In the development of library systems, the movement today is toward the so-called "total" or integrated system. This raises certain design and implementation questions, such as: what functions should be on-line, real time and what should be done off line in a batch mode; should one operate in a time-share environment or is a dedicated system preferred; is it practical to design and implement a total system or is the selective implementation of a series of applications to be preferred. Although it may not be feasible in most cases to design and install a total system in a single operation, it is shown how a series of application programs can become the incremental development of such a system.

Currently library mechanization is entering a new phase. The first phase, extending from 1936 to the mid-fifties, saw the development of a number of small, scattered, and essentially experimental Automatic Data Processing (ADP) library applications. These were punch card systems for purchasing, serials holdings lists and circulation control. During the second phase, which has been running now about 15 years, a large number of library applications have been mechanized. These include the production of catalog cards, book catalogs, periodical check-in, serials holdings, circulation control systems, acquisitions programs and searching of files, or
information retrieval. Systems librarians have been busy designing individual programs, building special computer stored files, implementing conversion of records and developing operating procedures for these various applications.

More importantly, they have been studying the library from a systems point of view in order to have a better understanding of the individual tasks performed and how they can be best accomplished with the available tools. At first concern was limited to individual applications in the library. Gradually some of the more perceptive systems analysts began to be concerned about integrating these various applications. Some simple examples are the generation of book cards for process control and circulation control as a by-product during the order-receiving cycle; the combination of subscription renewal, claims, and binding control with the serials holding program; the development of authority lists in book catalog programs; the simultaneous updating of accession files and circulation control files, etc. The purpose of many of these partially integrated programs was to reduce redundancy and make multiple use of single inputs.

The next step was to look at the library as a whole and consider it as a "total" or single, integrated system. Rather than building a series of independent applications programs, a number of libraries began to plan total systems in which the individual applications would be integrated segments. In the past year or two such efforts have been undertaken by the University of Chicago, Stanford University, Redstone Arsenal, the National Library of Medicine, Washington State University, University of Toronto, System Development Corporation, IBM and others (1, 2, 3, 4, 5, 6). It is this total systems concept which is the new and current development of library Electronic Data Processing (EDP).

At first, a total integrated system was conceived as a series of separate application programs utilizing separate files, but whose records have similar formats and field designators allowing for the multiple use of single inputs. A more advanced concept, however, calls for the construction of a single logical file, even though, physically, the individual record elements may be distributed over a number of tracks and storage devices. Operating on this central file are a series of program modules performing functions involving file building, searching, computation, display and printing. As each application is called for—that is, as the librarian prepares an order, receives an invoice, checks in a periodical, adds a call number, does some cataloging, charges out a book, etc.—the appropriate program functions are called into use. Attached to the file are a number of indexes or access points. One such program, for example, provides some eighteen indexes: author, permuted title, subject heading, descriptor, call number, invoice number, publisher, serial I.D., L. C. card number, borrower, etc.

It is not just coincidental that the development of the total integrated library system developed at the same time that computer hardware be-
came available that made it practical, especially in an economic sense, to operate a total library system. One of the basic elements of this hardware was the development of real-time, on-line, terminal-oriented, time-shared systems. At present, orders for on-line systems are increasing at such a rate that it is estimated in the June 23, 1969, EDP Weekly “that half of the computers installed by 1975 will be on-line systems.” Although there are a number of reasons why on-line, time-sharing and terminal oriented equipment made it feasible to build total library systems, the fundamental ones were that now the librarians could interact with their system and records and could, essentially simultaneously, perform a great variety of tasks.

The scientific and business communities have been quick to take advantage of these new capabilities. A number of computer manufacturers, software firms and service companies soon started to provide terminal oriented, commercial time-share services. By the beginning of 1969 there were some 35 such services in existence, serving over 10,000 customers; by the end of 1969 it is estimated there will be over 30,000 users.

Although these systems are often used essentially for remote job entry, their main attraction for users has been their on-line, conversational, real-time capabilities. The interactive, man-computer techniques made possible by commercial time-sharing services have been extremely valuable for problem solving applications, especially engineering and programming. However, the wide availability of text editing packages have also opened up these services for libraries. One of the first academic libraries to use such a service for preparing bibliographic records was the State University of New York at Buffalo (7, 8).

Many universities and industrial firms have developed their own time-sharing systems. A number of special libraries, notably those in IBM, were quick to take advantage of their in-house, time-share system to implement acquisitions, catalog input and library bulletin programs (9). The Defense Documentation Center over three years ago began preparing its bibliographic inputs on line. The SUNY Biomedical Network based in Syracuse does the same (10). The Washington State University library was one of the first academic libraries to implement an on-line acquisition program (11), and Midwestern University (12) and Bell Laboratories (13) now have on-line circulation control systems.

With the advent of time-shared, on-line capabilities and the potentiality of building total, integrated systems, librarians today who are planning EDP systems are faced with a number of design decisions: 1) Should the system be a real-time, on-line system or an off-line, batch mode operation, or a combination of both? 2) Is it desirable to operate in a time-share environment or is a dedicated system to be preferred? 3) Should one design a total, integrated system or should one selectively implement a number of individual applications? 4) If the decision is for an integrated system, how can it be incrementally implemented?
It is recognized that a program must be tailored to fit the available resources and that it is not always possible to build an ideal system. Nevertheless, design objectives must be established even though they cannot be immediately realized. If the ultimate objectives are understood, then the program development will be orderly and later reconversions will be kept to a minimum. Therefore, even though the design objectives may not be achieved for a number of years, they should be established so that current implementation can be carried out in a rational manner with some assurance that the system will grow and develop.

REAL TIME OR BATCH

Library operations have always involved a variety of interactive real time and batch mode procedures. Most operations dealing directly with the library patrons are, of course, in real time; reference question handling and charging of books are typical examples. Some technical processes, such as cataloging and searching for acquisition, are also essentially interactive, real-time operations. This means that the librarian completely processes each item by creating or updating a record or servicing an inquiry, one at a time, with little or no attempt to batch the identical operations for a number of items or inquiries. Other processing, however, such as preparing and mailing orders to vendors, sorting and filing charge-out cards, sending overdues, filing into the catalog, checking in periodicals, labeling, preparing binding, etc., is essentially done in the batch mode. In other words, batch and real-time operations complement each other, for whereas it is more effective to do some operations in real time, batching is more effective for other operations.

Librarians, therefore, expect and need both modes of operation. The actual distinction between these two modes is often lost in certain mechanized systems where everything is done in a non-interactive batch mode while interactive, real-time services are provided from printouts.

Many current library mechanized systems are really nothing more than processing techniques for producing the standard, hard-copy, bibliographic tools such as catalog cards, serials lists, book catalogs, orders, overdue notices and the like. Whenever the librarian wants to use the information generated by these programs, he consults the hard-copy files or lists. He does not interrogate a computer file directly. This approach has been typical of many other computer-based information systems.

When the first direct access devices (RAMAC) were made available for commercial and industrial inventory control, they were used primarily to update the records and to produce the inventory lists and card files which the user would consult for information. Later, as confidence developed in the machines, and terminals became available, the print-out lists and files were abandoned and the user began consulting the computer store directly.
Typically today in libraries using computer systems, inputs are processed in batches and outputs are produced in batches. Real-time services are provided from the print-outs: the catalogs, the on-order file, serials lists and so on. Even circulation control has been an off-line, batch operation. Although the charge-out may be made through a data entry unit, all that is actually accomplished at the time is that the transaction is recorded. It is only later that the transactions are batched and processed, the files set up for the loans, the discharges pulled from the file and the delinquencies handled.

Although librarians will not, in the immediate future at least, as readily give up their card catalogs and printed lists as business and industry are doing and as some enthusiasts believe librarians will (14, 15) — the queuing problem alone where the public must use the files would be very severe — some hard-copy files could be dispensed with in an on-line system. Certainly hard-copy files of circulation records, periodical check-in records, authority lists, on-order records and the like need not be maintained when these files are available via terminals.

Until now, practically all library machine processing, with a few exceptions, has been batched, off line and not interactive. In a non-interactive system, records are created and modified by manual preparation of work sheets followed by keypunching for data entry. In a library environment, for example, this means that the acquisitions librarian fills out an order work sheet that is given to a keypunch operator, who either prepares a decklet of punch cards or punches a paper tape or makes a magnetic record on tape. The cards or tape are then fed into the computer, the input is edited and errors noted and a proof copy is printed. The error messages and proof copy come back to the order librarian, who makes the necessary corrections. These are handed to the keypunch operator, who corrects and updates the record and inputs it again into the computer. If the operator has not introduced some new errors, the record is then processed. If she has, the record loops back again to the order librarian. The same story can, of course, be told about catalog records, journal and report records, and so on.

In an interactive on-line system, the originator of the information (in this example the order librarian) could key his data directly into the computer or could prepare a work sheet for operator input. The editing would occur at once by the terminal responding to each entry and verification or error messages would be returned immediately. The librarian or operator would enter the necessary corrections and upon acceptance of the record by the system would signal entry of the record into the file and the print queues as required. Also, during the preparation of the entry, the librarian would be using the terminal—presumably a display type terminal—to consult the files he needs, such as shelf list, orders outstanding, authority lists, etc.
A simplified flow chart comparison of an off-line and an on-line cataloging process would look something like that shown in Figure 1.

**OFF-LINE**

```
Catalog Worksheet → Revision → KP → Proof → Input Edit → Output
```

```
Correction → Correction
```

**ON-LINE**

```
Cataloger or Catalog Worksheet → Input Edit → Revision → Output
```

```
Error → Correction
```

**Fig. 1. Cataloging Process: Off-Line and On-Line.**

Although only a few library applications and no total library system are as yet on-line operations, a number of analogous operations are being carried out in other industries, such as order entry, inventory control, production scheduling, insurance policy information, freight waybilling, etc., so that one can make a few tentative assessments (16, 17, 18).

To begin with, in an on-line system a work sheet does not have to be prepared, and so the keypunch operator is eliminated. Because of the interaction of the originator and the system, all corrections and editing are accomplished at once, so that the turn-around time is very much less. Preparation of printed error messages and proof copy are eliminated and the total error rate is greatly reduced. Thus, although the reading-in of the individual records is slower in the on-line mode than in the batch mode, appreciably fewer messages need be read to complete a record in the on-line mode, making for more economical machine time. To this, however, must be added terminal and communication costs as well as the terminal supervisor program and the fact that most on-line work is done
during the prime shift, so that actual machine costs tend to be higher with the on-line system. Some, however, dispute this, claiming that, on balance, machine costs are equal.

Labor costs, however, are very much lower with the on-line system. As a general rule, computer input costs are 85% labor and 15% machine. Not only can a transcription clerk be eliminated, but the order librarian who prepares the original inputs on the terminal works very much more efficiently. Consulting hard-copy files and lists is more time consuming and less informative than interrogating machine files. In an on-line system, the librarian's necessary tools are brought directly to him and displayed rapidly and efficiently. He does not have to walk to the shelf list, the catalog or the on-order file and copy information. In a well-developed, sophisticated system some of the heavily used tools, such as the subject heading authority lists and class tables, would also be available from the terminal. Not only does the librarian not have to spend time going to the physical files, but since the information is computer stored, it is brought to him in a greater variety of forms and sequences than is available in the hard-copy files. For example, titles are fully permuted so that incomplete title information can be searched. Some systems librarians are proposing the use of codes and ciphers to search for entries, especially those with garbled titles (19, 20). All entries, including added authors, editors, vendors, etc., are immediately available even for uncataloged on-order items, so that searching is not restricted to main entries. It is not surprising, therefore, that clerks preparing computer inputs prefer working on line rather than off line.

One interesting discovery is that since operators can do so much more with on-line systems they tend to take more time to turn out a better product. Indications are "that significantly lower costs would have resulted if the time-sharing users had stopped work (i.e. gone to the next task) when they reached a performance level equal to that of batch users" (17).

Even with a circulation control system, there is higher system efficiency with an on-line operation. Every transaction, such as a charge-out or a discharge, is an actual inquiry into the file as to the status of the book and borrower and the answer is immediately available; therefore controls and audit procedures can be simpler. Elaborate error correction routines do not have to be provided in the program to identify improper inputs as has to be done with an off-line system. Incorrect loans are not made of restricted material, such as holds and reserves, or to delinquent borrowers. The system also acts as a locator tool for determining the location and availability of volumes. As a final note, on-line systems are necessary if effective networks are to be developed and decentralized services provided (21, 22).

The basic conclusion is that an on-line system can handle more work and provide more services at greater machine costs but lower labor costs than a manual or an off-line machine system. In view of the fact that
machine costs are coming down rapidly, while labor costs and throughput demands are forever rising, the future of the on-line machine system in the library looks very promising.

TIME SHARE OR DEDICATED SYSTEM

A number of librarians have had very unhappy experiences with data processing departments over which they had no control. Machines have been changed, schedules dropped, library jobs delayed or dropped for "higher priority jobs" and so on. One tendency, therefore, has been to try to get a library's own computer facility. But, as De Gennaro so succinctly summarizes it, "the economics of present day computer applications in libraries make it virtually impossible to justify an in-house machine of the capacity libraries will need dedicated solely or largely to library uses... Eventually, library use may increase to a point where the in-house machine will pay for itself, but during the interim period the situation will be uneconomical unless other users can be found to share the cost. In the immediate future, most libraries will have to depend on equipment located in computing or data processing centers... Experience at the University of Missouri suggests the future will see several libraries grouping to share a machine dedicated to library use... it seems reasonable to suppose that in the next few years sharing of one kind or another will be more common than having machines wholly assigned to a single library..." (23).

It is true that the small computers are getting more powerful and it is quite possible the day will come when small stand-alone computers will have the capacity to do all the jobs required by the library. For the time being, however, an on-line system supporting a number of terminals for a variety of tasks in the library requires a computer of a size which cannot be economically justified except for the very large libraries.

Also, one thing that is often overlooked is that implementing a large library system requires data processing technical support that is very seldom available on the library's staff. One need only look at the Information Systems Office of the Library of Congress, or the System Analysis and Data Processing Office of the New York Public Library to have some appreciation of the requirements for such technical support. Also, a large central system often has backup capabilities which provide insurance against breakdowns and interruptions.

The question really is not whether a library should time share or have a dedicated system, but rather whether or not the library has the necessary control over its segment of the total system. This segment is the library's property and its services are available to the library as set forth in the agreement made when the library became part of the data processing services.

Again, it must be emphasized that all this applies to systems which have to perform all library functions. Most libraries, however, in order to
get started and develop their programs, are beginning with small, stand-alone computers or are submitting batch jobs to a data processing center. Later, as their programs develop, they will have to upgrade their computer capabilities.

In view of the ultimate needs of a system which will support most of the major processing functions of a library, most libraries will have to have access to computer facilities whose full support they cannot economically justify. Time sharing, certainly for the immediate future, will be required for any on-line library system.

TOTAL INTEGRATED SYSTEM OR INDIVIDUAL APPLICATION

It is more economical to handle a variety of library applications by using a single file and a standard set of functional programs, than it is to provide a separate file and a separate set of application programs for each application. Not only is it more economical, but this total, integrated approach is, in its essential modularity, extremely flexible. Functions can be added, changed, or removed, and sequences can be re-ordered, so that the system can grow and change with changing needs and capabilities. Also, since the full record is available, if needed, for every application, added services, normally not feasible, are practical. For example, a circulation control system that, instead of having separate circulation files, keeps charge records in its central bibliographic file, can set a hold on all copies of a book, no matter where the copies are kept, as in the BELLREL system (13). Also, from a total record one can select various subsets and make different orderings to provide a variety of services.

The library systems currently being designed are essentially mechanized versions of existing manual systems. However, as experience is gained with these new systems, as more advanced equipment is made available, and as research and development provide new insights, these systems will evolve and change. For example, in some cases a major part of descriptive cataloging is becoming a part of acquisitions. The former compartmentalization in libraries is already breaking down. One should, therefore, be prudent and not lock up the system into tightly compartmentalized segments on the assumption that current file subsets will remain unchanged. It is advisable that each library activity have potential access to all system functions and to all records. In the present context, an activity may have no need for all functions, nor does it need the total record, but as the system develops it might very well need these added capabilities.

The problem, however, is that for a total, integrated system one must first build a complete structure including the file and all the functions—such as file building, search, compute, compare, display, print, etc.—as well as set up all the access points which are essentially indexes. In addition, all the overhead necessary for supervising the programs, managing the files, and monitoring the terminals must be provided for. To use an
analogy, one must first build foundation, walls and roof and install all plumbing and wiring before building any rooms. Consequently, the start up or initial investment is far higher than for implementing a single application program. Some who have undertaken the development of total systems did not fully appreciate this at first and have, as a result, had to replan their development programs.

Even if one could bring in a fully debugged program for a total system, there would still be the tasks of converting records, training staff, setting up operating manuals and working out procedures. Only as machinable records became available and the file grew and developed could various applications become operable. From a practical point of view, the implementation of a total system would have to be incremental; that is, once the basic system is installed, applications would have to be implemented one at a time and in some rational order. This is even more true where the programs for a total system have not been written as yet or where the library’s resources are such that it can only undertake one job at a time. From a practical point of view, one can develop and implement only one application at a time. Furthermore, as is often the case, the available equipment is limited and cannot do everything the library will ultimately want. It is necessary, therefore, to develop single applications and to design them in such a way that they can become part of an integrated system. It is also necessary to have a strategy and a plan to move up through the various levels of mechanization.

Today there are many who, although accepting a total, on-line system as a desirable goal, feel that it is impractical to consider because of costs and unavailability of equipment. A full analysis of economic change in terms of wage-cost rise and machine-cost decrease, of technologic improvement and of demand for added services, goes far beyond the limits of this paper. There is developing, moreover, a literature on these subjects (24, 25, 26, 27, 28, 29). Suffice it to say that an increasing number of librarians are becoming convinced that library mechanization is inevitable, that it will affect all operations of the library, that it will provide the highest level of service through direct, on-line, interactive systems and that, whatever today’s limitations may be, these changes are coming so fast that plans must be made now. These individuals are also convinced that whatever is now undertaken in the way of mechanization will evolve into an integrated system with many basic functions operated in a real-time, on-line mode.

IMPLEMENTATION OF AN INTEGRATED LIBRARY SYSTEM

Typically a library mechanization project will start with a single, relatively uncomplicated application that will not impact library operations very much, will require only a small amount of systems design and programming, and will run in a batch mode on a small equipment configuration. A typical example is the preparation of a serials holdings list. From
this first job, the librarian and his staff will become acquainted with data processing, will introduce the data processing personnel to some library requirements and will, hopefully, begin to develop procedures for working with the computing center.

Having passed this introductory stage, many librarians continued, as a rule, simply by developing the next application. Today, however, the more prescient ones are first assessing the total impact of mechanization and, having decided that their library will be mechanized, try to plan what their foreseeable goals are, then work out a plan to achieve these goals.

Having decided that the ultimate goal is a total integrated system for the whole library, which will provide real-time services and therefore must operate on line, the library planner will set priorities and work out a strategy to reach these goals. In some instances he can start designing a total system. In other situations, he does not have the resources to do so, but plans to make use of programs being developed for other libraries or of so-called standard, commercial packages, or programs which may be developed jointly with other libraries. He should realize that he can't just sit and wait for D-day when a total complete program will be wheeled in and a turnkey operation will be installed overnight. The lead time necessary for planning, training, conversion and installation is too often grossly underestimated, so that these preliminary preparations are neglected to the detriment of orderly growth and development.

Having established certain long-range goals, the librarian will tailor his current programs so that the library system will develop as smoothly as possible. He will try to keep the various subsystems and program segments as generalized and as modular as possible. He will structure his records so that they can ultimately be fitted together into a full bibliographic record. He will try to avoid using records so truncated that they will have to be discarded and recorded again later. He may, in fact, actually start with a full record that is comparable to his present shelf list or catalog card, even though there may be no need of the whole record for the current application. He will provide for a variety of print options, such as line width, number of lines, number of columns, etc., so that a separate print program will not have to be written for each product or to accommodate every change in style. He will try to organize his files so that the file structures and the record formats will not have to be radically changed when the system goes on line. He may store some of his records —his active on-order file, for example—on direct access storage devices. If he can, he will create access points to his large bibliographic file and store them on disk files too, even though he is currently operating off-line. Such direct access storage of indexes makes economic sense when very large files — and library files are large and grow very fast — must be searched or sorted. Aside from these immediate benefits, such a file organization requires little or no restructuring or record reformatting when
the system ultimately goes on line and becomes terminal oriented.

As early as possible, he will put his circulation control system on line. This is by far the cheapest and easiest on-line operation requiring the least investment and yet producing the most immediate benefits. Again, aside from the immediate benefits, this on-line operation represents an important building block for the ultimate total system. Aside from the current improved services, the experience of working on line and the opportunity to develop and refine processes and procedures will pay important dividends in the design and implementation of the total on-line system.

With knowledge of how he wants his system to develop, the librarian is now able to establish priorities and allocate his resources. The emphasis will be on file building, on capturing the record. Acquisitions programs or circulation control systems will come first. Work on the display terminal and communication will come later after searchable files have been built up.

In other words, an attempt is made to have a controlled growth through several levels of mechanization. A start is made with a simple, off-line, batch job. Then a beginning is made on building what is to become the main, central bibliographic file, the catalog. As soon as possible, parts of it are stored on direct access devices, so that it can be used more effectively and so that its structure will conform to the requirements of an ultimate on-line system. A simple on-line process is adopted as soon as feasible. Each application program uses standard functional modules in macro form and so on.

All this, of course, is highly oversimplified and may seem truistic to many. Nevertheless, there has been too much evidence of programs undertaken without adequate planning and of programs that have lacked continuity because adequate guide lines have not been established. Such failures are too often ascribed to changes in personnel or hardware. A project should be designed so that inevitable changes in personnel and hardware can be tolerated without its being wrecked. Therefore, the establishment of long-range goals can have a profound effect on the shape and success of current operations.

More and more librarians and systems personnel engaged in library projects are beginning to think in terms of total integrated systems. They are looking ahead and planning. They are designing and implementing their present applications not in a simple ad hoc way but as part of what is to become a total system.

REFERENCES


