7/ Automation in the Academic Library in the United States

CLAYTON A. SHEPHERD
Indiana University

The library traditionally has been an important link between the author and the reader in the process of information transfer, a chain which is initiated by the author and carried through from the publisher to the library acquisition of the material, its organization and cataloging, and finally, its circulation to its users. Each of the operations within the library complex need not be seen as discrete and independent activities, but rather as subsystems which are parts of the total service organization. These operations are often such that they lend themselves readily to automated processing, especially if the library is seen as an organism with closely interrelated functions.

The electronic computer and ancillary machinery, such as punch-card equipment, can be applied easily to many essential library operations. Although originally designed to perform mathematical computations, the computer has proven itself time after time in various applications oriented toward nonscientific areas, such as accounting, stock inventory control, personnel records, maintenance, etc. The computer has an advantage in such applications because it is ideally suited to process large amounts of material at high speeds and to perform its functions accurately. Thus, it is no accident that libraries turned to the computer to meet increasing demands of time and labor soon after its emergence as a business tool.

The university library, therefore, has been active in library

automation. Certainly the high volume of users the university library serves is an important reason, although there are many types of libraries serving large numbers. Further, the university library is likely to handle a broad spectrum of subject material (although this is true of other libraries too), and it often, in addition, builds extensive collections in depth in particular subject areas. More important, the university library is found in an environment of research and innovation. If there is a school for library science on campus, the library can turn to it for assistance in automation, and the library, in turn, can act as a fertile test bed for library school research. Further, many universities provide access to a computer on campus as a research tool, thus allowing a double opportunity for experimentation and the development of novel approaches to library operation. It is not surprising that in the university library environment some of the most significant strides are made in automation practices and procedures.

What is "Automation"?

In the mid-1950s it was determined that the electronic computer could be employed in specialized industrial operations that previously had been monitored and controlled, in some cases by entire staffs of highly skilled workers. Although this is only one type of operation to which the computer has been applied, it has been a highly significant one—especially since it has taken over repetitive tasks and has freed industrial personnel to concentrate on planning and development problems. The term "automation" has been traditionally applied in these situations; it calls to mind a device, monitored by a single individual, which controls an entire band of machines. Various process-control devices fall into this category.

Yet we must raise the question as to whether the term "automation" in its exact sense can be applied to library procedures. The application of computing equipment has certainly freed the time and aided the effort of library personnel, with the result that they have greater opportunity to concentrate on the various intellectual tasks so necessary to library operation. However, the

degree to which operations are automated—strictly "hands-off"—is open to question. The usual pattern of computer application in various library activities can be likened more easily to those found in accounting procedures and stock inventory control, which essentially are record-keeping functions rather than controlling functions. Thus, if we are to use the term "automation," we should use it guardedly so that we labor under no misapprehensions about the automatic nature of these processes. Perhaps "mechanization" would have been a better term, especially since the word implies a wide range of equipment and technique—the situation that holds in library contexts. But the term "automation" seems to be here to stay as far as library applications are concerned.

Automated Operations in the Contemporary University Library

A survey carried out in mid-1966 indicated that there were at that time 638 libraries using data processing equipment, and three-fourths of those were academic or special libraries. They represented 90 percent of the 942 libraries with firm plans for automation.[1] It would be impractical to catalog all the existing cases within the American university library complex that have undergone automation of one or more of their operations. In the first place, the literature is incomplete, and second, the picture is constantly changing. Therefore, only a few of the most recent examples in each area will be discussed, especially those with unique or noteworthy characteristics.

ACQUISITIONS

In both the acquisitions and serials ordering procedures in a typical university library, many operations are analogous to those in business and industry having to do with stock ordering, inventory control, and related accounting procedures. Not all automation programs in this area depend upon sophisticated computing equipment. The University of Maryland, for example, carries out its acquisitions operations on an IBM 407 accounting machine. An order deck is processed, including the generation of the purchase

order. Records for the cataloging department are produced, and an auxiliary listing for bill payment and management control is generated.[2] A combination of manual and machine operations to accomplish certain acquisitions procedures was reported by Ralph Shaw at the University of Hawaii Library.[3] In a system that has been in operation for more than a year and a half, better and faster control is now achieved than had been by the manual accounting system previously used, yet it costs about one-thirteenth as much. An example of economies that can be gained from automating several functions at once is illustrated by the system at Texas College of Arts and Industries. The library is able to use the computer at no cost for its acquisitions operations, since the cost is written off by the circulation and serials departments. In automating their acquisitions procedures, this university took the lead from the Pennsylvania State University system in incorporating the use of the change card, which furnishes cancellation data if an item is unavailable from a dealer, and which deletes all information pertaining to that order from the computer. If, however, an item is received, cost data are generated and input to the computer for accounting purposes; the change card then accompanies the item received to the cataloging procedures. The second change card receives the code number and, after keypunching, is used by the computer in providing a card acquisitions list.[4]

The University of Michigan Library was also concerned about costs in automating its acquisitions program. Feasibility studies had indicated that the system would be more costly at the time it was installed than the manual method. However, it was recognized that input volume could increase considerably without a concurrent increase in clerical staff. Rather modest equipment is being used there, as at other installations—in this case, an IBM 1460 computer.[5]

SERIALS CONTROL

In the survey previously mentioned, Jackson indicated that "the most usual current use of EAM [electronic accounting machine, i.e., punch card] equipment is for serials control." Serials control also seems to be the most prevalent area of application (in the university

libraries) using more sophisticated equipment, Eugene Jackson reports.[6] Indeed, because of the multiplicity of input items and the variation and irregularity of publishing period among serials, this area traditionally has caused the librarian a great deal of concern and has been considered fair game for automation. The haphazardness of the issuing policies of various publications has been pointed out by Roper. A serial may be identified by volume and year; by volume, issue, and year; by issue and year, and so on. In a 1965 issue of *American Documentation*, David Bishop and others posited a theory of publication that revolved around the related concepts of frequency and interval.[7] Publication patterns are often conceived of in terms of frequency, but serials of different frequencies may, in fact, have the same publication pattern. Conversely, serials with different publication patterns may have the same frequency. Irregular frequency has often caused problems in serials controls systems. The authors emphasize the concept of predictability, which they say is implicit in the idea of pattern. Thus, the computer is able to produce a check-in card that will predict certain characteristics of the next issue expected by the library. These check-in cards form the nucleus of the serials record system at the UCLA Biomedical Library. Fred Roper describes this feature of the system as follows:

> For each *title* the library currently subscribes to, the computer generates one card for the *next expected issue.* It is here that the UCLA system departs most obviously from the other systems in use. Rather than predicting the issues that should arrive during a specific time period, the computer recognizes that there *will* be a next issue and accordingly generates a check-in card for that issue, regardless of when it will arrive in the library. The check-in card will contain in most instances enough information to identify completely the particular issue expected, i.e., the elements of the internal number system needed to identify the publication.[8]

With such a system, the library apparently avoids a number of headaches arising out of the complex patterns of serials publication.

The interest in serials control and in making serials publications available to the user is exemplified by a project headed by Donald Hammer at Purdue University and funded by the Indiana State

Library in Indianapolis. The project is in the process of compiling a list of serials holdings at several academic and public libraries in the state of Indiana; this compiled list will be reproduced and distributed to all participating libraries. The overwhelming advantage in such a program is that it will provide an important communications link among the various libraries and will expedite the interchange of serials holdings not present in all libraries. Unfortunately, once again data input seems to constitute the difficulty, especially at Indiana University where there are over fifteen thousand titles in the serials collection and limited manpower with which to record serials holdings information for input to the computer. Nevertheless, the availability of such a list should more than outweigh the effort required in its preparation.

CATALOGING

The book catalog, of course, has long been recognized as an important adjunct to the card catalog, especially in locations in which the card catalog is not readily available. Many university libraries are generating book catalogs by computer; their style and use are somewhat similar, by and large, and will not be described here.

Richard Johnson, however, has reported on an interesting variation: At Stanford University, where the main entry concept has been abandoned, the system instead arranges works alphabetically by title under each subject heading and not by author. Experience indicates that this arrangement is quite satisfactory, especially for a selective collection.[9]

The most significant development in catalog automation over the past few years has been the implementation of the MARC Project by the Library of Congress.[10] The MARC Pilot Project grew out of a series of conferences, the first having been held in 1965, to investigate the feasibility of converting catalog data to machine-readable form. Sixteen libraries were selected to participate in the project; of these sixteen, eleven were universities with varying kinds of computing equipment. Since that time, magnetic tapes bearing information contained on LC cards, together with computer

programs which enabled their reproduction in LC card format, have been distributed. After certain modifications of the MARC format, the MARC II system was developed; at present tapes of this format are being distributed in a standardized form and are being used to advantage in the automation of cataloging procedures in various university libraries. One of the major advantages is that the basic bibliographic entry is already being provided.

These tapes are proving to be extremely useful, not only as a basis upon which to build an automated cataloging system, but as a source of information for selective dissemination of information programs as well. The practicality of the latter use has been demonstrated at Indiana University by a research project in which interest profiles of several faculty members from various departments were compared with the MARC tapes. Bibliographic information selected therefrom was distributed. This material was valuable to faculty members in keeping abreast of new publications in their fields of interest.[11]

Although Project INTREX will be discussed in greater detail later, it should be pointed out first that an important feature developed in conjunction with this project is the augmented library catalog. Alan R. Benenfeld, in a discussion of Project INTREX, indicated that the catalog is deliberately experimental:

> There are no ties to existing or past catalog structures which might otherwise constrain achievement of our experimental objectives. There is freedom to change both order and format of data elements to meet varying experimental conditions.

Second, the author indicates that the system is time-shared; that is, it is accessible to a number of users at remote locations who can have direct access to the catalog. Third, he indicates that it is "augmented," that is, that it goes beyond traditional catalogs and indexes in terms of the depth of information in the catalog. He suggests that "a large number of bibliographic data elements associated with diverse bibliographic forms have been identified and synthesized into a hospitable single catalog structure."[12] These features are especially significant in terms of Project INTREX's stated goal of providing a central resource for an information

transfer network to extend throughout the academic community. Since access to information is the touchstone of any library, especially one participating with others on a network basis and one in which a commonality of terminology is mandatory, such effort as is seen in Project INTREX will stimulate network development. The augmented library catalog as an object of research is being used to discover answers to questions of the commonality or differences in the needs of users in the university community, as well as to establish criteria for optimum search and display.

REFERENCE

One of the primary goals of the library is the retrieval of information. However, when the term "information retrieval" is encountered, it immediately calls to mind elaborate systems of indexing and filing organization and the complex Boolean combination of terms required to extract specific, detailed information from the file. Although such complex operations seem to be far removed from the traditional reference service provided by librarians, the two functions can be seen as different not in kind but merely in degree. Whereas the conventional definition of library automation excludes "information retrieval" operations, this area should not be ignored by library administrators, nor, more specifically, by reference librarians themselves, nor should it be ignored by the designers of automated reference systems.

Yet the need for information retrieval, per se, in the conventional library, has been seriously questioned. Andrew Osborn has put it this way:

> From the beginning with Vannevar Bush's paper of 1945 in *The Atlantic Monthly*, a two-pronged attack has been launched: on the one hand, libraries are branded as inept, inefficient; on the other hand, science and the computer represent knights in shining armor come to the rescue.

Osborn makes the point that high-speed information retrieval may be necessary in clinical and defense-related situations, but that in the typical university library context, time-consuming study and reading of source material makes rapid retrieval unnecessary.[13] Although

this same criticism—that is, that information is only very rarely needed "immediately" has been leveled at full-scale information systems oriented toward direct access and time-sharing capabilities, it certainly does not mean that no benefit would be reaped from applying certain principles of information retrieval to library reference work. Few would deny that there are shortcomings in the number and structure of access points to the information contained in a conventional library; and, although an increase in the number of access points would mean a corresponding increase in the difficulty of both generating and maintaining them, more effective access to the information would thereby be provided to the user. Thus, not speed, but accessibility, becomes a prime consideration.

One of the first experiments in automating a reference service was made at the Institute for Computer Research at the University of Chicago. Two hundred and thirty-four biographical books were categorized as to types of subjects included and as to the contents of the uniform entries they contained. Cherie Weil, the author, believes that by automating this phase of reference work the reference librarian could be freed for more intellectual activities. In so stating, she is recognizing one of the most powerful applications of the computer in library automation of this type: the computer and the human intellect can operate as a team, with complementary tasks resulting in optimum use of the library collection.[14]

Once again Project INTREX is providing valuable insights into library operations and into the possibilities of improving their effectiveness. The augmented catalog portion of the system, referred to earlier, is used in an experimental library storage and retrieval system, in which catalog requests can be made by specifying subject terms, authors, or titles, or any combination of these. The items retrieved by this initial inquiry may be narrowed further by requiring specific matches to be made on additional fields of the catalog records. Thus, a distinct parallel exists between library access and the development of search strategies of the type found in full-blown information retrieval systems. An additional feature permits dialogue between the user and the computer, which makes available to him a user's guide for instructional purposes. In this way

his method of gaining access to the file can be made more effective.[15]

CIRCULATION

One of the most prolific areas of development in the use of the computer and other mechanized devices in the university library is circulation control—a quite understandable circumstance since the circulation activities in a large library are likely to number several thousand a day and present a management-control problem of considerable magnitude. This is especially true in a library that supports a wide variety of users, with different circulation procedures for undergraduates, graduate students, and faculty members. Whereas circulation control is accomplished by the use of a variety of equipment, most applications are performed on the IBM 357 Data Collection System. In this system, a book is charged by requiring the borrower to present it to the circulation attendant along with his plastic identification badge. This card contains a punched code representing the student's identification number. In the book the pocket contains a pre-punched master book card, which is removed from the book pocket and inserted, along with the borrower's identification badge, into the 357 input station. The master book card furnishes the call number and accession number for each book charged. The input station punches the charge card, which contains all the necessary information.

A variation of this process is represented in the two-card system, in which a cartridge containing a pre-set due date is inserted into the 374 cartridge reader. The system then produces two cards, which include the information outlined above plus the due date. One of these cards functions as a charge card and keeps the circulation files up to date, and the second becomes a due-date card to be inserted into the book pocket. When the book is returned to the library, the due-date card then becomes the discharge card.[16]

The IBM 357 Data Collection System is currently operating in nearly fifty college and university libraries. Not only is it performing a creditable job in controlling the circulation activities of libraries, but it is supplying important additional information that otherwise

would not have been made available. For example, as a guide to more effective library management, an automated circulation system can provide information concerning the scheduling of personnel who staff the various public service desks. It can also be used to discover which areas of the collection are used most heavily, and it can find some of the characteristics of the patrons of the library as well as the proportion of faculty to student users.[17] Much of this information can be directly related to budget projections of the university, as well as to its acquisitions policy.

Libraries utilizing automated circulation systems differ as to their economic advantages. Many feel that little gain is made in personnel economies; the initial cost of the equipment itself, they believe, has to be considered. However, the advantages over a manual system must be recognized. Not only do automatic systems provide more efficient service to the users of the library, they also tend to be much more accurate in keeping records of the locations and due dates of books charged. Bruce Steward, in describing the circulation system at Texas A & M University Library, has suggested that "perhaps most significant of all is the fact that the present system can accommodate double or triple the present volume of circulation with only a minimum increase in personnel."[18] The significance of this statement cannot be overemphasized, in view of the tremendous continuing growth that is being experienced by the typical American university. It is incumbent upon systems planners, in designing automated library systems, to anticipate this growth and to accommodate it either by installing systems that can handle ever increasing workloads or systems that can be expanded in a modular fashion, without redesign and reprogramming.

Ideally, a circulation system should be capable of keeping immediately up to date, so that the location of all holdings of the library can be determined at any given moment. Essentially this means immediate updating of the master file upon the completion of each individual transaction. Such techniques have been developed at several libraries, notably the Illinois State Library, which uses a disc file in conjunction with an IBM 1710 computer.

As a matter of fact, several operations which are performed by the

conventional non-automated library are by nature real-time (that is, constantly current) in their systems characteristics. Many of these real-time characteristics have been lost, Audrey Grosch has pointed out, when various operations of the library have been mechanized.[19] However, there is no reason why a mechanized system cannot retain this characteristic, rather than the common system of depending on batch mode processing, with its coincident delay in turn-around time. Yet although this may be true from a systems standpoint, from the standpoint of economics, computing equipment that provides real-time capability costs significantly more than simple devices that operate strictly in batch mode.

NETWORK APPLICATIONS

A typical library is unable to develop an exhaustive, infinitely comprehensive collection of materials. Except in the areas of most common use, such as general reference works, the typical college and university library tends to grow more vigorously in specialties represented most strongly in the curriculum and research programs at the institution. This situation, added to the problem of increasing costs, has stimulated libraries to develop forms of cooperation that either amplify existing systems or depart completely from traditional methods of interlibrary loan. In this way materials can be shared more effectively, and the existence and location of specialized collections can be made known more effectively. Outstanding progress has been made in the development of networks connecting various campuses of a single university as well as in providing interchange among different universities. Certainly, in the latter case, the problems are more difficult, if for no other reason than that the participants do not function under the aegis of a single administration. Yet it is precisely in this area that the most spectacular work has been done.

One of the most extensive communications networks in the university library field is that of Biomedical Communication Network of the State University of New York, which was conceived in the fall of 1965. The SUNY Network is an on-line real-time

system, which links all three State University of New York Medical Center Libraries with each other and with the University of Rochester Medical Center Library and other university medical libraries in the area.[20] On 30 August 1968 the National Library of Medicine began experimental on-line computer communication with the SUNY Network, so that cataloging information could be exchanged with the participating libraries. An additional link is also being developed with a specialized information center, the Parkinson's Information Center at Columbia University, which is a part of the National Institute of Neurological Diseases and Blindness Specialized Information Center.[21]

Another significant advance in networks among university libraries is NELINET, the New England Library Information Network. Conceived in 1964, NELINET is designed to provide automated technical processing services to various libraries in the New England area. At present it is serving five state-university libraries—the universities of New Hampshire, Vermont, Massachusetts, Rhode Island, and Connecticut—and it is expected to extend to other university libraries in New England. It is designed to provide real-time access to the catalog information provided by the Library of Congress MARC Project, to produce catalog cards, book labels, and book pocket labels, and to provide automated order control for acquisition.[22]

Project INTREX is probably one of the most ambitious of the projects linking university library facilities by networks. Project INTREX was conceived at the Summer Studies Center of the National Academy of Sciences, at Woods Hole, Massachusetts, in a session held in September 1965. Its admittedly ambitious goal is to establish an information transfer system among a number of institutions "throughout the nation and perhaps the world." Based on work at Massachusetts Institute of Technology, Project INTREX "is expected to yield significant contributions toward the modernization of all large libraries and, indeed, toward the general improvement of information transfer." Providing some of the most advanced thinking in the field, the developers of Project INTREX recognize

the value of intercommunication among university libraries by means of the sophisticated equipment needed to provide communications flexibility and convenience.[23]

Further reports of cooperation among university libraries for information interchange are encouraging. They indicate the growing awareness of the need of scholars for more effective dispersal of materials and information. A consortium of five universities in Washington, D. C., has been formed to study the practicality of various levels of intercooperation. A jointly operated computer, to function in batch mode, was recommended for implementation as soon as practicable. A sophisticated computer system for real-time applications, including the tie-in with remote terminals, was also considered for the future, since, although the cost was too great for present use, such a computer system could grow out of the system suggested for the present.[24]

In 1968, a consortium was formed by eleven colleges within a thirty-five-mile radius of the University of Dayton. A computer was installed at Dayton to serve these colleges, which composed the Dayton-Miami Valley Consortium. The system is designed to store records on all volumes currently in the University of Dayton's libraries, as well as to provide research capabilities for the various student bodies. This action represents still another effort directed toward the sharing of facilities and collections. Network interchange also aids in the elimination of a great deal of technical services duplication.

Another significant method of interchange provided by modern technology is the facsimile transmission of materials. Facsimile transmission has reduced tremendously the turn-around time between request and receipt of materials, and it has caused a decrease in the amount of administrative procedure that otherwise would have been required.

Several university libraries have become interested in facsimile transmission systems, such as those connecting the campuses of Pennsylvania State University. Yet there seem still to be considerable problems in terms of resolution, lack of flexibility, and high cost. Such equipment as the Xerox Telecopier was designed for use over

ordinary telephone lines. It has several advantages, including those of convenience and cost. The Alden Company has developed a book scanner that seems to present the advantages of relatively high resolution and flexibility in speed.

The apparent difficulties in telefacsimile processing are suggested by the fact that, according to one source, only one working interlibrary facsimile system is in operation today—at Pennsylvania State University, among eight of its twenty campuses.[25] A major advantage of telefacsimile transmission is the rapid turn-around time it provides. However, research conducted at several universities, including The Massachusetts Institute of Technology, the University of Nevada, and the University of California, has indicated that an insufficient number of users need the turn-around time that telefacsimile can provide. A report discussing this research, published in 1968, indicates that emphasis should be placed on the improvement of existing manual systems, rather than on the attempt to develop faster turn-around techniques.[26] Naturally, in evaluating such findings, cost is not an insignificant criterion. The industry promises that it will not be long before it can provide inexpensive teletransmission. When it does become generally available for library applications, the consideration of fast turn-around may be seen in a different light—that is, in terms of economic feasibility.

Problems in University Library Automation

It was stated earlier that the university library, for various reasons, presents a fertile field for research and development of mechanized techniques. Those involved in systems design and development for libraries will find themselves facing many problem areas not encountered elsewhere. The following problems often appear in all stages of the development of automated processes. They should be clearly recognized by library administrators and systems planners.

SYSTEMS PROBLEMS

The overwhelming majority of the applications of automation within the university library environment have taken either one of

two patterns: Either a single function has been automated to the exclusion of all others, or a function has been automated prior to the others, with uncertain plans to extend automation to other areas. In the latter case, particularly, problems are likely to arise. The library is often seen as an aggregate of functions, each of which is highly compartmentalized in nature. In all too many cases, the interrelationships among these functions are overlooked, and a single operation is automated with total disregard for other library activities.

The pitfalls of such an attitude have been recognized through bitter experience and have made the concept of a "total systems approach" a somewhat hackneyed and overused term in the library automation field. As a matter of fact, there is nothing inherently wrong in mechanizing only one activity at a time—in many cases, this may be the only economical and practical method of proceeding with library automation. The University of Chicago Library is a good example of a university library, as are those at Washington State and Stanford, committed to the total systems approach. Yet they are libraries that implement their systems in modules. Other universities, such as Yale and Harvard, are using the evolutionary approach, which implies moving through various stages of development from the manual to increasingly complex machine systems. Such implementation is applied to the entire library complex, so that the services and operations of the library can be continued with a minimum of disruption.[27] There are, however, certain disadvantages to the evolutionary approach. It does mean that procedural changes could well be taking place simultaneously in all departments. Further, it implies several levels of system development—all too often we mistakenly assume that changeover from one system to another can be made on a one-to-one basis. Experience has shown, however, that without special consideration being given to file re-formatting and procedural restructuring to facilitate the use of the newly selected equipment, such a procedure is wasteful of the capabilities of computing equipment. Further, lack of such consideration leads to general discouragement and an unwillingness to

support further automation. Although making the big jump from manual to completely automated systems all at once has its own risks, it can successfully be accomplished through proper systems planning; we must realize that each operation within the library involves a relationship with all others, and that adherence to a "total systems approach" is vital to success.

A novel shortcut to library automation is outlined by Dougherty and Stevens in their report of an experiment using computer programs and data banks generated at one university and transferred to a second.[28] Specifically, the library of the University of Colorado transplanted the computer programs and data developed and generated at the University of Illinois libraries. Policy decisions as well as the inappropriateness of some of the computer programs seemed to present such problems that the process proved unpracticable. This experience is worth noting, however, on two counts. First illustrates the need for a process of systems design and development suitable to a particular situation. Although the operations of various libraries are often quite similar, it must be recognized that there are differences in procedure and policy, and that procedures and programs have not yet become so standardized that they can be readily transferred to different institutions. Dougherty's report, moreover, is worthwhile for a second and perhaps much more significant reason: it is a report of an attempt at library automation that was made *in an experimental context.* As a matter of fact, some areas of library automation have progressed so insignificantly that there is little justification for taking any steps to move directly into operational automation. Furthermore, in more than one case a library has gone ahead with ambitious plans for implementing automation programs (for either one or many library functions), and has found only after the expenditure of vast sums of money that the system was, in the final analysis, inoperable.

These experiences, in fact, are likely to have a chain effect: library planners sometimes seem to be transfixed by the glamour of the computer; they then describe impending or current attempts at implementation in somewhat optimistic, if not downright mislead-

ing, terms. Subsequently, if the system does not live up to expectations or, in fact, fails altogether, no further mention of it is made in the literature, since the attempt was made not as an experiment, but with a personal and professional commitment to its success. Other organizations, however, swayed by the tone of enthusiasm in initial reports, may well be misled, follow in their footsteps, and experience failure too. Evidence of the use of the scientific method, therefore, is often refreshing, because research was performed and acknowledged and because it was applied in the design of the systems. The setting up and testing of hypotheses, the construction of pilot systems, the use of modeling and simulation techniques and other such devices are powerful means of evaluation of systems before they are adopted widely.

A necessary prelude to systems design is an analysis of the existing system. A possible pitfall may be the conversion of the existing system to the mechanized system on a one-to-one basis; a thorough understanding of the goals of an operation and the methods by which it achieves those goals is essential. In too many cases the attempt to automate is made either by a librarian who has little knowledge of the equipment with which he is to be dealing, or by a computer expert who has slight insight into the problems of library operation and administration. Ideally, these two sets of skills would be combined in a single individual; rarely, however, is this the case. On the other hand, teams composed of librarians working with systems analysts have been successful. For example, Becker has reported the experience at Penn State in 1963 where, while in the course of automating its Acquisitions Department, the library used a team that included a professional librarian and an industrial engineer. The resulting new system was able to accommodate additional work load, reduce errors and time delays, and cause a significant reduction in unit costs.[29]

With all the talk of total systems, it appears that no completely automated library system is yet in existence. This objective has been achieved only partially.[30] Thus, information is lacking on the efficacy of total systems. One can only assess those instances in

which individual operations or groups of operations have been automated and speculate whether the success of the whole, when automated, will be greater than that of its component parts.

PROBLEMS OF DATA INPUT

In many applications areas, computer technology has far outstripped our ability to use it effectively. Particularly in one area—that of input of data—little practical change has been seen since the initial development of the computer itself. This is especially true in the context of the input of non-numeric information, especially bibliographic material, which tends to be lengthy and sometimes difficult and tedious to transcribe. The technology of optical character recognition is still embryonic; it may yet be several years before devices able to read accurately and quickly several fonts and type styles achieve full operational status. Until then, we will continue to be limited in our means of input of documentary material.

The most widely used conversion technique of written records to machine-readable form is keypunch. Other techniques have been used, such as paper tape typewriting and optical scanning of preprepared forms. The costs of various methods seem to be almost the same; one author points out, however, that input cost is, after all, relatively inexpensive when it is compared to the total cost of the automated library program.[31] But these costs seem to be estimated without regard to the larger questions of logistics and materials handling, although the author stated that in considering conversion of shelf lists to machine-readable form, the need of sending the material to a service bureau would probably not be a problem, since a library can continue to function for limited periods without portions of the shelf list.

To the three usual techniques for converting bibliographic information to machine-readable form are being added other methods showing considerable promise. The use of on-line input terminals in various forms has considerable potential for maintaining a consistently accurate and up-to-date library file, as well as for providing greater convenience for library personnel. This technique

is in use in the libraries of the State University of New York at Buffalo. In the conversion of the card catalog to machine-readable form, the IBM Datatext system was selected, and it proved to be quite flexible in both its editing capability and in its provision for upper and lower case character input.[32] Such on-line capability can be especially advantageous when considerations of editing and immediate turn-around of material are important.

COSTING

Experience has shown that cost projections for prospective areas of automation are often faulty. The tendency, unfortunately, to underestimate implementation and operating costs occurs so frequently that the library administration finds itself saddled with a system that has far exceeded its projected initial and operating costs. There are several reasons for this, not the least of which is the lack of systems planning that will allow the smooth, effective meshing of various areas as each is implemented. This point is made effectively by Johnson, in discussing the implementation of a mechanized book catalog system at Stanford University:

> The determination of actual costs is a difficult undertaking and a meaningful comparison of costs estimated during the planning process is filled with problems, uncertainties, questions of definition, etc. In a sense, it is impossible to make a meaningful comparison. An element measured during planning is not the same as the element actually achieved.[33]

Closely linked to this problem is that of justification of expenses incurred—a problem encountered not only in the university, but by almost every library in operation. Certainly, use statistics can be important in determining the effectiveness of a library in meeting its objectives. Yet there is no direct way, similar to industry's procedures for cost-benefit analysis or measures of return-on-investment, to determine whether an expense the size of that usually involved in library automation can be justified except, perhaps, in terms of greater efficiency and smaller staff. (The latter has generally not been the case.) These circumstances place the library administra-

tor in a precarious situation and make adequate funding of research and development projects difficult to justify. Libraries are operations in which the public seems to be unwilling to invest significant amounts of capital for research and development. They seem to have fairly low priority when compared with such activities as chemistry, physics, and defense programs, into which corporate and government funds are being poured in vast quantities.[34]

LANGUAGE PROCESSING

Even though a great many applications of library automation are oriented toward technical services, more and more attention is being given to information retrieval applications, including the input and processing of natural language. Bibliographic materials, to be sure, provide problems in many aspects of computer processing, if only because of their bulk and lack of precise format. A problem of another order exists, however, when the computer is expected to analyze natural language and to derive meaningful indexes therefrom for retrieval purposes. The automation of retrieval processes and the production of retrieved material has progressed a great deal further than that of pre-analysis of the printed document. The ambiguities inherent in language, both in terms of initial indexing and in terms of the user's dialogue with the system, need further analysis and resolution before effective information systems can work economically.

NETWORK PROBLEMS

If it is sometimes difficult for a single library to maintain internal consistency and smooth flow among the various sub-systems within it, it is infinitely more difficult to provide the ease of interchange and communication so necessary among various separate libraries within a network complex. The lack of compatibility not only between systems and hardware, but between the varying forms material takes in component parts of a network, creates difficulties. Such difficulties may, indeed, offset any benefits that might accrue from such interchange.

Inexpensive transmission facilities among the various components of a network are needed. The usual means of communication between computer components is by telephone lines. It often happens that greater expense is incurred (in a remote time-sharing communications situation) by the use of long distance telephone lines than by the use of computer time. The Federal Communications Commission and other agencies, fortunately, are studying this problem. Solutions, one hopes, will soon be forthcoming.

COMPLETENESS OF THE LITERATURE OF AUTOMATION

In planning for automation of one or more library processes, the administrator is likely to turn to the published literature for guidelines, and he is likely to seek information about those systems functioning most successfully. The literature, unfortunately, presents an imperfect picture of the experiences of many libraries in their attempts at automation.

It is too bad that library automation is undertaken so often not in a research or in an experimental context, but in one which expresses total commitment to the success of the project. If the former were more common, we would find in the literature a great many more reports revealing negative conclusions. Instead, we find such commitment to success that the literature is filled with accounts of pilot projects, descriptions of planned systems, and stories of automation that has progressed to one stage or another of operation. In many instances it is only by word of mouth that we learn that systems that were once reported as completely operational have since become defunct, that pilot projects have failed, and that plans for automation have not been carried through. In each case, the reasons for the breakdown should be openly discussed in the literature—negative results, in development as in research, are valuable and can lead to savings of time and effort. There are, of course, many college and university libraries successfully operating automated systems. The fact that more and more university libraries are turning to automation for solutions to their problems is ample justification for accurate reflection in the literature of the relative merits of innovative techniques.

The Future

When we review the progress toward library automation in the university during the sixties, we find important advances indeed. The common acceptance and wide availability of electronic data processing equipment, used both in batch and real-time modes, has been significant. Its use in business and industry has demonstrated its suitability for applications in which factors of volume, speed, and accuracy make special systems demands–factors relevant to many library operations.

To some, library automation has not proceeded as rapidly as it might have. Some of the problems discussed have deterred many attempts, and they have caused others to be stillborn. Such setbacks, however, have paved the way for the successful completion and operation of other automated systems, which in turn provide operations and management experience upon which even more efficient systems can be built. Others believe that conservative university library administrators have impeded progress in research and systems development. Although caution can be carried to extremes in some cases, careful planning for such expensive equipment must be recognized and applauded.

Yet strides have been made and will continue to be made. Technology promises to provide flexibilities and economies of a high order. Inexpensive communications media will vastly increase the practicality of remote access to centrally stored library files. Computer network technology and technique will allow more effective interchange of information between libraries. Indeed, long range plans are being made today for information networks that will link university libraries with other information facilities, such as industrial centers, public libraries, and government information centers. In this way the broadest range of information can be made available to as many people as possible. The improvement of image-transmission devices will contribute to the usefulness of such networks.

Thus "past is prologue to the future"; the successes of multi-

library network complexes found in the university are encouraging similar efforts elsewhere. Similarly, automation within individual libraries is leading to more widespread application of techniques already developed and to increased systems effectiveness. The benefits are sure to outweigh the cost and effort in terms of service to students and faculty, and the benefits should lead to an overall enhancement of the educational process.

NOTES

1. Eugene B. Jackson, "The Use of Data Processing Equipment by Libraries and Information Centers: The Significant Results of the SLA-LTP Survey," *Special Libraries* (May-June 1967): 317-327.
2. Carl C. Cox, "Mechanized Acquisitions Procedures at the University of Maryland," *College and Research Libraries* 26, no. 3 (May 1965): pp. 232-236.
3. Ralph R. Shaw, "Control of Book Funds at the University of Hawaii Library," *Library Resources and Technical Services* 11, no. 3 (Spring 1967): 380-382.
4. Ned C. Morris, "Computer-Based Acquisitions System at Texas A & I University," *Journal of Library Automation* 1, no. 1 (March 1968): 1-12.
5. Connie Dunlap, "Automated Acquisitions Procedures at the University of Michigan Library," *Library Resources and Technical Services* 2, no. 2 (Spring 1967): 192-202.
6. Jackson, "Data Processing."
7. David Bishop, Arnold L. Milner, and Fred O. Roper, "Publication Patterns of Scientific Serials," *American Documentation* 16, no. 2 (April 1965): 113-21.
8. Fred W. Roper, "A Computer-Based Serials Control System for a Large Biomedical Library," *American Documentation* 19, no. 2 (April 1968): 151-157.
9. Richard D. Johnson, "A Book Catalog at Stanford," *Journal of Library Automation* 1, no. 1 (March 1968): 13-50.
10. "The MARC Pilot Experience: An Informal Summary," Information Systems Office (Washington: Library of Congress, June 1968).

11. William J. Studer, "Computer-Based Selective Dissemination of Information (SDI) Service for Faculty Using Library of Congress Machine-Readable Catalog (MARC) Records" (Ph.D. diss., Indiana University, 1968).
12. Alan R. Benenfeld, *Generation and Encoding of the Project INTREX Augmented Catalog Data Base*, (Cambridge, Massachusetts: Massachusetts Institute of Technology, August 1968), p. 2.
13. Andrew D. Osborn, "The Influence of Automation on the Design of the University Library," *Library Planning for Automation*, ed. Allen Kent (New York; Washington: Spartan Books, Inc., 1965), pp. 57-59.
14. Cherie B. Weil, "Automatic Retrieval of Biographical Reference Books, *Journal of Library Automation* 1, no. 4 (December 1968): 239-249.
15. "Massachusetts Institute of Technology: Project INTREX Semi-Annual Activity Report, 15 March 1968 - 15 September 1968," *PR*-6 (Cambridge, Massachusetts: 15 September 1968).
16. "Three Systems of Circulation Control," *Library Technology Reports*, Library Technology Project, American Library Association, (Chicago: May 1967).
17. Floyd Cammack, and Donald Mann, "Institutional Implications of an Automated Circulation Study," *College and Research Libraries* 28, no. 2, (March 1967): 129-132.
18. Bruce W. Stewart, "Data Processing in an Academic Library," *Wilson Library Bulletin* 41, no. 4 (December 1966): 392.
19. Audrey N. Grosch, "Implications of On-Line Systems Techniques for a De-Centralized Research Library System," *College and Research Libraries* 30, no. 2 (March 1969): 112-118.
20. "New York Biomedical Communication Network," *Information Retrieval and Library Automation* 3, no. 9 (February 1968).
21. "National Library of Medicine On-Line With SUNY," *Information Retrieval and Library Automation*, vol. 4, no. 6 (November 1968).
22. William R. Nugent, *NELINET—The New England Library Information Network*, (Cambridge, Massachusetts: Inforonics, Inc., 1968).
23. Carl F. J. Overhage, and R. Joyce Harmon, ed. *INTREX, Report of a Planning Conference on Information Transfer Experiments*, (Cambridge, Massachusetts: M.I.T. Press, 1965).

24. "Automation for University Library Consortium," *Information Retrieval and Library Automation* 4, no. 7 (December 1968).
25. Harold G. Morehouse, "The Future of Telefacsimile in Libraries: Problems and Prospects," *Library Resources and Technical Services* 13, no. 1 (Winter 1969): 42-46.
26. William D. Schieber, and Ralph M. Shoffner, *Telefacsimile and Libraries: A Report of an Experiment in Facsimile Transfer and an Analysis of Implications for Inter-Library Loan Systems*, (Berkeley: California University, Institute of Library Research, 1968).
27. Richard DeGennaro, "The Development and Administration of Automated Systems in Academic Libraries," *Journal of Library Automation* 1, no. 1 (1968): 75-91.
28. Richard M. Dougherty, and James G. Stevens. *Investigation Concerning the Modification of the University of Illinois Computerized Serials Book Catalog to Achieve an Operative System at the University of Colorado Libraries*, (Boulder, Col.: University of Colorado Libraries, 1968).
29. Joseph Becker, "Systems Analysis: Prelude to Library Data Processing," *ALA Bulletin* 59 (March 1965): 293-296.
30. Grosch, "Implications of On-Line Systems Techniques," pp. 112-118.
31. Richard E. Chapin, and Dale H. Pretzer, "Comparative Costs of Converting Shelf List Records to Machine-Readable Form," *Journal of Library Automation* 1, no. 1, (March 1968): 66-75.
32. Frederick M. Balfour, "Conversion of Bibliographic Information to Machine-Readable Form Using On-Line Computer Terminals," *Journal of Library Automation* 1, no. 4 (December 1968): 217-226.
33. Johnson, "A Book Catalog at Stanford," pp. 13-50.
34. Louis A. Schultheiss, et al., *Advanced Data Processing in the University Library*, (New York: Scarecrow Press, Inc., 1962), p. iv.
35. Burton W. Adkinson, and Charles M. Stearns, "Libraries and Machines: A Review," *American Documentation* 18, no. 3 (July 1967): 121-124.