Verlagerung in die Bibliothek Zurück.

Ein großer Teil der alten Sammlung befindet sich heute in der Ostasien-Abteilung der Staatsbibliothek Preußischer Kulturbesitz in Westberlin bzw. in der Universitätsbibliothek Krakow.

Die Beschaffung von Antiquaria bzw. von Nachdrucken älterer chinesischer Werke zur Ergänzung der "Libri Sinici"-Sammlung war äußerst schwierig, zumal ein Ankauf über den Buchhandel nicht möglich war. Durch die Übernahme der chinesischen Sammlung Otto Frankes von der Universitätsbibliothek in Berlin im Jahre 1952 konnte, zumindest auf dem Gebiet der alten chinesischen Geschichte, manche schwere Lücke geschlossen werden. Diese aus 5 988 Heften und 336 Bänden bestehende wertvolle Spezialsammlung, darunter die Enzyklopädie Ch'in-ting-t'u-shu-chi-ch'eng, wurde als "Bibliothek Franke" in den Sinica-Bestand der Deutschen Staatsbibliothek eingearbeitet. Als weitere wichtige antiquarische Erwerbung aus dem gleichen Jahr ist das große Sammelwerk Ts'ung-shu-chi-ch'eng in 3 467 Heften zu nennen. Außerdem gelang es, eine Ausgabe des Ssu-pu-pei-yao neuanzuschaffen. 1954 erhielt die Bibliothek als Donum von der Nationalbibliothek in Peking das Ta-ch'ing-li-chao-shih, einen photomechanischen Nachdruck der Mukdener Ausgabe der großen Aktensammlung der Mandschu-Zeit, gedruckt 1937 in Tokyo, 1 220 Hefte.

Seit dem Ende des 2. Weltkrieges, werden moderne chinesische Publikationen fast ausschließlich über Tausch mit der Nationalbibliothek in Peking bezogen. Von Anfang der 50er Jahre bis Mitte 1966 erhielt die Orientalische Abteilung (1969 umbenannt in Asien-Afrika-Abteilung) auf diesem Wege kontinuierlich die wichtigsten Neuerscheinungen aus der Verlagsproduktion der VR China, darunter auch Neuauflagen älterer Werke. Aus dieser Zeit ist eine an Umfang und Inhalt beachtenswerte Sammlung chinesischer Literatur in der Bibliothek vorhanden, vorwiegend aus den Gebieten der Geschichte und Politik, der Kultur, der Sprach- und Literatur- wissenschaft. Die Asien-Afrika-Abteilung verfügt außerdem über eine gute Auswahlsammlung literarischer Werke sowie über die wichtigsten chinesischen Fachzeitschriften. Die über den Schriftentausch in die Abteilung gelangte Übersetzungsliteratur ins Chinesische konnte leider nicht, wie ursprünglich geplant, zu einer Sondersammlung ausgebaut werden, da ihre Anschaffung nicht den Erwerbungsrichtlinien der Bibliothek entspricht.

Während der Jahre der Kulturrevolution kam es zu einem Stillstand bei der Erwerbung von Literatur aus der VR China. Nach einem guten Anlauf in den siebziger Jahren gingen die Tauschsendungen aus Peking in den letzten Jahren nur noch sehr sporadisch ein. Im Bestand der chinesischen Sammlung der Asien-Afrika-Abteilung entstanden dadurch schwerwiegende Lücken, zu/mal wichtige Werke, darunter auch dringend benötigte Nachschlagewerke und Wörterbücher sowie eine große Anzahl von Zeitschriften, über Tausch nicht mehr bezogen werden können.

Die moderne chinesische Literatursammlung gehört zu den Beständen der "Libri Sinici"-Sammlung. 1960 wurde mit dem Aufbau eines neuen Sachkataloges für die nach 1945 eingegangene Literatur begonnen. Der Neuzugang ist teilweise nur provisorisch erschlossen bzw. noch nicht katalogisiert. Die Benutzung der unbearbeiteten Werke ist durch die systematische Aufstellung dieses Teils des Bestandes gewährleistet.

Die Bestände der "Libri Sinici"-Sammlung werden für die Leser im Lesesaal der Asien-Afrika-Abteilung zur Benutzung bereitgestellt.

Bibliographie

Auster, Guido: Die Orientalische Abteilung. In: Deutsche Staatsbibliothek 1661-1961. Leipzig 1961. Bd 1, S. 275-317

Hülle, Hermann: Die chinesischen Neuerwerbungen der Königlichen Bibliothek. In: ZfB (Zentralblatt für Bibliothekswesen) 1915, 32, S. 221-228

Hülle, Hermann: Die Fortschritte der Ostasiatischen Sammlung. In: Fünfzehn Jahre Königliche und Staatsbibliothek. Berlin 1921. S. 190-211

Hülle, Hermann: Neuerwerbungen chinesischer und manjurischer Bücher in den Jahren 1921-1930. Leipzig 1931

Klaproth, Julius: Verzeichnis der chinesischen und mandshuischen Büchern und Handschriften der Königlichen Bibliothek zu Berlin. Paris 1922

Kraft, Eva: Die chinesische Büchersammlung des Großen Kurfürsten und seines Nachfolgers. In: China und Europa, Ausstellungskatalog. 1973. S. 18-25

Kraft, Eva: Frühe chinesische Studien in Berlin. In: Medizinhistorisches Journal 1976, 11, 1/2, S. 92-128

Müller, Andreas: Anderer Teil des Catalogi der Sinesischen Bücher bey der Churfürstlichen Brandenburgischen Bibliothec zu Cölln an der Spree. Cölln 1863

Müller, Andreas: Catalogus Librorum Sinicorum Bibliothecae Electoralis Brandenburgicae. In: Fünfzehn Jahre Königliche und Staatsbibliothek. Berlin 1921. S. 193

Schott, Wilhelm: Verzeichnis der Chinesischen und Mandshu-Tungusischen Bücher und Handschriften der Königlichen Bibliothek zu Berlin. Berlin 1840

REVIEWS

Singular Listlessness
A Short History of Chinese Books and British Scholars
by T. H. Barrett

Prof. Barrett takes his title from Sir John Davis' description in 1822 of his fellow Britons' attitude to China and things Chinese, which he contrasted unfavourably with that of the French. Surveying the state of Chinese studies in Great Britain 167 years on, the phrase seems equally apt.

In this book, the first monograph ever written on the subject, the author conducts us at a brisk pace through the history of British sinology and Chinese book collecting, which persisted in the face of stubborn indifference from government and the educational establishment. The story is enlivened by a number of encounters with colourful and eccentric individuals, of the kind so often found in sinological circles. These include Peter Mundy, the Cornish adventurer who landed at Canton in 1673 and was possibly the first Englishman to learn Chinese, which was taught him by one 'Antonio the Caffer', a fugitive from Portuguese slavery in Macao; Shen Fu-tsung, who catalogued the Chinese books in the Bodleian Library and was remembered by King James II as 'a little blinking fellow'; James Flint, 'linguist' to the East India Company, whose Chinese tutor was executed for helping him to draft a petition to the Emperor in 1759; and Thomas Manning of Caius College, Cambridge, who about the year 1798 began to study Chinese to confirm 'certain ideas he had concerning the Greek particles', in pursuit of which he walked as far as Tibet, where he had an audience with the Dalai Lama.

The contributions of such scholars as James Legge and Sir Thomas Wade are well known, but in these pages we meet some of the lesser lights of British sinology, like Samuel Kidd, Legge's teacher, who catalogued San kuo yen i as a 'statistical work' and E. H. Parker, first (and last) Professor of Chinese at Manchester University, 'a prolific but resolutely mediocre scholar' (what pathos in that

'resolutely'). Even the redoubtable Herbert Giles strayed into dubious speculations about 'Traces of Aviation in Ancient China'.

Arriving at the twentieth century, Prof. Barrett paints a dismal picture, contrasting the fabulous endowments for Chinese studies in the United States with their abject condition in Britain, where lack of enthusiasm for institutionalised sinology had spread to scholars like Arthur Waley, who, canvassed for the Chair of Chinese at Cambridge, is said to have exclaimed, 'I would rather be dead.'

This book is well worth reading for its anecdotes alone, told in a humorous vein seldom found in sinological writing since the late George Kennedy. But the author has a serious purpose. Inadequate investment in library resources, lack of research opportunities and few openings for academic employment (the latter now improving, following a recent government initiative), mean that Britain cannot find native scholars to fill vacant posts. In the present climate of antiintellectualism and philistinism, it comes as no surprise to learn that Chinese studies are not flourishing. Prof. Barrett laments the loss of 'a generation of students' who, 'despairing of academic prospects have turned to diplomacy, commerce or posts overseas'. Perhaps, even in these disparate fields, Chinese graduates can contribute to mutual understanding and cultural interchange.

Here a cogent and impassioned argument is made for the study of Chinese culture, no less for practical reasons than for its intrinsic value, and for the maintenance of comprehensive library resources, without which serious research is impossible. Those of us who spend our lives in the 'daily grind of sinological librarianship' are fortunate to have so sympathetic and persuasive an advocate. The munificence recently showered upon Oxford may be a sign that his voice has not gone unheard.

This is an elegantly produced little volume, embellished on the cover with enigmatic blackboard jottings from James Legge's last lecture, piously preserved at Oxford. The collected signatures of eminent sinologues on the rear endpaper are a disappointment, and somewhat incongruously include 'Henry' Pu-yi's inscription

for Sir R. F. Johnston's retreat. The text unfortunately is marred by too many misprints for a book of only 125 pages; I noted eleven in addition to the nineteen on the publisher's errata slip. It is to be hoped that a second edition will be called for in which these minor defects can be corrected.

C.A.

Chinese Cookbooks
An Annotated English Language Compendium / Bibliography. by Jaqueline M. Newman

The book is a bibliography of all the Chinese cookbooks and pamphlets on Chinese cookery in the English language known to the author. The book includes 732 entries and most of them are claimed to be in the author's possession awaiting the eager researcher. Authors who use two different nameforms, for instance an English and a Chinese name have their works listed twice each having its own entry and individual number and with a crossreference to the corresponding item, this is also the case with books that are identical but were published with different titles.

All items are listed in alphabetically order by author. Those with no author are listed alphabetically by title. The style of citation is straightforward and follows the pattern: Author (last name, first name, middle name). TITLE. City and State: Publisher, year. Numbers of pages, size in centimeters (height by length), type of binding and ISBN and LC numbers (if available). Four binding styles are indicated: hardbound, paperback, spiralbound, and pamphlet.

Each item has a description varying in length from two lines to approximately 20 lines. A typical note is around eight to ten lines long and follows a style of annotation resembling this:
First the number of recipes, followed by recipe format or style. The author distinguishes three formats: (1) the Standard Style, which lists all ingredients first, the method following in separate numbered or numbered steps; (2) Action style,

which combines narrated action or method with lists of ingredients; and (3) Narrative Style, which includes the ingredients and their amounts within the narrated paragraphs. Then follows chapter or category titles, if given. These are spelled exactly as given. Next introductory material is detailed and the number of pages given. Any material following the recipes is detailed, the number of pages only included only if extensive. If there are glossaries, number of items and total number are given. Information is included, when applicable, about illustrations and photographs with artist's names given, if provided, for drawings. Indexes, end notes, and authors' biographies complete the annotation. Some books included are bi- or multi-lingual in all or part; for those, other language(s) are indicated as used. In volumes where Chinese is romanized, Cantonese, Mandarin, or Pinyin is indicated. Spelling within each annotation generally follows that used within the volume annotated.

There are also a few items containing few or no recipes, or combining Chinese and other kinds of recipes that have been included because they were considered important in other reference materials, and often are inaccurately listed as Chinese cookbooks, or if they are of some other particular interest. The book also includes an author-, title- and artist- index.

Unlike most of the books it is describing, it is printed on acid-free 250-year-life paper, has very clean typography and a few nice illustrations, but even though there is an index of the artists who have illustrated the books included in the bibliography, there seems to be no mention of those who made the woodcuts and drawings that illustrates the bibliography itself. Some of them seem to be copied from Chinese block prints and popular pictures and must come from many various sources.

This is certainly a very useful bibliography and one can only hope that there will be a similar volume dealing with cookbooks in the Chinese language in the future, including a section on 'western' cuisine naming the sources that formed the Chinese style of 'western' cooking as we can now see what has formed the 'western' under-standing of Chinese cuisine.

I have poured just enough soya on my rice now, and sprinkled my last drop of tabasco...so if you please pass the real Chinese recipes and let me have some genuine cooking at last.

L.F.

ANNOUNCEMENTS

Modern Chinese Studies at Oxford

In the largest benefaction so far given to Oxford University in its current fundraising campaign, the Hong Kong philanthropist Sir Run Run Shaw has agreed to provide £.10 million for the endowment of an institute in his name. The institute will be devoted to the study of 20th century China, for which appropriate library resources will be developed. It will not only provide training in Chinese social sciences as a vital complement to existing courses on Chinese language, history and culture, but will also provide a general understanding of modern China to a wider cross-section of graduates and undergraduates through new optional papers in other degree courses.

Sir Run Run Shaw is president of the Shaw Brothers Organisation. The group owns and operates cinemas throughout Southeast Asia and North America, and has its own film and television production studios in Hong Kong and extensive property interests in Southeast Asia, the United States, Canada and Great Britain. Sir Run Run himself has long supported the performing arts in Asia, and was until recently the chairman of the Hong Kong Arts Festival Society and the Hong Kong Arts Centre. His many educational benefactions include the recent establishment of Shaw College, the fourth college of the Chinese University of Hong Kong, which offers courses mainly in science and technology -.
He was awarded his knighthood for service to the community.

18:vi:90

Ongoing Projects

The Periodical On-line cataloguing project at the Far Eastern Library in Stockholm.

L.F.

The Scandinavian project to include all holdings of Far Eastern periodicals in Scandinavian Libraries and maintain them in a database was initiated in the fall of 1985 at Östasiatiska Biblioteket (Far Eastern Library in Stockholm).

What began as a modest list of titles using a simple database-program on a personal computer of the brand Macintosh has come to include all the holdings and bibliographical information of close to 3.500 Chinese periodicals.

The existing database is mainly the result of the joint efforts of three persons, namely the library assistants Thomas Nilsson, Lin Zhang Fredriksson and myself.

The collections of the Far Eastern Library in Stockholm and those of the Research Policy Institute and the Lund University Institute of East Asian Languages have so far been included.. Work has been started this spring to enter the periodicals of two other collections in Copenhagen and hopefully is to begin in Oslo this fall. At a meeting held in Copenhagen by Far Eastern Librarians from the Nordic countries in January 1987 it was agreed that the format used to register CJK periodical material in Östasiatiska Biblioteket would be used and promoted in Scandinavian Far Eastern Libraries as a whole and that catalogues of the holdings of the different collections should be published as supplements to the present catalogue that form the first two parts: **Chinese Language Periodicals in Scandinavian Libraries, part 1; Stockholm, part 2; Lund** 北欧远东图书馆中文期刊-第一部、斯德哥尔摩、第二部、龙德。

The collection of periodicals in the Far Eastern Library in Stockholm has become known to a large number of researchers outside Scandinavia and we hope that this series of catalogues shall become a useful reference tool for a larger number of those for whom these collections ultimately are kept and maintained. The rich holdings of the Far Eastern Library include many periodicals rarely or never seen outside China and copies of articles are frequently requested especially from the large amount of not so circulated or widely distributed periodicals.

The purpose of the database and how to obtain copies of the catalogues and datafiles.

What started out as a simple checklist at the Östasiatiska Biblioteket to deal with the rapidly increasing number of periodicals to come out of China has grown to include quite a lot of bibliographical detail and much, as we believe, valuable information has been added during the now more than three years we have been working on it. We hope it shall be useful and of help when trying to determine what magazine one is confronted with it's often very hard to trace magazines that have had numerous changes of both name and issuing administrative unit. These notes are continuously updated and will be further amended in later printouts.

Its prime function was to facilitate the hunting down of missing issues that had become more and more complicated. The large number of periodicals reaching our library through many different channels would have been impossible to control with the limited amount of staff that we have at our library and the decision to have an automated catalogue and subscription control has proven to be extremely useful.

In the process of merging the Far Eastern language collections in Stockholm we also for the first time got the incitement to make a thorough inventory of the previously so scattered journals. All the information accumulated on different sets of cards and lists were transferred to a simple database file using a computer and all holdings were checked against the physical volumes for accuracy. Each single volume was actually handled and as a spin-off there was a very large list of duplicates that has been offered to several libraries for exchange. Our hope is that other libraries shall do the same with their duplicate holdings.

Since it has also been a long standing desire by members of EASL to have a union catalogue of

the Far Eastern Libraries complete holdings of periodicals, I thought this was the right way to facilitate such a large project. First, an effort to get the major collections of a nation in a clear and simple format and then at a later point bring all this information together in a larger and more complete database for all the holdings in Europe.

The task of entering all the Swedish holdings is near completion, now only lacking the holdings in the University Libraries of Gothenburg and Uppsala. 1 am also very happy to report that Ostasiatishes Seminar Der Universität Zürich Chinesish Bibliothek can be added to the list of libraries who has successfully used this system to take inventory of their holdings.

We believe that adding local holdings to this file should be vastly simplified due to the large quantity of information already entered and therefore these files are free to anyone in the Far Eastern Libraries of Europe with the hope that they will undertake a similar inventory of their own holdings. Hard copies of the file will be sold at a nominal cost and are available from the Far Eastern Library.

The complete formatted file with Chinese characters can be obtained from our library at the cost of the material it is reproduced on. The files can be obtained in Filemaker™ format on 3.5" floppy disks formatted for the Macintosh or in the ASCII form as raw text either on Macintosh or DOS formatted floppies. If you don't run a Chinese system you lose all the pretty Chinese characters but will nevertheless give you access to all other data. If you specify what field delimiters you prefer for import to your system there are ways of preparing them accordingly. For such a 'raw text only' copy there will be a special charge to cover the handling and file transfer to other systems.

Some records, although they must be very few, may still contain comments in Swedish. My advice to you is simply not to think about them since we never meant them to be there in the final version. The terminology to describe the periodicals have been taken from the Chinese language and are in their transcribed form. The reason for this is to avoid to translate into English some special terms that would, apart from looking awkward, most likely also create confusion. I also beg that you if you wish to use the file and after receiving a copy encounter any errors, you would have the kindness and take the time to let us know about them and not only correct them locally. A file of this proportion is prone to have at least a few mistakes in it even after having been checked over and over.

However, we believe that the holdings should be correct and reliable with an exception for the Daily Newspapers. Most of these have been checked issue by issue as well but it is hard to give this dirty and very time-consuming work the right air of heroism to find enough people willing to undertake it at all libraries and in view of staff limitations I'm not sure myself that it is always possible to defend oneself doing it!

If you have any comments and if you feel that information should be dealt with differently , you can submit these to me at any time. One of my wishes is that we shall be able to find a format that all our libraries can agree upon, so let us discuss it!

Requirements for catalogueing and registering Chinese monographies and periodicals using the system adopted by the Far Eastern Library in Stockholm.

The successful project of catalogueing all Chinese periodicals and the improvement of software to deal with Chinese information, has made us confident enough to automate the catalogueing monographies as well

The setup will run on any Macintosh with at least two and a half megabytes of internal working memory and a harddisk.

The Macintosh must run the Chinese operating system, known as Hanzi-Talk™. There are two versions of this system, one for simplified Chinese (jiantizi 间体字 using the GB- code standard), and one for unsimplified, (fantizi 繁體字 using the Big-5 code standard).

The software to create the databasefiles is called
Filemaker™ and is available from CLARIS and
to facilitate searching and better printing the
program Hypercard™, that comes free with every
Macintosh, has been adopted. With Hypercard it
has also been possible to design an easily explained
interface for the occasional user, so that searching
can be done from a dedicated machine in the
catalogue-room.

For wordprocessing and manipulations of larger
texts as well as generation of high quality format-
ted bibliographies, a program called Nisus™ by
Paragon Concepts Inc. has proven to be the most
qualified.

The databaseprogram is easy to use and flexible
to extend it has been developed by **lingo io** in
close connection with the library staff and can
easily be modified to meet local requirements. A
database can have unlimited 'views' and can easily
be adjusted for various searching and updating
preferences.

Some standard layouts have already been designed
and are available for free on request together with
more than 6.000 records of catalogued Chinese
monographies and more than 3.000 Chinese
periodicals.

Should the provided format not suit the needs of
ones library or if special demands on catalogue-
cards or other output should be requested, although
it is certainly easy to learn how to modify and
create new designs, these formats and printing-
routines as well as custom-made multiscript
systems can be ordered from **lingo io**.*

* **lingo io** *is a company that specializes in Multiscript-
systems, software adaptions and handling of large
databases. One of the companies strong points is the
expertise in Library applications and CJK information
systems.*

*Its main field of activity is to design systems and aid
companies and institutions as well as private end users
to set up and maintain systems capable of handling
multiple scripts especially Chinese, Japanese, Korean
and Roman.*