Accessing Japanese Data Bases on Personal Computers in the United States

Scott Edward Harrison

Follow this and additional works at: https://scholarsarchive.byu.edu/jeal

BYU ScholarsArchive Citation

This Article is brought to you for free and open access by the All Journals at BYU ScholarsArchive. It has been accepted for inclusion in Journal of East Asian Libraries by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Reference librarians working in Japanese studies are showing more and more interest in gaining access to data bases and other sources of electronic information from Japan. This is not an easy task due to the complexity of telecommunications networks and differences in computer operating systems in the United States and Japan. Until very recently, one needed to have a Japanese-language computer (such as an NEC PC 9801 series), communications software, extension board, and printer, or a Macintosh with a Japanese operating system in order to log onto a data base in Japan. Significant investments in hardware and software, complicated problems involving telecommunications, and the amount of specialized technical expertise required to set up and run such "arcane" systems understandably discouraged all but the most determined and financially able groups (or individuals) from establishing their own transpacific links. Due to new developments in software and telecommunications, the situation has changed considerably and East Asian libraries can now consider relatively easy access to some forms of Japanese electronic information without spending large sums on specialized hardware and engaging the services of a Japanese computer expert.

In this article, I will discuss what is entailed in accessing Japanese online data bases and CD-ROM products from Japan in this country. My purpose is to provide information as to how a library might go about establishing access to these data bases and about what to expect in the way of hardware/software requirements, costs, setup procedures, and possible difficulties. This is not a data base review article, although I will speculate on the utility of some products to librarians in Japanese studies collections.

Before discussing data base access, however, it might be beneficial to provide some brief background information on the nature of Japanese personal computers for those who may not be fully aware of the important differences that exist between the computers of Japan and those of the United States. First of all, it is necessary to define two terms. A "Japanese" personal computer refers to Japanese-language computers made in Japan, running under a Japanese operating system, using a Japanese keyboard and display, and designed for domestic use. It does not refer to Japanese-made computers for western languages sold in the United States by NEC, Epson, Toshiba, and others. In broad terms, an "American" personal computer refers to IBM PCs and compatibles running under the English-language version of MS-DOS.

Many people ask if Japanese-language personal computers can run American software written for IBM PCs and compatibles. In most cases the answer is no. Basically, this is because personal computers intended for western languages are designed to process letters and symbols which occupy one byte of space in comparison to Japanese computers which are designed to process graphic characters which occupy two bytes of space each. The two systems are not compatible because of differences in display requirements, coding of information, keyboard structure, and many other complex factors. In addition, Japanese computers have proprietary operating systems which are different from those in the United States. Consequently,
with the few exceptions noted below, an ordinary Japanese PC cannot run American software or operate in conjunction with programs designed for the IBM PC family of computers, nor can an American PC run programs written for a Japanese computer.

The first exception to this is a class of Japanese personal computers known as architecture extended (AX) machines. These computers, initially made by companies such as Matsushita, Sanyo, Sharp, and Mitsubishi, were designed to run American software for the IBM AT as well as Japanese programs which were written for or enhanced to run under the AX operating system. Unfortunately, individual design differences in each company's line of machines resulted in partial incompatibility with each other in the Japanese mode and difficulties running English-language MS-DOS software.\(^1\) Newer AX machines have incorporated many improvements and the level of incompatibility has been reduced. However, users still experience problems running some versions of Japanese and American software.

The other exception is a recently-announced machine from AST Research Japan which claims to run both MS-DOS applications and NEC PC 9801 series software.\(^2\) It is called the AST DualStation 386SX/16 and features a 110 MB hard drive and 2 MB of RAM. The computer automatically selects the correct operating system depending on the application being used. Consequently, it is possible to run English IBM software and Japanese NEC 9801 series software on one machine, but not at the same time. It cost approximately $4,500 in Japan.

The AX computers and the AST product are not sold in the United States and will not be supported or serviced here by their respective companies even if brought from Japan by individuals. This, combined with high price tags, make them a risky investment for any library.

There are two additional AX "solutions" to the problem of making an American personal computer function as a Japanese PC. Both are available in the United States from Pacific Software Publishing in Seattle, Washington. Readers are invited to contact Mr. Richard Stratton at (206) 232-3989 for more information.

The first solution is a hardware device called the Prosise AX Kit. It consists of a board that plugs into a PC equipped with either a 286 or 386 processor and a multi-Sync monitor. In conjunction with a Japanese keyboard, it emulates the AX operating system. This board also allows the transposition of Japanese text over graphics. The board costs $700 and an AX-compatible printer runs between $700 and $1,500.

The other solution is a piece of software known as the AX VGA/S Development System scheduled to be released later this year. It is supplied on a floppy disk and apparently will enable any PC equipped with a 386 processor and VGA display to function as a Japanese AX computer. No special keyboard is required. It is anticipated that this software will function with Epson dot-matrix and HP Laserjet printers; but there are no guarantees being put forward. The price of this software is expected to be between $100 and $200.

---

\(^1\)Richard F. May, "It Must be Getting Better, Because It's Been So Bad," *PC Report* (May 1990): p. 43.

It should be remembered that an American PC operating under the AX VGA/S system can run only Japanese software that has been written for, or specifically enhanced for, the AX system. This means that an ordinary copy of Ichitaro (a popular word processing program) for NEC computers will not work on a PC running AX VGA/S. On the other hand, an AX version of Ichitaro or an NEC/AX version will work. This is an important distinction that should be kept in mind.

It is for the most part true that there is presently no one standard for Japanese personal computers. With very few exceptions, there is nothing like the IBM compatibles in Japan. NEC machines run on NEC's own proprietary operating system. Fujitsu personal computers run on a Fujitsu system. They are not compatible with each other or with computers made by Sharp, Toshiba, or Hitachi. Consequently, a software package intended for NEC personal computers will not run on a Sony or Fujitsu computer. As mentioned above, the AX machines are the first attempt to change this. The current market situation in Japan resembles the personal computer market in the United States in the early 1980s when every computer maker had his own version of the CPM operating system and there was very little "compatibility."

In comparison, personal computers in Taiwan are based on the American IBM operating system and have the same hardware requirements. Both Chinese and English software packages from Taiwan can be run on American IBM and compatible personal computers, although not without problems.

Fortunately, it is no longer necessary to have a "Japanese" computer to gain online access to Japan's data bases. It has been possible for quite some time for Macintosh computers running under the Japanese operating system called KanjiTalk to connect to computers in Japan. Now, thanks to a text editor and telecommunications package called KanjiComm, IBM PCs and compatibles can also connect to data bases in Japan. Both types of computers can execute searches, display the results in Japanese, download information, and print it out in Japanese on printers made for the U.S. market. Transpacific electronic mail in "real" Japanese is also possible for IBM and MAC users in the United States.

Basic hardware and software requirements for an online data base connection to Japan are as follows (the presence of a modem is assumed). The requirements for IBM PCs and compatibles are: an AT-class machine or higher with hard disk, EGA, VGA, or Super VGA, and the KanjiComm software package. Macintosh computers need a hard disk and at least 2 MB of RAM, the Japanese operating system KanjiTalk, and EGTalk, a telecommunication program. Libraries fortunate enough to have an NEC 9801 series PC, IBM PS55, or similar Japanese computer can also connect to most Japanese data base services provided they have the correct modem and appropriate Japanese telecommunications software.

At the present time KanjiComm costs $200.00. The menus are in Japanese, as is the main manual. A summary manual is available in English. MacKanji (which includes KanjiTalk and a text editor) retails for $99.95. The Mac LC and Ilsi require

3 Japan is a Mirror Image of the U.S.," P. C. Letter (June 18, 1989).

4 Readers interested in this topic are referred to Ken R. Lunde's excellent article, "Electronic Transfer of Japanese," in the September/October 1990 issue of ATArashii. KanjiComm also supports an electronic mail function.
a later version of KanjiTalk which lists at $400.00. EGTalk, which is usable on any Macintosh, is $299.00.

IBM and compatible computers can print search results in Japanese on the HP Laserjet II or higher, HP Deskjet series, and some dot matrix printers such as the more popular models in the IBM and Epson lines. Macintosh users can print Japanese with the Apple LaserWriter, HP Laserjet III (with Apple Talk), and the Apple Imagewriter. Japanese computers need Japanese printers, cables, and possibly some type of interface card.

One of the easiest and most inexpensive Japanese data base connections available to libraries in this country is the G-Search service (formerly known as Hi-net). This service is offered through Japanese Language Services (JLS) located in Boston. This company also sells the software packages mentioned above and can be of assistance in gaining telephone access to TYMPAS (a packet switching network affiliated with TYMNET) which enables one to connect to G-Search in Japan by placing a local call in the United States. Interested parties may contact:

Mr. Takeshi Tokushige, Technical Support Manager
Japanese Language Services
186 Lincoln Street
Boston, MA 02111
Telephone: (617) 338-2211
Fax: (617) 338-4611

There is also the option of dialing direct to Japan using AT&T or other companies such as MCI and Sprint. In this case, the subscriber would pay for the international call and receive an invoice from Japanese Language Services for data base use. At the present time, it is possible to arrange a TYMPAS connection through JLS, use a JLS account to logon to TYMPAS, access G-Search with one's own account and password and receive one statement from JLS for telecommunications charges and data base usage. The subscriber pays for the local telephone charges.

The initial cost for G-Search is $200.00 which includes establishing an account, issuance of a password, manual, and one hour of support in the Boston office of JLS or by phone with the customer paying for the calls. There is a monthly minimum of $75.00. A charge of $1.00-3.00 is incurred each time a data base is accessed. Users may anticipate TYMPAS charges at 25 Yen per minute for connect time and 50 Yen per K-byte of information displayed. Setup charges for a TYMPAS account are approximately $75.00.

Access to G-Search on an IBM PC or compatible is accomplished through KanjiComm which dials the number to TYMPAS and establishes contact in much the same way that Procomm does in an English-language environment. KanjiComm also makes it possible for the PC to receive and display Japanese. G-Search is available twenty-three hours a day and offers access to some forty Japanese data bases of which the following are currently available to subscribers in the United States as the screen menu below indicates:
Other data bases, such as WHO - jinbutsu jinzai joho from Nichigai Associates, are also available for a separate charge.

In order to enter a data base, one simply types in the appropriate code found at the lefthand side of each data base listed. In the Yomiuri Shinbun kiji detabesu, for example, one can search any word in any article using the "free term" index. Keywords (assigned thesaurus terms) can be used to produce a narrower search. Another approach involves selecting a subject portion of the data base, such as economics, and then searching the subset with "free terms" or keywords. It is also possible to qualify searches by date and certain other restrictors, although I understand that this can be time consuming. All search requests are typed in romaji and converted to kana by the system. It is not possible to construct a search with Chinese characters and one must be aware of the potential problems caused by homonyms and formulate search strategies accordingly.

I brought a few actual reference inquiries to JLS to see how the G-Search Service would respond. I will summarize our experience with three out of six search results.

The first question we dealt with involved a request for information regarding public reaction to recent changes in the abortion laws in Japan. We were not able to find anything substantial related to legal matters and abortion. We did, however, discover that many articles retrieved were about abortion in countries other than Japan, as might be expected for a newspaper that covers international events. The importance of a well-constructed search soon became apparent.

The second request was for current articles by or about Kuroiwa Jugo. We searched the Yomiuri data base for the free term "Kuroiwa Jugo" and within less than a minute received sixteen citations ranging in time from November 1986 to July 1991. Figure 2 below shows what the search request looks like on screen:
Citations consist of headline, date, edition, page number, section of the paper in which the article appears, presence of a photograph, and the total number of characters in the article. Citations for articles appearing in 1991 are given below:

- 000014  (19910121TYM10002)
  [歴史小説]「白鳥の王子ヤマトタケル＜大和の巻＞」黒岩重吾著（寄稿）
  91.01.21 東京読売朝刊10頁 書評面 写有（全550字）☆

- 000015  (19910415TYE09005)
  吉川英治賞の贈賞式 文学賞の平岩弓枝さん「老父が感激」
  91.04.15 東京読売夕刊9頁 文化面（全665字）

- 000016  (19910729TYM11003)
  [執筆時間]作家・黒岩重吾さん 望遠鏡の先はいにしえの国
  91.07.29 東京読売朝刊11頁 国際経済面 写有（全692字）

The third request came from a graduate student investigating the antismoking movement in Japan. A search on the free term "ken'en" (meaning antismoking) produced seventeen citations. The addition of "tabako" as a free term produced 1,651 citations. Combining the two results produced fifteen citations ranging from January 1987 to July 1991. Four of the citations did not deal directly with Japan. Although one would need experience in the use of free term searches, combinations, and qualifiers to assure maximum coverage, the preliminary result we obtained provided the researcher with access to articles concerning the creation of antismoking posters, a proposal (demand) to prohibit workers from smoking in the Setagaya Ward Office, a decision to place eleven female high school students on probation for smoking, and finally, a call for participation in a smokeless Goh tournament.

The cost in yen for the session which included the last two searches mentioned above and some others not reported here is as follows:

<table>
<thead>
<tr>
<th>Service</th>
<th>Cost (yen)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to the Yomiuri database</td>
<td>¥300</td>
<td></td>
</tr>
<tr>
<td>Headlines displayed</td>
<td>¥1,040</td>
<td>(52 @ 20 yen each)</td>
</tr>
<tr>
<td>Full text requested</td>
<td>¥600</td>
<td>(4 @ 150 yen each)</td>
</tr>
<tr>
<td>Total</td>
<td>¥1,940</td>
<td>(US$14.07)</td>
</tr>
</tbody>
</table>
The table above does not include telecommunications charges which add approximately $10.00 to the cost of the searches. This represents about an hour of connect time and, it should be emphasized, includes time taken by my questions and requests for variant search strategies on the same question.

Plans are in progress to add thirty more Japanese data bases to G-Search by the spring of 1992. It is hoped that easy access to the entire G-Search service without supplemental charges will be available in the United States sometime late in 1992.

Another major data base service available in this country is the Nihon Keizai Shinbun's Nikkei Databank, also known collectively as Nikkei Telecom. There is an English-language data base and a Japanese-language data base which differ in content and in their fees. In order to connect to the Japanese data base, one may use an IBM PC or MAC in conjunction with access software supplied by Nihon Keizai Shinbun. Naturally, a Japanese computer may also be used. The Japanese data base is divided into four components. Of these, the Sōgōban and the Nikkei News Telecom are of most interest to East Asian libraries. The former is concerned with financial and market news, while the latter is a general news and reference data base with a business emphasis. Access to these data bases can be accomplished through Nikkei's NEEDS-NET communication nodes in New York, Los Angeles, and Washington, D.C. or through TYMNET or ITT.

The initial setup fee for any of the four data bases is $280.00. Subscribers may then choose from three different fee plans. The flat rate for the Nikkei News Telecom data base is $1,000 per month and allows for unlimited access. The "Type 1A base rate" consists of a monthly charge of $400.00 plus $0.50 per minute of connect time. The "Type 1B base rate" requires a monthly payment of $100.00 and connect charges of $1.00 per minute. In addition, there are various charges that apply to the use of certain data bases for subscribers in the "Type 1B" category. I understand that it is not possible to download information to a disk. Consequently, search results need to be read on the screen or printed online. For more information, readers may contact:

Mr. Mark B. Duff  
Sales Department  
Nihon Keizai Shinbun America, Inc.  
1221 Avenue of the Americas  
Suite 1802  
New York, NY 10020  
Telephone: (212) 512-3600

CD-ROM data bases have many attractive features but present some different access-related problems than do online data bases. CD-ROM applications require software which tells the computer how to search the disk, to format information, and to display it on the screen. In the case of Japanese CD-ROM products, the software is written for Japanese computers and cannot be run on an American machine. In addition, a Japanese CD-ROM drive using the format required by the CD-ROM application is necessary. This drive must be capable of being connected to the PC and one needs to have the correct cable(s) and interface board. Most Japanese ROM drives also require CD-ROM extensions software. Many factors have to come together for a CD-ROM application to run properly; the PC itself must be able to run the application software and the CD-ROM drive, display, and interface board must all work together. There is absolutely no guarantee that a hardware/software configuration put together to run one specific CD-ROM application
will run other CD-ROM products off the shelf without any problems. However, if one has a "major" hardware configuration, such as an NEC 9801 V series PC and an NEC PC-CD 103 ROM drive with proper CD-ROM extensions and interface board, there should not be any serious difficulty running most CD-ROM packages as long as the correct version of the search software is selected.

Such equipment could be purchased from dealers in Japan, but there would be no service or support in this country and a stepdown transformer would be needed for equipment designed to run only on 100 volts. The import duty on personal computers brought into America from Japan is 3.9 percent.

In theory it could be possible for an American PC running AX VGA/S software to support a Japanese CD-ROM drive. However, this has not yet been tried and it remains to be seen if the two components can be connected to work together properly. The availability of CD-ROM packages for the AX system will also be a primary concern. There are also CD-ROM packages available for the Macintosh in Japan; however, it appears that availability of data bases such as J-BISC, Asahi Shinbun data base, and various periodical and statistical indexes is quite limited. At this point in time, Japanese equipment seems to be the only sure option for accessing the widest possible range of Japanese CD-ROM products.

To repeat what I said above, libraries in this country may turn to Japanese Language Services for an NEC configuration similar to the one described above. It is not an inexpensive proposition and prices may vary depending on exchange rates and availability. The following figures (F.O.B. Boston) provide a good indication of what to expect:

<table>
<thead>
<tr>
<th>Item</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEC PC (640 K, 2 1MB disk drives)</td>
<td>$2,700</td>
</tr>
<tr>
<td>40 MB hard disk</td>
<td>1,400</td>
</tr>
<tr>
<td>Color display (includes cable)</td>
<td>1,098</td>
</tr>
<tr>
<td>Printer (80 column)</td>
<td>1,200</td>
</tr>
<tr>
<td>CD-ROM drive (with interface board)</td>
<td>2,300</td>
</tr>
<tr>
<td>CD-ROM extensions</td>
<td>200</td>
</tr>
<tr>
<td>Total</td>
<td>$8,598</td>
</tr>
</tbody>
</table>

In addition, the CD-ROM drive will, for the immediate future, require a stepdown transformer to change 110-120 volts to 100 volts. All of the above equipment, except for the CD-ROM drive, can be serviced in the United States. JLS will arrange service for the CD-ROM drive which would need to be done in Japan. Technical support for the entire configuration is available by phone or in Boston from JLS.

CD-ROM drives purchased in Japan for $500-$1,000 (plus consumption tax if applicable, shipping and handling, insurance, and import duty of about 3.7 percent) can be attached to newer versions of Japanese computers (three or four years old?) which a library may already have. Needless to say, specifications for cables, interface boards, and software should be carefully checked. Persons planning to add a CD-ROM drive to older NEC, IBM, or Fujitsu equipment would be well advised to consult with the manufacturer in Japan and carefully verify matters of compatibility and equipment availability before making a purchase. Whatever approach is taken, it is very important to make sure that the CD-ROM applications one wishes to use are available in a version designed for the computer on which they will be run.

The main advantage of a CD-ROM application over online services is that there is no charge for connect time and no need for telecommunications equipment. While
it is difficult to predict the amount of time and, therefore, of money that use of an online data base would cost, a CD-ROM data base represents a fixed expenditure. However, CD-ROM data bases can be expensive and the costs of buying different packages can quickly mount up. An individual library would have to weigh the advantage of somewhat limited access to the wide range of data bases available through an online service to the unlimited use of a few carefully selected CD-ROM products.

How do costs compare for access to the print and electronic versions of the *Asahi Shinbun*? The following figures give the economic side of the picture. Amounts quoted are in US dollars for a yearly "subscription":

<table>
<thead>
<tr>
<th>Service</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard copy (New York satellite edition)</td>
<td>$940-1,000</td>
</tr>
<tr>
<td>Reduced-size edition <em>(shukusatsuban)</em></td>
<td>$490-500</td>
</tr>
<tr>
<td>Microfilm</td>
<td>$2,200</td>
</tr>
<tr>
<td>CD-ROM (1989 and after)</td>
<td>$876</td>
</tr>
<tr>
<td>Online</td>
<td>varies</td>
</tr>
</tbody>
</table>

Based on the roughest of estimates, if a library used G-Search solely to search the *Asahi Shinbun*, one might receive eight searches each averaging twenty-five headlines and one full text article for the $75.00 monthly fee. TYMPAS charges might come to $35.00 for two hours of connect time. This would amount to $110.00 per month, or $1,230.00 per year. Of course, the online service may not be used exclusively for searching the *Asahi*, in which case the "cost" for online access would decrease.

The next question that appears to be asked fairly often is whether the hard-copy newspaper subscriptions could be eliminated in favor of the CD-ROM or online versions. In the electronic versions, there is no way to browse or "read the newspaper." A full page of the paper cannot be called up on-screen and viewed. Moreover, advertisements and other ephemera cannot be seen and researchers lose the ability to check for related but unknown events by scanning a set of issues for a given period of time. With the CD-ROM version one loses timeliness. The online version does not include the full text of articles copyrighted by individuals. For these reasons alone, it would seem unlikely that the electronic versions could be used in place of a subscription to the hard copy—at least for the purposes that most East Asian collections subscribe to daily papers. If a library could rely, for example, on CRL’s microfilm set of the *Asahi* for historical research and browsing, the reduced-size edition or the microfilm could be eliminated for the years after 1984. For some libraries, online access might even be a viable alternative to hard copy subscriptions to all but one major daily newspaper.

In the final analysis, newspaper data bases are more useful as an extremely powerful access tool rather than as a cost containment device. If the data base use patterns experienced by general reference librarians hold true for Japan, the electronic newspaper and magazine indexes alone will prove invaluable additions to the Japanese reference collection.

The online data base services also provide access to information that is difficult and expensive to obtain in hard copy. Some data bases provide information available only in electronic form, the acquisition of which may be viewed as analogous to new books and serials.
The number and scope of Japanese data bases is growing rapidly and will continue to improve as personal computers gain popularity in the home and workplace. We now have the ability to access a portion of this valuable pool of information and make it available to our clientele.