APPLICATIONS OF DATA PROCESSING AT THE CANADIAN NATIONAL RESEARCH COUNCIL LIBRARY

Jack E. Brown

In the world of libraries, the Canadian National Research Council Library is an unusual and perhaps unique institution, for it performs two closely related but often-times conflicting and incompatible roles. One of these roles is that of a science and technology library and documentation center serving a large group of scientists and engineers engaged in pure and applied research in many areas of science and technology. The other role is that of a National Science Library serving the entire scientific community of Canada.

The NRC Library, as it exists today, consists of a main library which houses the bulk of the Library's half-million volumes, and six smaller and more specialized collections serving several divisions of the National Research Council (NRC) which are located four miles from the main building. The main library acts as the nerve center of the Library system with administrative services, acquisitions, cataloging, classifying, and binding centralized at this point. The branch libraries operate primarily as working collections which, in most instances, duplicate parts of the main library's holdings. By means of close cooperation between the various library units, the maintenance of a union catalog at the main library, good telephone communication, and the use of a station wagon which shuttles back and forth several times a day between the main and branch libraries, the entire resources of the system are coordinated for ready access. The Library has a staff of seventy-eight, twenty-seven of whom are professional, and an acquisitions budget of $200,000.

For purposes of this meeting, it is unnecessary to describe in detail the resources and services provided by the Library—suffice to say the NRC Library is much more than simply a repository for the world's output of scientific and technical literature. It is a dynamic organization which utilizes every means at its disposal to provide the Council's scientific and engineering staff with the publications and

---

Jack E. Brown is Librarian of the Canadian National Research Council Library, National Science Library of Canada, Ottawa, Canada.
information required in their day-to-day work. These same resources and services are extended to scientists and engineers anywhere in Canada, by means of interlibrary loans, through the provision of photocopies, and by means of a Science Information Service geared to compile bibliographies, carry out literature searches, and answer requests for scientific and technical information.

The NRC Library, as with most other scientific and technical libraries worthy of the name, is endeavoring to keep abreast of the latest developments in the field of mechanized storage and retrieval of information. Key members of the staff are encouraged to attend pertinent training courses and seminars, and one member of the staff whose formal training embraces chemical engineering and mechanized systems of documentation, has been designated Library Systems Analyst. His specific assignment is to determine, in collaboration with the librarians, those operations which can be made to function more effectively through the use of automatic data processing equipment.

During the past four years, the NRC Library has been experimenting with the use of electronic equipment to solve specific problems. The scale of experimentation is indeed modest as compared with similar activities being conducted in many United States libraries. However, we must learn to walk before we can run, and attention has been concentrated on the improvement of those essential operations which, because of sheer volume of work involved, were failing to achieve their objectives.

At present, automatic data processing equipment is being used successfully in the following operations:

1. Preparation of a list of subject headings for use in one of the branch libraries, with revised editions to be issued at regular intervals.
2. Preparation of complete lists of serials held by the NRC Library, and issued annually.
3. Preparation of periodic and cumulated lists of NRC publications, together with author and Keyword-In-Context (KWIC) subject indexes.

As I discuss these three operations, I trust you will keep in mind that the work was carried out within the limits set by existing staff and budget. No additional allotment was provided or extra staff hired. A key punch machine (IBM-26) was acquired by the Library, but all other machines required in the operation were available either at the National Research Council or at other government departments in Ottawa.

Our first attempt to employ electronic equipment in a Library operation was in the preparation of a printed list of subject headings covering the fields of aeronautics and mechanical engineering.
Lacking experienced guidance, we made many mistakes before the system was operating to our satisfaction.

Three years ago the Library started work on a revised list of subject headings used in indexing technical reports received by the Aeronautical and Mechanical Engineering Branch Library. The headings and sub-headings were typed on 3"x 5" cards, but the ultimate aim was to prepare a list of headings which could be revised and reissued as often as was necessary and with a minimum of repetitive work. Various techniques for obtaining lists from the cards were evaluated, and it was decided that punched paper tape would best meet our needs. This decision was in some measure influenced by the accessibility of three Flexo-writer machines.

In theory the method should have worked; in actual practice it created more problems than it solved, due largely to our inexperience and inadequate guidance. The tapes, when consolidated, proved to be incompatible, and single letters, parts of words and whole words failed to appear in the printed list. These errors were, of course, quickly discovered and the printing operation halted. Furthermore, we found the updating of the tape, to incorporate new headings, to be a cumbersome and frustrating task.

At this point the Ottawa office of IBM became interested in our problem and offered their assistance. They suggested converting the paper tape directly to magnetic tape from which a printout could be obtained through the use of the IBM 1401 computer. Since it was the first time the IBM office had tackled such a project, they offered to do the work for a nominal sum and we agreed. Many programming difficulties were encountered during the various steps in the conversion from punched tape to computer printout, but eventually they were overcome and the final results justified the adoption of IBM equipment.

New subject headings and corrections are prepared on punched cards and the magnetic tape updated at regular intervals. The preparation of new editions of the list of headings is now a relatively simple and inexpensive operation. The chief cost lies in the multilithing of sufficient copies for distribution to other interested libraries.

The second project, the preparation of a complete list of serials held by the NRC Library, has been described in detail in an article in Canadian Library. For this reason, and because similar procedures are used in several U.S. libraries, I shall limit my discussion to the main features of the project.

Because of the nature of scientific publishing, periodicals and other serials constitute the major portion of the NRC Library's collection and account for approximately 80 per cent of its total holdings. At the present time, the Library receives more than 10,000 different serial titles. The preparation and publication of up-to-date lists of these serials is a formidable task, but one which must be continued if Canadian scientists are to be made fully aware of the material available to them.
Until 1958 a complete record of periodical titles and holdings was published in book form at three-year intervals. The lists were placed at strategic points throughout the main library and its branches, and in the offices of the various divisions. Later, as the national responsibilities of the NRC Library expanded, copies of the lists were sent free of charge to university libraries and, for a nominal sum, to other interested organizations.

The size of the periodical collection has now reached the point where, with the staff available and by conventional methods, it is no longer possible to issue up-to-date lists at reasonable intervals. The Library examined various mechanized systems to determine which of these, if any, could be used to solve this dilemma and, in June 1963, embarked on a system using IBM punched cards and related automatic data processing equipment. During the preliminary stages of development, it was found that, with a little more effort, it was possible to assign codes to each title which would facilitate the preparation of lists of selected titles on the basis of subject, country of origin, language, subscription agent, and other categories.

The layout of the IBM card is as follows:

Sort groups.—In order to maintain an alphabetical arrangement of titles, and to permit resorting of the file, each card or set of cards is assigned a number. This number sequence (columns 2-6) allows for a listing of 99,999 titles. As new periodical titles are received, additional numbers are assigned to columns 7-8. This allows for the insertion of 99 titles between any two existing titles and in alphabetical order. Thus, up to ten million titles may be listed in the alphabetical-numerical sequence. Since more than one IBM card is required to describe the title, a numerical code in columns 9-10 ensures the proper sequence of cards within a set.

Text.—Columns 11-18 contain the LC classification and Cutter number. Columns 27-80 record the title of the periodical, the holdings of the main library and its six branches, and any necessary notes or "see" references. The information punched in these columns determines the number of cards required for a given title—usually four to five cards per title.

Subscription agent.—The majority of the NRC Library's periodical subscriptions are handled by eleven agents located in various parts of the world. Here, in column 19, an alphabetical-numerical code has been assigned to each agent, leaving fifteen additional letters available for use at a later date.

Subscriptions.—A numerical code, in column 20, is used to indicate whether a journal is received as a paid standing order, as a paid subscription renewed each year, received at irregular intervals, or received on exchange, as a gift, or through membership in a society. A numerical code, in column 24, records the expiration date of each paid subscription by month.
Language.—The twenty-one major languages in which the periodicals received by the Library are printed are indicated by an alpha-numerical code in column 22. International journals which contain papers published in a wide variety of languages are not coded.

Translations, abstracting and indexing services, bound volumes—a yes/no number code, in column 23, records those journals which are English or French translations of journals published originally in some other language, abstracting or indexing services, and journals which are bound on a current basis.

Holdings.—A numerical code, in column 24, records titles held by the main library and/or any of the Library's six branches. The code also indicates duplicate sets held in the reserve collection at the main library.

Country.—An alpha-numerical code, in column 25, indicates the country in which a journal is published. A combination of the information recorded here and in column 22 permits the preparation of lists of journals published, for example, in Russia but printed in another language.

Ideally, all information to be coded should be keypunched in one operation. Because of staff shortages this was not possible, and columns 11-18 and 27-80 were punched first to record the LC classification and Cutter number, the title of the serial, the holdings, and so on.

At this stage, the resulting cards were run through an IBM 407 to obtain a complete printout of all titles and holdings. The first printout was done on 11" x 12-1/2" sheets, with space left at the lefthand column for the insertion of coding symbols. Sets of sheets from this printout were distributed to selected members of the staff for proofreading and the assigning of appropriate codes. Errors or omissions were noted on the work sheets and the sheets forwarded to the keypunchers for the preparation of corrected cards.

Upon completion of proofreading and keypunching of codes and corrections, the complete set of cards was ready for preparing the final and master list of serials, again by the use of the IBM 407. The printout was done on 15" x 18" sheets which were then reduced by Xerox camera to 8-1/2" x 11" duplimat plates. Duplication was carried out by means of multilith machines. If a reduction is not required, the master copy, of course, can be printed directly on to duplimat paper.

New titles, changes of title, or changes in holdings, together with the pertinent coding, are recorded on specially designed 3" x 5" cards. The layout of these cards enables the key punch operator to prepare the IBM cards without further instruction from the librarian.

Supplementary lists of new serial holdings, for internal use, are run off at four-month intervals. The new cards are then incorporated into the master file preparatory to the printing of a revised and complete list of serials.
I have indicated earlier that this mechanized system enables the Library to meet requests for lists of selected journals which hitherto could not be satisfied at reasonable costs by conventional methods. For example, the Library has prepared such lists as journals held by each branch library, mathematical journals (or any other subject), Russian language journals, journals published in China, abstracting and indexing services covering all subjects or selected subjects, journals received on exchange or as gifts, and journals whose subscriptions are handled by a specific agent.

The third and most recent application of data processing equipment in the NRC Library has been to prepare a list of all papers published by NRC personnel, and other publications issued by the National Research Council. The procedures used are similar to those described above and can be dealt with in less detail.

The NRC, as a publisher of scientific and technical information, is best known for its seven Canadian journals of research: Canadian Journal of Physics, Canadian Journal of Chemistry, Canadian Journal of Biochemistry, Canadian Journal of Botany, Canadian Journal of Physiology and Pharmacology, Canadian Journal of Zoology, and Canadian Journal of Microbiology. It is also the publisher of many separate monographs and scientific series, but the majority of the reports written by the Council's scientific staff are published as papers in international scientific and technical journals.

It has been a relatively easy matter for the Library to issue periodic lists of NRC publications. On the other hand, the preparation of cumulated lists of more than 8,000 publications has become a task with which the Library could not cope by conventional methods. The success of the serials operation prompted the adoption of these same techniques to solve this new problem. Because of the descriptive nature of the titles of the papers, it was further decided to compile a KWIC index.

The preparation of IBM cards recording all bibliographical information and NRC numbers for some 3,000 papers published between 1958 (the date of the last cumulation) and December 1963, was completed in three weeks. A preliminary bibliography for proof-reading purposes was run off on an IBM tabulating machine, and a list of 250 non-significant words keypunched. The latter cards and the title cards were turned over to the IBM Data Processing Center and the KWIC index prepared. Punched cards are now being prepared for all papers published between 1918 and 1958, and a complete list of NRC publications will be issued.

Once again, it is worth noting that the preparation of the master file of IBM cards offered no saving in time or money as compared with the typing of cards or lists. However, with the completion of the master file, it requires only a few hours to prepare supplementary lists of NRC publications or complete cumulations, each with author
indexes and KWIC indexes. The compilation of lists of papers by author or by publishing journal is also a simple matter.

In conclusion I must emphasize that, in each of these projects, our aim was not to reduce the costs of existing operations; rather, we were seeking new procedures which would enable the Library to provide several needed services which could not be performed by conventional methods and without an increase in staff or budget. The fact that, through the use of automatic data processing equipment, we were able to provide these services at no additional costs, was all the encouragement we needed to extend our experiments to such areas as circulation control, acquisitions, and the preparation of printed catalogs and accession lists.

The Library staff has become conversant with the techniques and possibilities of data processing machines, and looks forward to the time when mechanized systems of information storage and retrieval will become a reality. We have learned to walk, and hope that we may soon be able to run.

REFERENCES