academics, local historians, church members, genealogists and journalists. Linguists and sociologists have found the records of the Diocese of the Arctic and the papers, photographs, and translations of a number of prominent missionaries including Edmund Peck, William Bompas, Isaac Stringer and Archibald Fleming to be a valuable primary source.

Many challenges remain for the GSA beyond caring for the materials in its custody and contributing to the improvement of Anglican archives in Canada. These include encouraging and assisting other denominational archives, co-operating with Anglican archives in other countries, collaborating with secular archives as they expand the horizons of archival service, and the winning of acceptance and respect for this sort of non-government endeavour.

Marion Beyea
General Synod Archives
Anglican Church of Canada

FOCUS: The Machine Readable Archives Divison of the Public Archives of Canada

In 1973 the Public Archives of Canada (PAC) established a new division, the Machine Readable Archives Division, to provide archival services for computer-generated records. Although not the first national repository to establish such a programme—its counterparts in the United States and Sweden having already done so—the PAC was the first archival repository in Canada to tackle machine readable records. This was in keeping with the Public Archives Act which provides for the acquisition of all historically significant archival material "of every kind, nature and description."[1] Furthermore, in the federal government's Public Records Order, the definition of "public record" includes "tapes, computer cards, or other documentary material, regardless of physical form or characteristics. . . ."[2] And finally, it was in keeping with the "total archives" approach of the Public Archives of Canada.

The Machine Readable Archives (MRA) has as its mandate the appraisal, acquisition, preservation, and reference service for machine readable records of historical and long-term value created by the federal government and those of national significance created in the private sector. To carry out this mandate, the Division has nineteen employees, including data archivists, computer science personnel, and clerical staff. It is physically located in the Larivière Building in Hull, a considerable distance from the main PAC building, resulting in many of the disadvantages and inconveniences which such remoteness implies.

A new archival medium required fresh approaches, procedures, and policies. Under the direction of Michael E. Carroll, the staff of the Division devoted considerable time during its first three years to developing proper methods and standards. Some similarities and differences which are apparent when this medium is compared with more familiar archival media can best be seen by outlining the Division's main functions: appraisal and acquisition, processing, conservation, and public service.

1 Public Archives Act, R.S., c. 222, s. 1.
2 Public Records Order, P.C. 1966-1749.

APPRAISAL AND ACQUISITION

Over the past few years, because of limitations of human and financial resources, the MRA has concentrated almost exclusively on the acquisition of machine readable data created by the federal government. Yet, even in the public sector, there are vast amounts of information in machine readable form covering such varied aspects of national life as employment, crime, disease, immigration, emigration, climate, geology, food production and consumption, housing, transportation, communication, and the cost of living. In fact, full Electronic Data Processing (EDP) costs in the federal government for 1976-77 were identified as $283 million, with projections for 1977-78 of $323 million and for 1978-79 of $353 million.

When all the systems have been set up, the procedures for the acquisition of machine readable records will follow in many ways those used in the acquisition of paper records. These procedures will form part of the federal government's EDP records management programme as outlined in chapter nine of the government's *Guide on EDP Administration*.[3] As part of that programme, the Office of Records Management Services of the Public Archives of Canada undertook in 1976-77 an inventory of machine readable records in some sixty-seven government departments. These departments have now been asked to schedule their machine readable files (collections of related data which have been coded and prepared for computer use) and to submit these schedules to the Archives for approval. It is through the operation of these schedules that MRA will acquire most of its records in the future.

Several problems have arisen in connection with the scheduling of machine readable records. First, many departmental records managers do not have training in the EDP field and are therefore unfamiliar with the medium. Second, many do not have control over machine redable records in their departments since the creators and users of the records in many departments work independently. The latter are somewhat suspicious of our intentions, knowing very little about the PAC and even less about MRA. To overcome some of these difficulties, commencing in April, 1978 the data archivists will go into several departments and actually assist in the scheduling of machine redable records.

The majority of MRA's holdings, however, have been acquired not as a result of the scheduling procedures of the EDP records management programme, but rather as a result of archivists contacting departments directly through our ad hoc acquisition programme. An additional inventory or survey undertaken in 1975-76 by a special secretariat of Treasury Board has assisted us in identifying and acquiring machine readable records in government departments and agencies. With this information, MRA has been able to acquire nearly three hundred files.

With paper records it is common practice to wait for the expiry of the retention period on the schedule before acquiring the records, a practice avoided, if at all possible, with machine readable records. Magnetic tape is a very vulnerable medium, easily affected by strong climatic changes and mishandling. Moreover, it is mandatory to locate as soon as possible the supporting documentation necessary to use the data, for quite frequently the archivists have had to prepare documentation manuals using information received orally from data users. The longer one waits to acquire a machine readable file, with its related documentation, the greater is the possibility of the information being lost. However, departmental restrictions are always maintained when documentation is being sought. Because adequate documentation is so important, staff

3 *Guide on EDP Administration for Departments and Agencies of the Government of Canada* (Ottawa, 1974).

of the Division have been working on a sub-committee of the Government EDP Standards Committee to develop proper documentation standards for implementation throughout the federal government.

As with textual records, appraising machine readable records for their historical and long-term value is no simple task. Such factors as the nature and reliability of the data, the completeness of the documentation, the size of the file, the estimated processing and preservation costs, as well as the complexity and format of the file are all considered. Although divisional guidelines have been established to aid archivists in the selection process, a degree of subjectivity still comes into play, for an archivist's educational background and work experience in certain subject areas will undoubtedly influence this decision. Like other divisions within the Archives Branch, MRA is beginning to hire archivists with specialized education and training for specific subject areas. While several MRA staff members have history degrees with emphasis on quantitative methodology, most have received their educational training in political science, economics, sociology, and other social science disciplines which make use of the computer to undertake quantitative analysis.

PROCESSING

It is in this area that the work of the data archivist differs most dramatically from that of the traditional archivist; here his or her specialized background is most noticeable. During the processing stage, the acquired records, whether punched cards, discs, or another computer format, are converted to reels of magnetic tape, the storage medium used for all material in the Division. One file can comprise any number of reels of tape, although all are in the standard format of nine track, sixteen hundred bytes per inch, Extended Binary Coded Decimal Interchange Code (EBCDIC).

In processing a file, the data archivist can spend considerable time verifying the background documentation with the data on the file. The archivist must ensure that a file is exactly as the documentation describes it, that is, that the number of variables and cases, code types, and file structure or layout are as indicated. Verification quite often reveals that the documentation is inadequate, requiring a search for further information which may exist only in someone's head, or is widely dispersed. In some cases, it is even necessary to refer to the original questionnaires or interview schedules if they exist. The verification is often done using certain statistical packages, which are systems of computer programmes designed for analysis of data, with which the archivists are familiar, the most frequently used being SPSS (Statistical Package for the Social Sciences). Occasionally, special computer programmes have to be written for particularly difficult or complicated files. Errors found in the files are noted in the documentation, but usually not corrected.

A fully processed file now consists of two archival copies made on certified tape as well as a fully documented codebook or manual which the archivist prepares according to a standard format. The codebook contains the title of the file and the name(s) of the principal investigator(s), the source of the file, the date(s), a description of the content of the file, the methodology used in creating the file, information about the processing of the file which would be useful to the researcher, and so on. It will include a copy of any published or written report prepared by the department as a result of the creation of the file, as well as transfer and access authorization forms. Because the PAC does not have its own computer facility, all of the computer time required to process a machine readable file has to be rented from commercial service bureaus, an expensive undertaking for which the Division expends the biggest portion of its non-salary budget.

CONSERVATION

While conservation is of the utmost importance to the Machine Readable Archives, it is conserving information rather than physical documents that is significant. MRA's storage medium, magnetic computer tape, is very compact and can be easily and inexpensively copied with no loss of information. Since it is vulnerable to destruction if handled or stored improperly, a number of measures have been taken to minimize any loss of data. Two copies of each file are made on high quality magnetic tape that has been properly certified. These copies are stored in separate physical locations, one on the premises and the other outside Ottawa. A proposal to have a further backup copy on Computer Output Microfilm (COM) is being considered. As microfilm is a fairly stable medium, this would allow re-creation of a file if both tape copies were lost.

MRA's tape library is environmentally controlled and dust free, with a temperature of 21°C± 2°C and a humidity level of 50 percent ±10 percent. Divisional staff are currently working with the Government EDP Standards Committee in developing standards for the care, handling, and storage of magnetic tape. It is hoped that these standards will be implemented eventually throughout the federal government. Certain tape library procedures have been established which should ensure data retention. In order to maintain the pressure within the tape mass to a constant eight ounce per square inch tension, each tape is precision rewound annually by using the Division's two tape cleaner/precision rewinder machines. Because of the natural deterioration of the medium, one copy of each file is recopied every five years. However, the frequency of both precision rewinding and recopying will depend upon the results of quality checks.

An important part of MRA's conservation programme is concerned with keeping up to date with the latest developments in, and finding a more suitable medium for, the long-term storage of machine readable data. Microfilm, particularly the Computer Input Microfilm (CIM) process, is one alternative, although it is currently very expensive. Another is the use of a laser beam to record digital information on discs. The next few years could bring about some very significant developments in this area.

PUBLIC SERVICE

The differences in educational background between most MRA and other archivists at the Public Archives of Canada is reflected in the clientele which the Division serves. The majority of our researchers are social scientists, the numbers being somewhat restricted because they require the use of a computer installation. A further distinction is that most of our inquiries are by mail or telephone rather than by personal visit.

The Division provides three basic services at reasonable costs: data copy, data extracts, and data analysis. Data copy includes providing the researcher with an exact copy of the requested file in addition to the necessary supporting documentation; data extracts are an exact copy of a portion or extract of a file; and data analysis involves generating statistical runs from a data file.

In providing public service, the question of confidentiality frequently arises. Some files which have had no restriction placed on them by the donor contain data by which respondents can be identified. Although MRA's policy is to provide researchers with as many files as possible, there is also a clear obligation to protect the identity of respondents. This will be accomplished by a data access archivist whose duties will consist of examining all files before they are released for potential disclosure of respondent identities and editing anything necessary to eliminate such a risk. This could be done either by deleting all personal identifiers from the record or by performing certain analyses desired by the researcher and providing only summary data.

In the important area of public service, MRA has discovered a close relationship with the library profession. The cataloguing rules and formats for machine readable files, developed by a sub-committee of the American Library Association, have now become part of the Anglo-American Cataloguing Rules and it is according to these rules that MRA's files are catalogued. This should make the referencing of a data file as easy as that of a book, and may also permit the eventual integration of MRA's holdings in a descriptive manner with other media in the Archives Branch.

The documentation manuals or codebooks are additional research aids. These are currently being microfilmed and microfiche copies will be provided to interested researchers. In the future, it is possible that a copy of the catalogue card and a microfiche copy of the codebook will be deposited in major archives and libraries across the country.

During the 1978-79 fiscal year, the Machine Readable Archives will continue to list its accessible files in the *Social Science Data Inventory* published by the Data Clearing House for the Social Sciences. In addition, the first of annual holdings catalogues will appear. And, finally, a manual will be published to provide an overview of the functions and operations undertaken in the Machine Readable Archives.

It is hoped that the procedures developed by the Machine Readable Archives will be of some assistance to other Canadian archival repositories in their acquisition and preservation of machine readable records of historical and long-term value, for it is really only a matter of time before other repositories begin collecting computer-generated records, if they are not already doing so, since more and more information is being placed in some machine readable format by public and private governing bodies, organizations and institutions. The establishment of MRA is a further example of archivists recognizing their responsibilities through the acquisition and preservation of records which, though they may be different and unfamiliar, nonetheless document many significant aspects of Canada and its people.

Harold Naugler
Machine Readable Archives
Public Archives of Canada