Minicomputers in Libraries 1979-80

Audrey N. Grosch
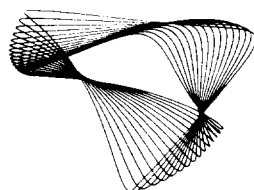
Knowledge Industry Publications, Inc.

Minicomputers in Libraries, 1979-80

by Audrey N. Grosch
University of Minnesota

Professional Librarian Series

Minicomputers in Libraries, 1979-80
by Audrey N. Grosch

Contents

Preface

The purpose of this book is largely instructional or tutorial. It emphasizes hardware and software that are generally available throughout the United States and Canada and which form the basis of any potential or existing library minicomputer installation.

Librarians and information center managers should have a grasp of some of the technical matters relating to minicomputer systems including the design philosophies upon which systems may incorporate this class of computer system.

Students of library and information science, specifically those who are not also concentrating on becoming systems specialists, should also have a basic source of information on this topic to use during their exposure to courses in library automation and systems design.

The author presumes the reader has a basic familiarity with beginning data processing concepts, but is essentially inexperienced within the systems area and requires some learning tools for self-instruction or as an adjunct to beginning level courses in library systems aspects. Consequently a glossary of terms follows the text which attempts to use language as non-technical as this topic permits.

This volume is arranged in two distinct sections. Chapters I through V are primarily non-technical descriptions of available systems applications and procurement procedures. For those who wish to know somewhat more about the technical specifications of computer systems, chapters VI through IX provide additional background. This latter material may also be useful in understanding vendors or in-house systems analysts as well as arming non-technical library system decision makers with the questions which might be asked of the technical people.

Audrey N. Grosch

St. Paul, Minn.

I

Overview — The Use of Minicomputers in Library Applications

In the late 1960s minicomputers first attracted the eye of some information systems practitioners concentrating on developing library automation systems. As far as can be determined, simultaneous interest was taking place in the U.S., Great Britain, Canada and Australia as libraries moved into the 1970s. Some of the pioneer institutions that employed minicomputers were the University of Minnesota, Oxford's Bodleian Library, the University of British Columbia and the University of Sydney.

However, the early use of minicomputers by libraries was constrained by both hardware and software limitations. Although some hardware limitations were removed by 1972, when Digital Equipment Corp. introduced its PDP 11/40 model minicomputer, software was still a major problem. For example, there still existed no operating systems to allow online real-time use.

Libraries developing online applications in the conventional large computer environment have very sophisticated aids at their disposal which make it possible to move immediately into the design, coding, debugging, testing and installation of a system. Operating systems with file management and peripheral equipment servicing routines exist for these systems in a tested and serviceable form. A variety of higher level programming languages are available that can speed program creation. Software which assists a programmer in testing and debugging programs is also available. But in the minicomputer development era of the early 1970s, software was well behind hardware development. Software of the type and quality found in conventional computers was either nonexistent or too primitive to permit actual development of the interrelated applications found in many library data management activities — activities which demand fast response and quick handling of large numbers of transactions. For example, a circulation system must interact with a cataloging system to obtain bibliographic data, but it must also give rapid response to a charge even though dozens of charges may occur each minute.

Today, on the threshold of the 1980s, this picture has changed. In the minicomputer field, and the very rapidly developing microcomputer area, both hardware and software

have matured to the point where library applications can be developed practically and economically using readily available hardware and system software products from a variety of manufacturers in the U.S., Europe, Japan and Great Britian. In some cases these foreign products equal or surpass U.S. manufactured equipment, e.g., Facom and Okitac in Japan and Siemens in West Germany.

To gain some idea of the use of small scale computers in libraries, i.e., micro, mini and midicomputers, a questionnaire was sent to subscribers of *Advanced Technology/ Libraries* newsletter as well as libraries known to have interest in such systems. Questionnaire returns were predominantly from domestic libraries, with scattered foreign responses. The results are to be found in the Directory of Installed Systems in the back of this volume. Certainly, this Directory cannot be considered to be complete, particularly as new systems are being installed almost daily. However, the questionnaires do provide the basis for discerning some trends in the application of minicomputers. This initial chapter highlights the findings from the questionnaire.

Few librarians who answered the questionnaire on behalf of their libraries sent any system documentation. Therefore, the questionnaire responses produced general information about the system's hardware and software environment and the applications performed. It may be helpful to point out that the questionnaire was structured to determine: how the system was developed, i.e., whether it was a commercial, turn-key or in-house system; whether the system operates independently as a stand-alone system or as a front-end computer sharing processing tasks with another or host computers; whether system documentation existed; and whether the library would be willing to consider arrangements with another library interested in its software for replication (unless the software was a commercial proprietary product).

The remaining sections of this chapter address commercially developed systems, both North American and foreign, and then examine some significant work in non-commercial developments. Omission of a system does not imply that the system was of lesser significance but merely that insufficient information was available. Some systems not found in the Directory may not even have been reported at all.

TRENDS IN COMMERCIALLY DEVELOPED SYSTEMS

United States and Canadian Firms

Initial minicomputer installations, for the most part, have been concentrated in the area of circulation control. Particular emphasis by commercial system vendors is very apparent here. These basic circulation systems capture data online using optical character recognition (OCR) or bar codes scanned by a hand-held reading device called an optical wand reader. File updating is also performed online for these transactions. Patron or staff inquiry of these systems is usually through an author/title search key or a numeric identity number if the book is in hand. This inquiry is performed by using a visual display terminal with an attached keyboard.

Major commercial vendors of these systems are CL Systems, Inc., DataPhase Systems,

Inc., Systems Control, Universal Library Systems, Ltd., Geac Canada Limited, and Gaylord Bros. Although the Gaylord system uses a minicomputer, it is actually a front-end processor to a host computer at Gaylord which is shared by Gaylord system users. The other systems use a variety of minicomputers alone or in network configurations for very large multi-library systems. An example of this latter type of system is the North Suburban Libraries system near Chicago. In this case, CL Systems, Inc. LIBS 100 installations in individual libraries are connected to and communicating with LIBS 100 systems at other libraries.

It is not the objective here to recommend any single system. Such a choice depends upon many factors which each library itself must consider. However, certain trends in the field of commercially available systems are apparent. Among them are:

● Competition is increasing, with the expectation that in the short term at least, even more firms will enter the field, particularly with new applications for libraries of different types.

● Present firms will build stronger customer bases, permitting them to gain experience in many different operating environments, and making their systems constantly more efficient in serving their users' needs.

● The latest commercially available technology is being used in these systems, but at the same time they are being built to facilitate later expansion for other functions.

● Program reliability and customer response, particularly with the larger companies, are improving.

● New functions, such as cataloging data entry and direct online user catalog searching, have been or are being added by most of the commercial vendors, thus easing the task of keeping the circulation/catalog data base current with library holdings.

● Connections with other bibliographic systems, although still primarily accomplished by processing magnetic tape data from another system, are moving in the direction of direct, online computer-to-computer interaction.

● As the power of minicomputers rapidly increases, the available mass storage and main memory capacities have grown to the point that many libraries, funds permitting, can now have fully online catalogs and in-process files.

● Foreign firms, particularly Canadian, are now attempting to enter the U.S. market. If their efforts are successful and their systems can be supported in the field to compete with U.S. companies, foreign suppliers may be expected to offer some exciting new technology. On this point alone, investigating foreign systems is important for the future.

As an example of the competitive situation, five years ago CL Systems, Inc. then called CLSI, was the only marketer of a workable and successful minicomputer-based stand-alone circulation control system, which in several installations also performed film

booking and acquisitions. But since 1976, DataPhase Systems, Inc. has provided systems to Dallas County (Texas) Community College District, Wichita (Kansas) Public Library, Tacoma (Washington) Public Library, Houston (Texas) Public Library, and Oral Roberts University, Tulsa, Oklahoma, among others. Meanwhile, CL Systems, Inc. grew from its initial system installation at the Cleveland (Ohio) Public Library to over 100 installed systems at the end of 1977 and over 200 by the end of 1978.

Each of these companies is introducing devices representing the latest technology into its systems. For example, DataPhase was the first to offer a choice between OCR and bar-encoded labels for book and borrower identification numbers. In late 1977, CL Systems, Inc. introduced a touch screen visual terminal as a device for the public to make inquiries. This terminal, manufactured by Information Dialogs, Inc., one of several terminals currently offered with touch screen capability, allows users to make their choice of displayed alternatives for input by merely touching the appropriate spot on the screen. This affords a whole new approach to access to a computer system for a general public having varying social, psychological or intellectual acceptance of computerized systems.

Another matter of importance in these commercial systems is their move to more flexible record formats. Initially, CL Systems, Inc.'s LIBS 100 used a short, fixed length bibliographic description for each title. Anticipated for introduction in early 1979 is a full Library of Congress MARC record capability which will be offered as a part of their applications software. Obviously, this will enhance their system's functional capabilities, particularly in academic or research libraries which require in-house systems capable of storing and accessing full cataloging records.

One of the latest commercial entries into the U.S. library market is a system offered by Informatics, Inc.'s Library Information Services Group to support L.C. MARC cataloging. It is aptly named Mini-MARC. Full L.C. MARC cataloging records are stored in the floppy diskette attached to a Computer Automation LSI-220 processor packaged into a desk-sized work station. A Hazeltine Co. visual terminal is used to search the MARC catalog record file by main entry, title and L.C. card number. Retrieved records can be modified in any way desired by the user and, with the addition of a special program package, catalog card sets may be printed out on an optional printer. Also, magnetic tape may be substituted for floppy diskettes output. The L.C. MARC data base resides intact on floppy diskettes much like a set of phonograph records in their dust jackets and is updated by weekly mailings of MARC on this medium. The appropriate diskette is chosen by scanning an author/title index which gives the correct diskette for a range of L.C. card numbers.

Moreover, Mini-MARC has as another option full telecommunication capabilities with linkages to systems such as CL Systems' LIBS 100 circulation system, Baker & Taylor's BATAB acquisition systems and soon OCLC's cataloging network. Informatics intends to offer future application packages that will permit this presently single function system to grow as its users may require.

Other Firms

Outside of the United States and Canada several innovative commercial systems are being developed and installed in libraries or information centers.

The CAIRS (Computer Assisted Information Retrieval Systems) retrieval package, offered in Great Britain, was developed for the Leatherhead Food Research Association by Libra Information Systems, Ltd. This turn-key system includes a choice of hardware appropriate to the size of the installation and all software necessary to begin actual use of the system by a library. One installation outside of Leatherhead has been made in the corporate library of The Streetley Co. in Nottinghamshire. CAIRS offers the following functions in each of its five versions: data entry, inverted file subject searching, author searching, accession number location, document or file searching, two library acquisition list outputs and complete file listings, key word index listings, SDI processing, serial character string searching and full record editing/updating.

Searching in this system employs Boolean logic and can be qualified by date. The system also provides for an online thesaurus of terms and implementation of some automatic indexing of documents.

The smallest version of CAIRS is 1A which utilizes floppy diskette storage and a microprocessor in a single, one location installation. This system stores an index to over 2000 items. The largest version is CAIRS 5C which will contain 700,000 bibliographic records and uses a large minicomputer and removable media disk drives in a multi-terminal application.

This system appears to be well designed and affordable for automation of bibliographic retrieval functions for special libraries having patents, reports, reprints, in-house documents or records to access. No such comparable commercially available system is being offered by a U.S. vendor, although the CAIRS system employs Texas Instruments micro- and minicomputers in its respective versions.

Another British system, ADLIB, is designed with the special library and information center in mind. It is being offered both as a turn-key hardware/software product or as a time-sharing service by Lipman Management Resources, Ltd. Thus, very small to large diversified libraries can utilize this system in whichever of these manners would be most cost effective for their file sizes and transaction volumes. ADLIB comes from: an ADaptive management system for special Libraries and Information Bureaux.

ADLIB can store full or partial bibliographic records for monograph, report, preprint, serial or audio-visual materials. Searching is by Boolean operators, i.e., and, or, not, linking keywords or by interrogation of simple rotated string subject indexes. Selective dissemination of information (SDI) and retrospective searching are both accommodated in this system. Online thesaurus maintenance and the ability to accept British MARC format bibliographic data are features of the system that lend appreciably to its utility for a variety of libraries. Circulation control is also handled by the system via data entry from visual terminals. Output can be generated on the line printer or on

magnetic tape for COM processing for micrographic catalogs. The hardware system supporting ADLIB is a PRIME 300 central processor with a minimum of 128K (128,000) bytes of memory, with peripheral equipment chosen appropriately to the size of the system.

A final commercial system which was discovered through the questionnaire was the VUBIS system, a joint development of the Free University of Brussels and Interactive Systems, N.V.S.A., also in Brussels. This system employs Digital Equipment Corp. PDP 11 minicomputers and can be arranged to support up to 1.4 million titles. The software supporting this system is the MUMPS (Massachusetts General Hospital Utility Multi-Programming System) operating system, data management system and application programming language. The language version is that supported by Digital Equipment Corp. (The DataPhase Systems, Inc. system previously discussed also employs MUMPS but with a version supplied by Medical Information Technology, Inc.)

VUBIS is an interactive cataloging input and public catalog inquiry system which is being planned to also include a circulation control subsystem. Acquisitions and serials check-in are expected to be implemented in 1979. This system has several features which make it quite unique and responsive to the environment of its first installation.

A prime feature is multiple language interaction. At the Free University a user may inquire in French, Flemish or English. In a university which has a French- and Flemish-speaking clientele separated into essentially two user groups, this affords each language user equivalent access to the system. Interactive Systems is planning on marketing the system to other European countries and is prepared to handle German and other European language versions.

The second unique feature of this system is the manner in which subject searching and access is provided. In this scheme, keywords are assigned within the Universal Decimal Classification (UDC) scheme so that one may search by keyword tied to appropriate UDC numbers which are then employed to locate the item on the shelves. One may also search directly by UDC, starting with a general number and receiving keywords associated with this number and moving through more specific related numbers, narrowing the search in this manner. This is an appropriate technique for libraries using the UDC classification but one which would apply only to very few libraries in the U.S. or Canada that do not use either L.C. or versions of the Dewey classification.

NON-COMMERCIAL SYSTEMS IN THE U.S. AND CANADA

Not surprisingly, most libraries having minicomputer systems acquired their software from commercial vendors. But some libraries have systems which were not supplied by outside vendors on a turn-key basis. These libraries either developed their own systems in-house using their own staff or contracted with outside software or systems engineering firms.

Again, the tendency in this area is initially to develop a system concentrating on a specific application such as circulation and then move to incorporate further applica-

tions. Libraries which have invested considerable effort and expense in software for a large-scale shared computer have tended to acquire minicomputers to provide front-end local processing capability to remove some of the host computer central load and improve responsiveness at the library end.

An example of this kind of minicomputer installation is at the Hennepin County (Minn.) Library (HCL). HCL employs a Digital Equipment PDP 11/34 with 96K words of memory to perform online cataloging data entry and serve as a remote job entry station for its link to the Hennepin County IBM 370/168 computer. But Hennepin County is planning on extending minicomputer usage to support public catalog access and circulation.

Another example of the front-end type of configuration is the medical library net-work serials management system PHILSOM III, developed by the Washington University School of Medicine Library. PHILSOM III is designed to employ a PDP 11 computer at each site for local holdings file maintenance. This is an alternative to remote batch job entry or mailing of sets of punched cards to update local holdings files on the central system. Also, a PDP 11 serves as a terminal communication handler for the network coordinator's terminal for the basic data file, which is maintained centrally for all network member libraries.

University of Minnesota's MILS

The first U.S. development effort using a stand-alone minicomputer as a host system designed for library applications was funded at the University of Minnesota Bio-Medical Library in 1972 by a three-year National Library of Medicine grant. After a continuation grant, the effort is now totally funded by the University of Minnesota through state funds. This system is simply called Minicomputer Library System (MILS), an acronym following in the tradition of the Minnesota Union List of Serials (MULS).

MILS began in the early minicomputer era when no vendor software was available and few hardware systems afforded the necessary features to permit easy use. MILS thus developed independently of Digital Equipment Corp.'s operating systems. MILS can be used with any PDP 11/34 or larger PDP 11 series minicomputer. In its Bio-Medical Library installation at Minnesota, it uses a large PDP 11/40 configuration. In its new St. Paul Campus installation, a PDP 11/60 system was installed in late November 1978.

MILS software employs multi-purpose program routines. Figure I-1 provides a simplified diagram of MILS software organization.

System operations handle functions such as input/output procedures, data conversions and service of interrupts, which are messages generated by users at terminals connected to the system. Applications are defined by the procedures software which direct the other portions of the system to execute the tasks required to build a user application. Terminal management involves video display and character printer devices for communication, security and main memory allocation. The interface routines set parameters to process data between its internal and external coded forms. Preparation of data
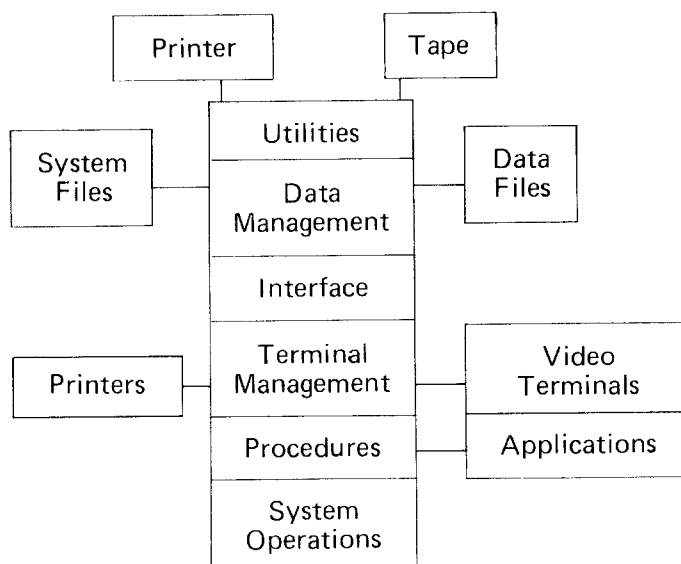
Figure I-1: MILS Software Organization.

for file input or output is controlled via the data management module. Special purpose activities such as file regeneration, magnetic tape conversion and printed record outputs are handled via the utility module. These modules are organized operationally into a system component and a library component.

By 1978, after six years, all of the system component modules and some of the library modules are operational in the Bio-Medical Library system. On the library application side, components for monograph cataloging, data entry, editing and conversion are operating together with an online searching capability by author, title and various numeric coded fields. The Bio-Medical Library has a record conversion in progress for monograph titles which started by using magnetic tapes of selected L.C. MARC records for works identified as MARC titles. Acquisitions/accounting and a user online cataloging inquiry application are being designed in the Bio-Medical Library system.

At the new St. Paul campus installation, development of a serials management application is scheduled for 1979 and further systems programming work is also planned. The University of Minnesota Libraries plan further MILS installations as funds and staff permit, all using common software and applications tailored to the requirements of each individual user library. Ultimately, communication links will be created among each library's processor, creating a local University of Minnesota Libraries network system.

Since 1977, the Lister Hill Center for Bio-Medical Communications has been developing a full functioning online library management system using Data General minicomputers and the MUMPS software system. Its system design, however, unlike MILS, is even more generalized and less dependent upon the present hardware. Thus, versions could be more easily created for different minicomputers than with MILS, although MILS can accommodate any configuration based on a PDP 11/34 or larger PDP11 computer.

MILS was designed to take maximum advantage of the DEC PDP 11 series hardware and assembly language during a time when limited memory made such an approach imperative. At the National Library of Medicine (NLM), a system for acquisitions and an interface to NLM online files are operational. Circulation and several other applications were being programmed in late 1978. This work is also being performed to meet the needs of the U.S. Army library, which has contracted for this system as its basic supporting system.

Both of the above institutions should bear watching as they are trying to achieve modular, flexible systems able to be replicated within their own institutions and have gained considerable support. Yet both of these systems use very different design and programming approaches.

Other Non-Commercial Systems

Another longtime user of minicomputers is the University of Pittsburgh, which uses a PDP 11/10 and an 11/34 along with a DEC System 10 shared large-scale computer for acquisitions, accounting, payroll, circulation and some online searching and complete in-process file control. Book form catalog production processing is also handled via this system. Similarly, the University of California at Berkeley General Library has used Datapoint minicomputers for online cataloging and data entry for approximately four years.

Although many systems in the U.S. and Canada depend upon Digital Equipment Corp. and Data General Corp. hardware, many other manufacturers are also represented. For example, Hewlett-Packard Series 3000 minicomputers are employed in the circulation system designed and installed at Virginia Polytechnic Institute and State University. One decided attraction of this system is its data base management facilities provided through a package called IMAGE and its associated retrieval/inquiry facility, QUERY. These are discussed further in Chapter IX. A powerful programming language called SPL also makes this a very attractive machine host for online library applications. Another Hewlett-Packard user is the Claremont College (Calif.) Honnold Library, which has a reputation as one of the pioneering users of data processing in its library. Also, industrial libraries such as Allied Chemical Corp. and a school system in Fairfax County, Virginia use Hewlett-Packard systems for their library support. In Canada, the International Development Research Centre Library has developed a multi-function system around a Hewlett-Packard 3000 which supports acquisitions, payroll, cataloging, serials check-in and claiming and full author, subject, title Boolean searching. This system will be offered to developing countries' governments and is in use by at least one other Canadian agency, the Sport Information Resource Centre, Ottawa.

The City of Cerritos (Calif.) Public Library, through a system development contract with Tamas Associates, has installed a custom designed system run on a General Automation 440-DS minicomputer programmed in commercial FORTRAN. The National Library of Canada uses equipment from IV Phase Systems, Ltd. (its full model name is the 7009-M System IV/90) with five magnetic tape units and 14 terminals for Canadian MARC file data entry and file building. Later processing is done from these magnetic tapes on a large IBM system.

FOREIGN DESIGNED NON-COMMERCIAL SYSTEMS

Of interest to special libraries and information centers which must produce abstract bulletins or published indexes from relatively large bibliographic data bases, the STATUS II system developed by the United Kingdom Atomic Energy Authority is used there and at the BNF Metals Technology Centre's Grove Laboratories in Wantage, England. The STATUS II replaces a system which requires a shared, large scale computer operating in batch mode for searching and file processing.

STATUS II runs on a Prime 300 minicomputer under the PRIMOS operating system at BNF. FORTRAN programming language is used to implement STATUS II. Originally, the software at Harwell was implemented by an IBM 360 and a Digital Equipment Corp. PDP 11/45 minicomputer. Only two man months' effort was needed to transfer the software over to the Prime 300. indicating how readily the software lent itself to another hardware system.

In this system a user is provided a searching capability using Boolean operators and word proximity on both full text abstracts and bibliographic structured indexes. Thus, both free text and controlled vocabulary searching are supported. A structured thesaurus can be handled by the system to support the controlled vocabulary approach, as are macro, word truncation, synonym and homonym handling facilities. One may also search by values imbedded within text, a feature particularly useful for numeric data files. Full editing facilities exist which allow the user to add, delete, change or replace any elements of information defined in the files.

From this online system, BNF produces magnetic tapes of abstract text for computerized photocomposition by author, subject and other index products. Online, the system may be used for retrospective searching and entry of bibliographic data or production of current awareness output based on user interest profiles. A large retrospective file conversion project is also underway to move data from its former batch system called ASSASSIN.

In Switzerland, the Hauptbibliothek der Universität Zürich Irchel has installed a system supplied by Interdata, AG which is used for acquisitions, cataloging data entry and bibliographic file maintenance. It employs a minicomputer by Interdata built around an Intel 8080 model microprocessor. The software is named TEX 400 and the configuration is a single user terminal with two floppy diskette units which provide low cost direct access storage on removable media. It is a preprocessing system essentially for data entry. with the production then finished on another university computer system via batch processing.

Australian libraries were very early users of minicomputer technology. At the University of New South Wales Library, a Digital Equipment Corp. PDP 11/40 operating under the Bell Laboratories-developed UNIX operating system is used for the cataloging system. Five video terminals provide data entry to the system, with additional applications being planned over the next several years.

Another system for cataloging is in place at the Australian National University, which uses an Interdata 8/32. Its processor is equipped with a very large main memory and disk storage which will undoubtedly make its minicomputer able to add many other applications.

For those interested in an Australian system based on a Digital Equipment Corp. PDP 11/34 minicomputer using the RSTS/E operating system and implemented in FOR-TRAN IV, a cataloging system with book fund accounting and order writing is operational at La Trobe University Library. Both book and card form catalog production is supported via this system and some programs also run on another computer system. In development is an acquisitions capability, of which the book fund accounting and order writing routines will form a part. The system is also used to list and process gift materials. This system performs some functions as an independent processor, but also communicates with a DEC system 10/70 host computer.

An even earlier use of minicomputers for cataloging support, perhaps even the first, was that of the Bodleian Library in Oxford, England, using a DEC PDP 11/20 for the conversion of a 1.25 million title catalog. This system performs online data entry and editing prior to processing by a large shared computer.

Although only one system was reported from Japan through the questionnaire, there may well be quite a few similar systems throughout the Orient, which run on Japanese micro- and minicomputer models. The Gunma University Library in the city of Maebashi uses an Okitac-4500C minicomputer in a small single user configuration serving a library with 300,000 volumes and annual serials subscriptions of 6000 titles.

In contrast with most other systems, however, the Gunma installation shows a departure from the prevalent trend of online processing back to batch processing for acquisitions, accounting, cataloging and serials control functions. The single terminal is used for an online circulation system. Thus two jobs — one real time, the other batch — can be serviced at the same time in what is known as a multi-processing system.

CONCLUSION

This overview has attempted to highlight some of the information gleaned from the wealth of data found in the bibliography, as well as to provide some information about systems which exist but may not have been formally reported in scholarly and professional journals. It should generally serve to stimulate the interest of librarians and encourage them to look at systems which are moving libraries in new directions, toward greater responsiveness and improved manageability. Ideally, it will motivate librarians in dealing with new problems and provide the knowledge needed to contend with these problems.

The next chapter examines two separate approaches to developing library applications and their respective merits and problems. Subsequent sections address various ways open to a library for actually creating systems. Toward this end, librarians should be aware of what is involved in proposal or bid development, the various types of modular low cost hardware, peripheral equipment devices and system software necessary to the creation of actual application programs.