EAL95 Information Technology for Chinese Studies Experience at the East Asia Library, University of Washington

Yeen-mei Wu
Advances in computer and communications technology have had a revolutionary impact on information handling. Computer databases worldwide can be rapidly searched for information, including full text. The Internet provides instant access to an astoundingly wide range of information.

Nevertheless, some areas in the world have few electronic resources. This may be due to lack of financial resources or technological support, or lack of appreciation of the value of those electronic resources. Furthermore, most of those unfamiliar with the technology often do not know that any of these electronic resources are available to them. The gap in international network coverage and the lack of agreement on information network exchange among some regions of the world have rendered less valuable the idea of a global computer network. However, several projects are now in existence to maximize opportunities for information exchange among various regions. The Pacific Neighborhood Consortium (PNC), initiated by the University of California at Berkeley, is one of the groups taking the initiative to fill the gap and create a standard for information exchange. The Asia and Pacific Special Interest Group (APSIG) of the Australian Library and Information Association is another that conducts the "Survey of databases on Asia" and makes information available to all.

People managing the exchange and productive use of information for different cultures and languages, especially in nonroman scripts, need to overcome some problems. As the PNC has pointed out, the following two issues are of critical importance: (1) the exchange of information about knowledge bases and the ways in which communication with these knowledge bases may be most effectively developed, and (2) the ways in which access protocols may be developed and adapted to permit transparent access to a variety of knowledge bases under conditions of a known and familiar operating system.

Among the several component networks, Internet is the largest; it is still expanding rapidly. According to the Internet Domain Survey in April 1993, 41 percent of the Internet hosts are in the United States while the rest are maintained by various groups in fifty-seven other
countries. The ever increasing number of Internet users, and many newly accessible library catalogs connected to the Internet, give the impression that the whole world is networked.²

Since information is now generated in electronic form, librarians are facing the challenge of implementing the electronic library. Many universities are setting up electronic libraries on the campus network by storing local resources, mounting databases and CD-ROMs, and connecting remote resources via the network. New information technologies applied to an information-rich environment require librarians to learn new ways to access, evaluate, select, and apply information. We must improve our information flow and provide our users with a broader view of worldwide research development.

Today's users as well as librarians have to acquire information literacy. Information literacy education for electronic resources, such as workshops, lectures, demonstrations, and hands-on practice, is an essential part of library learning in order to utilize resources more effectively.³ Two workshops held in 1994 through e-mail subscriptions are worth special mention: (1) Go-Pher-It Workshop, subtitled "Go-Pher-it: your passport to the Internet", and (2) Roadmap for the information superhighway.⁴ Another workshop entitled "Link workshop", a training course of the World Wide Web (WWW) and the HyperText Markup Language (HTML) could also be contacted by e-mail on 16 and 30 October 1995. All three workshops are offered by Thomas P. Copley at http://www.crl.com/~gorgon/

AUTOMATION AT THE UNIVERSITY OF WASHINGTON LIBRARIES

On-line Cataloging: The University of Washington (UW) began Washington Library Network (WLN) on-line cataloging for both monographs and serials around 1977. In 1978 UW became a member of Online Computer Library Center, Inc. (OCLC) and shifted only the monograph portion of cataloging to OCLC; the serial portion of cataloging remained under WLN until 1988. As OCLC technology advanced, UW migrated to the OCLC Prism system in 1991.

Local System: GEAC was implemented as a local system from July 1988 until June 1993, when it was replaced by the Innovative system (Innovative Interfaces, Inc.). Since July 1993 Innovative, as an integrated library system of one virtual file, has added functions beyond that of GEAC. These functions include ease-of-use features in cataloging, acquisitions,


⁴ The first workshop has already repeated its session and a total of 55,000 people have subscribed to the second workshop; these facts are indications of the real need for such networked workshops.

37
circulation, serial controls, and course book reserves. This information is publicly accessible through LCAT (on-line library catalogs) which is part of UWIN. Although UW did not acquire Innovative's distinguished feature of being capable of displaying nonroman scripts in the on-line public access catalog (OPAC), the Computing & Communications (C&C) and the Library are jointly designing such a feature under the UWIN environment. All western-language holdings are on the UWLIE except for some of the items in special collections, government publications, maps, and nonbook materials. UWLIE also displays the on-order status of titles.

**Databases:** In the 1970s UW introduced the DIALOG and ERIC on-line systems, at that time operated by the Reference Division. As early as 1986, UW acquired MEDLINE, Aquatic Sciences & Fisheries Abstracts (ASFA on CD-ROM), and INFOTRAC. Soon after it acquired LEXIS/NEXIS and others. The Twenty-five Dynastic Histories Database was among the first databases added in the early stage of database development (installed in December 1989). Since 1990 there have been rapid increases of locally loaded databases, including MEDLINE (current and backfiles), PSYC, INSPEC, OCLC WorldCat, Expanded Academic Index, National Newspaper Index, and Business Index. All these databases are accessible through a system called UWIN, and most of them are self-guided.

**Library System:** Introduced in 1993, UWIN, the University of Washington Navigator, provides UW's campus-wide information network. This includes access to the UW on-line catalog and other bibliographic databases (UWLIE) which carries a complete list of databases of the UW libraries that can be found through UWIN's "Computer & Electronic Database Information." There are over 350 database files, grouped under names of the libraries and including EAL. WILLOW/UWIN provides a uniform graphical interface to a wide variety of networked information resources through the Internet.

**Information Literacy Education:** The UW Libraries arrange extensive staff training and retooling programs for library staff that deal with the ever changing information technology in the libraries, and furnishes a "Support Notebook" for each library unit with updates made regularly. The Libraries provide numerous guides for library users. Since 1994 the Libraries have collaborated with the Computing & Communications and with Undergraduate Education in a pilot project called Uwired to introduce the UW Freshman Interest Group to new computer technology, electronic communication, and information navigation skills. UW is considered one of the libraries more advanced in automation.

**INFORMATION TECHNOLOGY FOR CHINESE STUDIES AT THE EAST ASIA LIBRARY**

The UW Libraries' conscious development of and emphasis on library automation have laid the foundation for an early start of the EAL's automation program. As a pioneer in the area of automation among East Asian libraries, EAL has performed many distinguished tasks.
EAL's Automation Past and Present

In 1980 and 1981 with a grant from the Ford Foundation EAL cataloged 4,100 serial titles into the WLN database.


In 1986 EAL and the Computing Center of the Academia Sinica signed an agreement to initiate a joint project to develop a prototype of a Chinese full text processing system and to design an integrated library system. Consequently, in 1989 EAL acquired the Twenty-five Dynastic Histories Full Text Database, making the EAL the first institution other than the Academia Sinica to own that database.

In 1987 EAL became one of the first and the largest institution to join the OCLC CJK subsystem. Between 1989 and 1990 OCLC converted 15,000 CJK records for EAL. As a result the number of EAL's titles on-line increased significantly. Although currently only approximately 50 percent of EAL's collection is on-line, we expect the percentage and number of on-line records to continue to increase because EAL now operates in the advanced OCLC CJK Plus version 2 Prism system.

Since UW Libraries launched the GEAC local system in 1988, filing of catalog cards has been eliminated at the EAL's public catalog drawers. Because the on-line CJK records were in romanized form, an OCLC CJK station was provided until 1993 in addition to the GEAC Station for users to search for titles using CJK characters.

EAL's participation in installing new automation services and new automated work procedures intensified when INNOPAC replaced the GEAC system in July 1993 and the UWIN/WILLOW interface and its connection to Internet resources became available. As a result EAL increased its Personal Computer Workstations from three in 1990 to twenty in the fall of 1995. There are also four OCLC CJK workstations, three monitors for the Histories Database, PC monitors for the use of databases in each of the CJK languages, and five X-terminals for public use.

Under a Title II-C grant EAL completed in 1990-91 the OCLC CJK on-line cataloging of 3,000 titles of the prose literature and local gazetteers of the Ch'ing dynasty in its collection.

EAL now has a total of fifty-four Chinese databases, all but one of which were acquired between 1992 and 1995. It is "the only place with a sizable collection of local databases," according to the ARL Foreign Acquisitions Project Report on Chinese Materials, issued in July 1994.
The CCCIICJK software\(^5\) was installed in 1994 on most EAL staff workstations to support the display of CJK characters of the on-line catalog, although the capability to display Korean characters has not been fully developed. It also enables staff to enter CJK characters in the Innovative Interfaces, Inc. Acquisition Subsystem when generating the CJK order slips.

The Computing and Communications and the Non-ASCII Character Set Advisory Group (NACSAG) of the UW Libraries have been working together toward the on-line catalog display of the nonroman scripts through WILLOW which supports graphic displays. The recent C&C successful demonstration of the display-only of the CJK on-line catalog (using the CJK fonts provided by the University of California at Berkeley) has raised high expectations, although there are problems still to be resolved.

**Software Supporting Chinese Work:** Several software programs have been acquired since 1994 to support the work required in this new library technological environment.

- Microsoft Chinese Windows 3.1 (MS CWin)
- Hanin
- Twinbridge Chinese version (TB Chinese)
- Eten 3.1
- UC DOS
- Fox Pro

With the support of MS CWin, the viewing of e-mail correspondence in Chinese, news delivered in Chinese (Big5) from various newsgroups, and Internet resources of Chinese texts becomes possible. Hanin, using the *pinyin* input method, has made easy the composing of Chinese letters and documents. TB Chinese allows quick compilation of bibliographical work involving both English and Chinese. For instance, our *1994 Yearbook of Chinese Studies* was compiled with TB Chinese.

Eten supports all the databases which EAL has acquired from Taiwan and Hong Kong. UC DOS supports most of EAL’s databases produced in China. The database of the Universities Service Centre Collection of the Chinese University of Hong Kong requires Fox Pro software in addition to Eten. We are yet to find a Japanese software program to operate the database *Ming tai teng k’o lu hui pien* which was produced in Japan.

**Resources for Chinese Studies:** "Libraries without walls" has become a reality since the development of UWIN’s campus network in 1993. Library users may access (find, save, 

---

\(^5\) Produced by JOIN Company in Taipei, this new version 2 software adheres to the East Asian Character Code (EACC) and the Chinese Character Code for Information Interchange (CCCII) coding schemes for CJK characters and contains 22,000 characters with four fonts.
send, and upload or download) resources from the library as well as from their offices and homes. And the resources they cite may be those located on the campus or even in other countries.

There are over 400 western-language databases in UW Libraries (mostly bibliographical indexes) and many Internet resources for faculty members and students of Chinese studies. The following are especially useful:

**Cari UnCover Reveal Service:** This service covers more than 13,000 journals and contains records describing articles and their contents. Over 4,000 current citations are added daily. Chinese titles are in Wade-Giles romanization and the article titles are given with English translations. The preliminary count of journal titles in Chinese or in Chinese studies is over 180. The service provides article access and delivery for a fee.

**Center for Research Libraries (CRL):** The Center contains research materials for international studies, including periodicals, newspapers, archival materials, and foreign doctoral dissertations among other kinds of material. This information access is especially welcomed by the UW Libraries which, as a member of the CRL, has the privilege of borrowing materials from its collection for up to one year.

**Foreign Broadcast Information Service Citation Index (FBIS):** FBIS contains citation indexes to a wide range of foreign television and radio news broadcasts and foreign press news. The FBIS index leads the user to the full printed edition of the China series and the East Asia series (including Taiwan and Hong Kong) located in EAL. The *FBIS Daily Report China* at EAL is one of the most frequently used items.

**OCLC WorldCat:** This database consists of over thirty million catalog records and holding information for materials owned by libraries around the world. OCLC WorldCat, a self-guided database, displays CJK titles in romanized form. Those who wish to see CJK characters are advised to use the OCLC CJK database which is identical to OCLC WorldCat but requires a librarian's assistance.

**CHINESE DATABASES**

**Bibliographic Index:** OCLC CJK and RLIN CJK are two of the most comprehensive and well-known on-line databases that provide millions of catalog records of books, serials, maps, and other materials in the library collections. For Chinese periodical articles there are two

---

6 Reported on 9/7/94 by Maureen H. Donovan to H-ASIA, a listserv group.
useful bibliographical indexes in electronic format which have just become available for subscription:

1. Chung-hua min kuo ch’i k’an lun wen so yin kuang tieh (Index to the Periodical Literature in CD-ROM)\(^7\) is produced jointly by the National Central Library and the Industrial Technology Research Institute in Taipei. The latest edition, dated September 1995, covers articles published between January 1977 and December 1994 from more than one thousand periodicals published in Taiwan.

2. Chung wen she k’o pao k’an pien ming she chu k’u (Index to Chinese newspaper and periodical literature on CD-ROM)\(^8\) is produced by Shanghai Municipal Library. Although the producer sent announcements of this important database in the spring of 1995, no one has viewed the demo. Most of the libraries are still contemplating its acquisition.

Some bibliographic indexes are accessible on the Internet, mostly through the Taiwan Academia Network Libraries (TANet) connection. These include T’ai-wan wen hsien tzu liao lien ho mu lu hsi t’ung (Union Catalog of the publications of Taiwan documents), T’ai-wan wen hsien ch’i k’an lun wen so yin hsi t’ung (Periodical index of the T’ai-wan wen hsien quarterly), and Chung-hua po shuo shih lun wen so yin (Index to the ROC Ph.D and Master’s theses). However, access to those bibliographic indexes and other useful resources provided by the institutions listed in TANet requires a password.

**Full Text Retrieval Databases:** Of the fifty-four Chinese-language databases in the EAL, fifty-two are full text retrieval databases and forty-two databases deal with classical titles. However, none are on the network. A complete list of titles can be searched on UWIN:

**Steps for looking up the list of the EAL databases in UWIN:**

1. Telnet uwin.u.washington.edu
2. Select: UWIN
3. Select: Libraries
4. Select: Library Information
5. Select: Computer & Electronic Database Information
6. Select: East Asia Library

---

\(^7\) This title, acquired by EAL in 1994, is an easy-to-use electronic index with comprehensive subject coverage.

\(^8\) The paper edition of this database is entitled *Ch’uan kuo pao k’an so yin.*
The Institute of History and Philology, Academia Sinica, has produced more than forty databases on classical Chinese works. They are networked and available for the use by members of the Institute only. One of the titles, the *Twenty-five Histories Database (HD)* has been acquired and installed by four institutes outside of Taiwan. As of this writing, none of the other titles are available for purchase; most of these titles are not accessible outside of the Institute. Its Computing Center is constantly improving their new technology and upgrades the search methods of databases created by them. For example, they have significantly shortened the search time for the entire full text *HD* from thirty-seven minutes in 1990 to only seven minutes in 1995, currently in its fifth version.

The Institute of Chinese Studies at the Chinese University of Hong Kong (CUHK) compiled and published in 1994-95 two series of the *ICS Ancient Chinese Texts Concordances*, a total of twenty-six titles. Twelve titles of the first series were also available in database format in July 1995. In searching for words and terms in its full text database, users not only can retrieve citations but also a variety of words and terms found in other editions of the same title.

Professor Chen Yu-fu, an independent scholar in Taiwan, has produced more that thirty databases of classical works and Buddhist canonical works. The full text databases do not include annotations or commentaries, only the original texts.

A number of institutes in China have produced valuable full text retrieval databases of classical literature. Few of those databases are available for purchase. However, some databases, such as *Index of Chinese journals*, *People’s Daily*, and *Hung lou meng*, are available for purchase. EAL has acquired *People’s Daily* of 1993 and 1994, but the search functions are less efficient than in other databases.

**Database Manual:** Chinese-language databases often come only with Chinese-language manuals (for both the setup instruction and operational instruction). Commands on the monitor screen are in Chinese as well. The key-in methods of the Chinese characters for different database searches are often different also. Users need more time to learn and need more assistance than when they use western-language databases. EAL compiles easy-to-use operational instructions of each database group (databases by the same producer can be searched by the same search method). However, the National Central Library and the Institute of Chinese Studies, CUHK, provide the most detailed guides in both Chinese and English for the use of their databases. The *Twenty-five Dynastic Histories Database*, while the best know and most often used database on and off the UW campus, has been provided in March 1995 with a revised instruction manual in Chinese which may still be too complicated for some users. An article about EAL’s *HD* and its users’ survey provides detailed descriptions⁹ to those who wish to learn more about it.

---

Survey of Database Bibliography: More Chinese databases have become available dealing with both classical and modern periods since I conducted the preliminary survey\textsuperscript{10} in 1992. It is my desire to update my survey list of the Chinese databases in the near future. More and more Chinese scholars are now interested in Chinese databases, as evidenced by the increasing requests for database demonstrations each year and by the frequent inquiries about "Chinese E-text" in the e-mail program. It is important that China studies librarians, when gathering information on Chinese databases, be familiar with the various hardware and software requirements and the search function of each electronic resource in order to be competent selectors.

INTERNET RESOURCES

Asian Library Resources: UWIN's Internet gopher connection enables users to reach sixteen academic library catalogs in Taiwan and three in Hong Kong. The setup of Mosaic also allows users to be connected to different sites in Taiwan, Hong Kong, and China. EAL uses Microsoft Chinese Windows 3.1 (Big 5 character set) and CCCII CJK (EACC character set) software packages to access Chinese data. However, EAL's experiences with this has been mixed. On TANet we were able to search seven library catalogs out of fifteen; four libraries required passwords and three gave no response. We had difficulty in establishing a connection with another one because we ran out of time at each attempt. These results were compiled after several tests. EAL had no problem connecting with Hong Kong library catalogs. Because of the password restriction, many useful Chinese data and files were not accessible for noninstitute users.

Usenet: At the CEAL's "Workshop on Internet for CJK" in March 1994, participants received a helpful "starter address book" of the Chinese Internet resources. This book lists thirty-nine groups under the "Discussion Groups (ListServs)", and nine groups under the "Usenet-News Newsgroups". In July 1995 Hartmut Bohn of Germany updated the list in his "China/Chinese-related mailing list" with periodic addenda after that. For those who would like to obtain an updated list, it can be gotten at: ftp.lrz.muenchen.de in pub/culture/east-asia/netinfo/MailingLists.C\textsuperscript{11} These lists provide detailed descriptions (such as its purpose, how to subscribe) of each group, so users may select and subscribe to the groups related to their research and interests. UW provides a long list of the newsreader programs available on the campus computing system. Users may use their Pine e-mail system to subscribe to the newsgroups of their choice. The research value and the intellectual level of the contents of these groups in the electronic mail system vary a great deal. We have to make choices of how much time to devote to reviewing these overwhelming number of resources. Among these resources only two groups, "alt.chinese.text.big5" and "CCMAN-L", communicate in Chinese and MS Cwin software at our EAL workstation enables viewers to read the Chinese text.


\textsuperscript{11} E-mail dated 7/20/1995 from Hartmut Bohn to Multiple recipients of list CCNET-L, subject: Update: China/Chinese related Mailing Lists.
**Netscape and Home Page:** The Mosaic and Netscape graphical browsers have been installed on many staff workstations. The international studies librarians are encouraged to produce a Home Page for their areas. As of this writing, EAL is developing a Home Page for EAL's basic information as well as a Home Page for each of the regional studies in EAL. The EAL Home Page will be ready by the **end of 1995.** Our URL is:

http://www.lib.washington.edu/libinfo/libinfo/libunits/east-asia/

Many interesting web sites related to Chinese studies have already been established. Interesting choices include the following:

**Academia Sinica WWW Service Home Page.** http://www.sinica.edu.tw

The "Chinese Full Text Retrieval System" in the Home Page enables UW users* to access the *Twenty-five Dynastic Histories database* and the *Thirteen Classics database.*

*The institutions which have acquired these databases from the Academia Sinica are eligible to use them on the Internet.*

**CITAS (China in Time and Space) Home Page.** http://weber.u.washington.edu/~citas

"CITAS data set includes vectorized base maps of China, georeferenced socioeconomic data, bibliographic resources . . ."

**Yuan Zc Institute Home Page.** http://www.yzit.edu.tw/

Under "Chinese literature" one may find the retrievable full text of *Ch'üan T'ang shih* (Tang Poetry). Poems and biographies may be retrieved by author or keyword.


This page contains (1) links to texts in Chinese philosophy and classical literature, (2) links to bibliographies and e-mail lists relating to Chinese philosophy, and (3) links to other relevant web sites.

**Zhaoming Wenxuan.** http://www.pristine.com.tw/wenxuan/wx01.html

This Zhaoming Wenxuan Project is involved with inputting the text of *Chao-ming wen hsuan.*


The major contents of this web site include (1) links to Chinese-language related FTP sites, (2) sources of Chinese text files, and (3) scholarly resources on Chinese conferences and other matters.
CONCLUSION

The UW Libraries implemented a new library system in 1993 and has since provided a uniform graphical interface to a wide variety of networked information resources. In the process of implementing current technology, we are constantly challenged by changing information technology, abundant electronic resources, and the need for information literacy. Our new networked library system has introduced new ways of communicating and exchanging information, conducting research, and gathering resources. We have come to realize the meaning of "the library without walls" and the disappearance of geographical distance. The electronic resources for Chinese studies are developing rapidly, with new and valuable information constantly becoming available from different corners of the globe. However, both librarians and users need additional technical skills to achieve the transparent access to nonroman scripts through the Internet. The East Asia Library is keeping pace with the development of the virtual library.