

# The Prospects For Mechanization

BY PHILIP M. MORSE

AT FIRST SIGHT, a large library appears to be a "natural" for the application of data processing techniques. Many library operations appear to be routine actions which could easily be mechanized; since libraries are notoriously understaffed the mechanization would not create unemployment; more fundamentally, since a library is "just a collection of data," it should be peculiarly amenable to the newer techniques of data processing; in addition libraries are growing so rapidly they must mechanize or they will become unmanageable.

Closer examination, however, brings to light a number of basic difficulties impeding rapid mechanization. These difficulties will surely be solved, and the large libraries of the future will almost surely be using electronic data processing techniques to a large extent in their operations. But the full application of these techniques may take several decades. In fact it may turn out that it is easier to "automate" banks and insurance companies than libraries.

This does not mean that a start should not be made soon.

In fact, unless large university libraries begin soon to mechanize, their operating effectiveness will deteriorate with increasing rapidity.

To see why this is so, and to see where a start should be made, let us list some of the activities of a library which may be improved by mechanization:

1. Handling the present contents of the library. This includes the operations of circulation, control of material on reserve, keeping an inventory of the book stock and maintaining it in good physical condition.

*Dr. Morse is Director, Operations Research Center, Massachusetts Institute of Technology. This article, in slightly modified form, constituted a portion of a report, with recommendations, on plans for the development of the M.I.T. library system.*

2. Bringing new material into the library. This includes accession of new or duplicate books; reception of serials such as periodicals, reports, and continuing publications; cataloging and other processing to ready the material for use.

3. Helping the user find the material he desires. This includes the activities of the reference librarian and the maintenance of various catalogs and book lists.

4. Making the material more available to the user. This includes the providing of microfilms and other copies and, eventually (perhaps) the transmission of desired excerpts over wire to distant users.

All of these activities are, of course, interconnected. Improvement of any part of one can help the others. Automation of one part must be planned to be compatible with possible later mechanization of other parts.

In all of these operations the need for mechanization depends on the size of the library. In a small library, not only are mechanical and electronic aids uneconomical, but they are also not (or at least not yet) as efficient as the librarian. As long as the collection is small enough for one trained person to keep control of the material and to maintain personal relations with the users, the standard library techniques are still effective and satisfactory. When the collection grows beyond a certain size or the users increase in number and range of interest beyond a certain

degree, there seems to be a sudden change in the character of the library and of its service. The larger mass of material makes it hard for the user to find what he wants and hard for the librarian to keep track of the material, and the larger number of users and their wider variety of interests decrease the personal contact between librarian and user. At some stage, the librarian finds that he no longer has control over his material and no longer has contact with the user, and the user soon loses the feeling that this is *his* library, and begins to misuse it.

This stage has been or is being reached in many larger university libraries. Today, such libraries are not able to be as much help to students and faculty as they were twenty years ago, and they are being treated with correspondingly less appreciation by their users. It appears unlikely that the trend could be reversed by splitting the larger collections into smaller departmental libraries. Most library users at present cover too wide a range of subjects to be satisfied with the coverage of one department. Many users would have to travel back and forth between several departmental libraries, or else each departmental library would have to broaden its coverage with consequent duplication of books, periodicals, and space. These library collections have passed the point where simple, personal control of material can suffice. More complex methods of operation and control, whether by machine or by people, are needed.

The experience of the science library at Massachusetts Institute of Technology is an illustration of this point, an illustration which is not at all unusual. Initial results, from a survey now in progress, seem to indicate that when a user comes to get a book or periodical which is listed in the catalog, in about one quarter of the "tries" he does not find the item where it is supposed to be. Part of the time, of course, the material is not there because another user has it or because the material is being bound or repaired. But in a nearly

equal number of cases the library is not sure where the material is; perhaps it is mis-shelved, perhaps it is not yet shelved, or perhaps it has disappeared, unrecorded, from the library. More than half of these "lost" books eventually turn up again, but meanwhile they are not available. Indications are that the "frustration factor" for this library—the mean fraction of times the user cannot find the material he desires—has roughly doubled in the past five years. If the present value turns out to be as large as one-quarter, then it is larger than it should be, and if it ever rises over one-half the library will rapidly lose its value as a repository of *accessible* reference material. Less detailed investigation indicates that a similar situation obtains in many other libraries.

The basic difficulty seems to be lack of control, in the engineering sense of the word "control." There seems to be no built-in feedback to the present system, no regular and frequent way by which the library staff can check its own operations to see whether the operation is actually running the way it is supposed to be running. At present, for instance, the usual way in which the average library learns that one of its books is "lost" is to be told by a user, who wished to use the book, that he could not find it; there is no regular procedure for the library staff itself to learn of the loss so it could be made good *before* the prospective user turned up. Such a feedback, resulting in a continual awareness of "how the system is doing," is a necessary part of any well managed operating system. In small libraries it can be maintained without procedural systematization, by the personal efforts of an intelligent and interested librarian. In larger libraries it can only be achieved by planned procedures of recording, systematic checking, and inventory. These procedures must soon be instituted if the growing collections of most university library systems are to continue their usefulness; already some of the pro-

cedures probably can be carried out more cheaply by machine than by person. The introduction of machine data processing techniques into the Class 1 type of operations (controlling the present contents of the library) is thus a necessary concomitant of the growing size and complexity of library operations and is now required if the larger university libraries are to continue as systems with adequate feedback controls. With appropriate planning, the techniques and equipment to be used can be made compatible with those other aspects of operation of a large university which must shortly be mechanized, so that the library will not need to defray more than its proportional part of the cost. For example, there are many parts of the university activities (student registration, infirmary records, some financial transactions) which could be simplified for mechanization by providing each student and faculty member with a machine-readable identification card. These cards could be used in any of several possible systems of library circulation control which are now available. The library can arrange its record-keeping and control procedures to share the same equipment used by other parts of the university administration.

These procedures should be designed to enable the library staff to know accurately and immediately "how the library is doing." Several systems have been worked out, for example, which enable overdue notices to be printed out automatically<sup>1</sup> and which can print out accession lists and other special lists of use to faculty as well as to library staff. Presumably any of these systems could be programed to provide shelf lists which would simplify frequent inventorying of the collections

<sup>1</sup> ALA, Library Technology Project. *Study of Circulation Control Systems* (LTP Publication no. 1 (Chicago: the author, 1961)), 133p.

Henry Birnbaum, *General Information Manual, IBM Circulation Control at Brooklyn College Library* (White Plains, N.Y.: International Business Machines Corporation, Data Processing Division, 1960), 32p.

R. S. Booser, "The Use of Data Processing Equipment for the Control and Circulation of Magazines," *Special Libraries*, LI (July-August 1960), 297.

so that missing items could be replaced before they are next needed (or their cards removed from the catalog if they are not to be replaced). And any of them could keep records of the use of individual books, which are at present seldom collected but which could be of importance in purchasing and discarding books. Whatever system is to be introduced would, of course, need to be tailored to fit the university's particular requirements, and should be accompanied by a thorough replanning of library procedures by a task force of machine experts and library staff which will take full advantage of the new data made available by the equipment.

For example, it might be possible to institute a planned program of retirement of less-used material from the more accessible parts of the library to more centralized (and less costly) stacks, thereby reducing the continual pressure to increase the size of library branches. It is likely that such a retirement program, by holding the number of books and periodicals on open shelves down to those items which are used more than once a year (for example), would result in an actual reduction of time spent by the user in finding the material he wants. But to implement such a system would require more available data on book use than most libraries can now easily collect.

It is therefore suggested that plans for mechanization of some of the operations of Class 1 should be first priority items for university library staffs. Such plans will require coordination with plans for mechanization of other university record-keeping and control operations. The increasing diversification of courses, schedules, and classrooms, for example, may require greater mechanization of registration procedures, room assignments, and class record controls than now exists. Standardization of identification cards, record cards, and processing equipment will materially reduce costs of the total changeover.

Once plans for mechanization of operations of Class 1 are under way, consideration of the operations of Class 2 (introduction of new material) could become active. Equipment and procedures for some of these operations are now being developed.<sup>2</sup> Data generated by the mechanization of Class 1 operations will provide useful and necessary inputs to Class 2 operations; for example, lists of "lost" books can be scanned to decide which should be reordered, and indications of high use factors can be utilized to order duplicates while the demand is still high. Much of the clerical work required in duplicating catalog cards and other listings could be carried out automatically if the original material were recorded on a punched card or magnetic tape rather than typed. Alternatively, the card could be microfilmed and copies made photographically; but this would not allow the sorting operations which can be made on punched cards or with taped data. Similarly, many of the control operations in ordering material and paying for it could be mechanized.<sup>3</sup>

Many of the basic actions of Class 2 operations, however, cannot yet be mechanized. Someone still has to decide which books to order, or reorder; a trained cataloger has to prepare the original card for the catalogs, and so on, though the work of these persons can be considerably lightened by the output of the mechanization of Class 1 operations and the techniques just mentioned. Eventually all large libraries should band together to prepare a standard catalog "card" in one place, to be used by all; but this must wait until a consensus appears regarding the form and format that this card should take.

<sup>2</sup> General Electric Company, Defense Systems Department. *A Final Report on Improving Information in a University Library* (prepared under contract with the University of Illinois Chicago Undergraduate Division (Washington, 1961)).

R. E. Durian and H. S. White, "Simultaneous Preparation of Library Catalogs for Manual and Machine Applications," *Special Libraries*, LII (May-June 1961), 231.

<sup>3</sup> General Electric Company, *op. cit.*

The prospects of mechanization of much of the operations of Classes 3 and 4 (helping the user find the material and making the material more available to the user) are still further off, though a few aspects show promise of early development. The basic difficulty seems to be the plain fact that information printed on a piece of paper is more convenient to read and use than is the same information projected optically or electronically on some screen. A catalog card can be carried over to a shelf to copy, a book can be scanned or studied anywhere and from any posture; any presently competing methods of displaying information are so handicapped in comparison that it would seem unwise to plan to depend on them until one is forced to. Thus arises the basic dilemma of present-day libraries, which will have to be solved in the next two decades.

This dilemma turns up in every aspect of operations 3 and 4. It is (so far) so much better to take out a copy of a bound periodical to study an article in it than it is to try to read a projected microfilm copy that many users refuse to use the latter. On the other hand, demand for such bound periodicals is so great that, even at present, they are not allowed out of many libraries and, because of the great demand (and probably because of the restrictions), these volumes are rapidly wearing out (and occasionally being willfully dismembered), so that even greater restrictions shortly will have to be placed on their use. A card catalog of a few score drawers, or even a few hundred, is comparatively easy to use, but what happens when the catalog grows to tens or hundreds of thousands of drawers? And while books are so useful, how does one store, control, and even find books if there are ten million of them?

It is thus not enough to devise an electronic catalog or to copy all periodical pages on microfilm. Until copies of the catalog entry or the periodical page can be made *on paper, quickly, and cheaply,*

the pressure for use (and even misuse) of the actual printed material and for continued extension of the card catalog, ad infinitum, will be well-nigh irresistible. Copying equipment is being used, of course, in many libraries, but it usually requires the book or periodical itself to make the copy, with consequent increased wear on the more popular items. It must be possible to devise a better, quicker, and cheaper copying process than this. When it arrives it will then be appropriate to plan the mechanization of some or all of the operations of Class 3 and 4.

By that time many other problems should be nearer solution than they are now. The very large number of questions which now are unanswered regarding the way journal articles (or even their abstracts) should be stored (electronically or microphotographically?), how they should be classified (what should be the descriptor language? should the author or a cataloger prepare the classification?) and the practical details of programming the operation of searching among millions of items, may at that time be answered *in toto* or in part.<sup>4</sup> It will then be easier to agree on the sort of information retrieval system the library should install, and to determine whether the university can afford to get it.<sup>5</sup>

Thus, for the less humdrum aspects of library mechanization, it is likely that

<sup>4</sup> A. H. Trotier, "Introduction to Mechanization in Libraries," *Library Trends*, V (October 1956), 191-92.

R. P. Bristol, *Closed-Circuit TV Equipment as Used in a Decentralized Library Situation* (Charlottesville: Alderman Library, University of Virginia, 1958), 131.

Avco Corporation, Crosley Division, *Technical Investigation of Elements of a Mechanized Library System: Final Report No. EW-6680*, January 11, 1960. (Boston: Avco Corporation, Crosley Division, Electronics Research Laboratories, 1960), 110p.

<sup>5</sup> "Mechanization of Service and Functions in the Library of Congress" (mimeographed report, March 24, 1960).

E. F. King, "Electronic Transmission for Inter-library Loans," *American Documentation*, XI (January 1960), 32.

there will be some decades of watchful waiting on the part of most university librarians before a decision can be made to mechanize the informational aspects of their libraries. There is a concomitant requirement, however, for libraries to take an active part in trying to find answers to the questions raised in the previous paragraph. Unless various methods can be tried out in practice as they are being developed technically, no real progress can be expected. But such experimentation is costly, since it cannot be tried out on a piecemeal basis. More funds from federal as well as private foundation sources will be required to build and experiment with the various proposed solutions. Some libraries will have to offer themselves (or some major portion of themselves) as guinea pigs. Only by such trial and error can we be in a position to decide quickly and accurately, when the time comes to do so, how Class 3 and 4 operations can be mechanized.

Meantime it will be necessary to start now to mechanize the less romantic parts of the library operation, those of Classes 1 and 2, if there is to be an operable library when the time does come for the storage and retrieval aspects of the operation to be mechanized. And, even more important, it will be necessary to learn what the library operation actually is, which books are being read and how often, what library users do and what they desire, in greater detail than is now known about most libraries. The librarian need not be a computer expert, but when the time comes to mechanize if he cannot specify quantitatively what the data processing equipment will be expected to do and in what operational environment it must work he will infallibly get the installation he deserves.

# Automation and the Library of Congress

FIVE YEARS AGO an internal committee of the Library of Congress was charged to study potential applications of electronic processing equipment to library procedures. . . . The library took the stand that it was not immediately concerned with mechanizing a few operations; it sought a plan to provide a blueprint for its actions during the next 5 or 10 years. . . .

On April 23, 1961, the Council on Library Resources, Inc., announced a \$100,000 grant . . . for "a survey of the possibilities of automating the organization, storage, and retrieval of information in a large research library . . . not only from the point of view of the functioning of an individual institution but also from that of a research library whose activities are interrelated with those of other research libraries."

To undertake this investigation the Librarian of Congress invited Gilbert W. King to head a survey team of technical specialists. . . . According to the grant from the Council, the survey was expected to result in a statement "of the feasibility of mechanization of research library activities and of requirements for such mechanization." The report obviously cannot do more than provide guidelines for consideration by the library administration. The team, in accordance with the stipulations of the grant, considered both current technology and foreseeable developments. . . . The survey team reached the following conclusions:

1. Automation can, within the next decade, augment and accelerate the services rendered by large research libraries and can have a profound effect upon their responsiveness to the needs of library users.
2. Automation of bibliographic processing, catalog searching, and document retrieval is technically and economically feasible in large research libraries.
3. The retrieval of the intellectual content of books by automatic methods is not now feasible for large collections, but progress in that direction will be advanced by effective automation of cataloging and indexing functions.
4. Automation will enhance the adaptability of libraries to changes in the national research environment and will facilitate the development of a national library system.
5. Automation will reduce the cost-to-performance ratio; however, the Library should aim at the expansion of services rather than the reduction of total operating costs.

. . . The immediate objective of automation will be to solve the pressing problems that face research libraries, among which are problems of bibliographic organization and control. In the long run, however, the most significant effect of automation will be the focusing of the services of the library on the individual user for the optimal satisfaction of his research needs. . . . —Section I, "The Automation Survey: Background and Conclusions." ■ ■