

Electronic Records and Archival Description

by CATHERINE BAILEY*

Unlike other associations in the information management field, such as the Society of American Archivists (SAA), the Association of Records Managers and Administrators (ARMA) and the International Council on Archives (ICA), the Association of Canadian Archivists (ACA) has not spent much time, since its creation in 1975, in studying "the issues relating to the processing, management, and acquisition of electronic records."¹ Furthermore, the Association was not able to respond positively to overtures made a few years ago by the Archives Committee of the Canadian Historical Association (CHA) and its Committee on Computing in History, regarding a joint ACA-CHA policy on the conservation of electronic records. Nor was the ACA able to answer requests by archivists and institutions for material relating to "standards to govern electronic information storage media; standards to govern electronic exchange of information; [or] policies and/or procedures concerning the management of electronic information."² On the verge of the 1990s, the Association had still not addressed in any depth the electronic records issues of the 1970s and the 1980s.

For that reason, in 1990, the Executive established a Select Committee on Electronic Records, which was "charged with the task of studying the issues related to the management of electronic records and advising the Executive on what sort of permanent body (Standing Committee or Special Interest Section) would be most appropriate, as well as on the agenda that such a body should pursue."³ After a slow start, the Committee was reconstituted following the Annual Conference in May 1991. The mandate and goals of this revamped Select Committee however, have not been altered. It is still charged with conceptualizing the issues relating to electronic records and offering recommendations to the Executive on the kind of permanent body best able to deal with these issues on a long-term basis, as well as suggesting to that body those areas of concern which are worthy of further study. It will not, in this first exploratory year, be offering solutions to key problems or developing usable standards of any kind.⁴

For this special issue of *Archivaria* devoted to archival description, a few preliminary observations may be appropriate as to the impact of electronic records on description. One of the most important issues is the need to determine the basic form and purpose of electronic records description. In his review article, "Easy to Byte, Harder to Chew:

The Second Generation of Electronic Records Archives," Terry Cook argues that early data archivists, isolated from their "regular" or "traditional" archivist colleagues, turned to associated professions such as statisticians, sociologists and librarians for assistance in developing models for description. Since most of those users were less concerned with the context of computerized information, what emerged was a system of archival description of machine-readable datafiles which was essentially a library cataloguing system.⁵

There are two points which may be raised regarding the use of library-style cataloguing for electronic records. First, is it at all necessary to describe the physical characteristics of the "original" archival electronic record, as is done for maps, drawings, plans, or audio-visual materials, when electronic records are routinely copied and reformatted onto different media, often at the instigation of the archives acquiring or later conserving the records? More importantly, should archivists continue to describe electronic records using this kind of system? While librarians and archivists both focus on the informational content of the "item" being described, library description usually proceeds within the limitations of that single, isolated unit (a book, a map or a single datafile). Archivists, on the other hand, derive a significant portion of their content (and, of course, context) description from the organic evolution of interrelated, similar items into a conceptual whole. A library descriptive system ignores both the context of creation of the records and their relationship(s) to other records, and thus such a system can never describe archival material *fully* without being significantly revised.

In recent years, there has been an increased awareness in the archival community of the need for a system of description which reflects the provenance-based, interrelated nature of archival records regardless of their medium. The Planning Committee on Descriptive Standards of the Bureau of Canadian Archivists has recently established the Computer Files Working Group to draft the chapter of *Rules for Archival Description (RAD)* relating to electronic records. It is hoped that this new chapter will provide a set of rules that will encourage the proper contextual description of electronic records within the entire multi-media record-keeping system of their creator, and not simply entrench the library cataloguing rules of the past.

There are also other, more opened-ended issues relating to electronic records description. Several examples were suggested through a recent lecture at the National Archives of Canada given by Dr. Mark Olsen of the Center for Information and Language Studies, University of Chicago, on the uses to which some scholars are currently putting electronic records. Dr. Olsen claimed that there have been three kinds of electronic records usable as sources for researchers in numerous disciplines: conventional "flat files," databases, and textual information. He stressed that today's more sophisticated researchers, familiar with computers and computer files are no longer content with analysing and re-analysing single "flat" datafiles, which comprise the majority of archival holdings of electronic records to date, but are becoming more interested in relational data, in hypermedia combining audio, visual and textual material, and in textual electronic records as "texts" to be subjected to word, phrase and meaning analysis. Textual materials formerly available only on paper are also being encoded into vast databases, and then made available worldwide through wide-area networks. Dr. Olsen predicted that in the foreseeable future, researchers will be greatly interested in such formerly non-archival records as electronic mail, electronic bulletin boards and the documents in office automation systems. He further speculated that researchers such as diplomatic historians may even

want to see the word-processing codes for office automation systems, in order to determine the nuances of key documents.

This is not futuristic, but a present reality at the National Archives of Canada. In 1989, the Archives acquired a number of electronic systems from the Trade Negotiations Office (TNO), originating from word-processors and local area networks which were used to produce, transmit and store documents used during the Canada-United States free trade negotiations. Researchers are now contacting the Archives with such requests as "all documents in which the subject of water was mentioned in the course of free trade negotiations." A free-text search for the relevant word(s) carried out on these electronic records using sophisticated "text analysis" software packages reveals a series of "hits," that is, a listing of the documents which mention the requested subject. Unfortunately, restrictions on the material under the *Access to Information* and *Privacy Acts* and the lack of a proper researcher interface which respects those restrictions, means that for now the responsible archivist is the only person who can carry out a search on behalf of the researcher. If individual documents can be retrieved using a software "overlay" on the original data, moreover, what role, if any, is there for additional description, let alone standardized description, by archivists? It seems clear that the levels of description, and the consistency of their application, will vary greatly among fonds, series, files and items, especially for electronic records.

Archivists have already spent a great deal of time and energy deciding how best to deal with the appraisal of electronic records, but little thought has been given as to how the material which is acquired is to be described. How should the TNO records be described in finding aids? Does a simple description of the name and nature of each of the pieces of supporting software and a logical data model suffice? Should the responsible archivist explain the difficulties researchers will encounter in undertaking to search the records themselves? Should they show them how to search? Should there be any mention of the initial processing and copying of the records necessary in order to make them accessible? How can the links between the textual documents and the office automation system records from this Office, be made when most archival description is now conceived in media-specific ways and formats? These are just a few of the descriptive issues raised by the use of office automation systems that will need to be addressed in the near future as archives acquire more and more such records.

Dr. Olsen also remarked that there has been and always will be a degree of mediation between the archivist and the researcher when it comes to computer records; as the TNO case shows, this is true, if only for the reason that the technology needed to search the records, while respecting the provisions of the *Access to Information* and *Privacy* legislation, cannot allow direct researcher access to the data. Olsen also suggested, however, that this mediation will become more and more substantial as time goes on, that archivists will become an *essential* intermediary between the user and the records. Given the sophistication of the software needed to access more complex electronic records, and the greater complexity of related system documentation, should archivists aim their descriptive tools more specifically to assist particular clientèle, such as those few highly computer-literate and specialized researchers of whom Dr. Olsen speaks? Or should they simply treat electronic records the same way they have been treating textual documents for years, as archival fonds requiring different depths and levels of description according to the complexity of the records, not of the anticipated use of them?⁶ Is description driven by researcher's needs or by the medium of the records? What level of complexity in the

electronic record must be reached before archivists increase the amount, or change or customize the description presented to researchers? If an electronic record acquired by the archives is subject to long-term access restrictions, should the archives develop a two-tiered hierarchy of description: one for public use and one for the staff use, in order to ensure that vital descriptive information necessary for future handling of the material is not lost?

If archivists are to maintain their role as collectors and guardians of society's documentary heritage in a world which increasingly relies upon computers and electronic records for its daily activities, questions such as these, relating to archival description, must be discussed in both theoretical **and** practical terms. The boundaries of the discussion need only be limited by technological developments and our own imagination and capacity for learning.

Notes

* I would like to thank several people for their help on this piece. Cynthia Lovering, who is a member of the Bureau of Canadian Archivists' Computer Files Working Group, and I discussed and revised many points. Paul Marsden and Rick Brown provided information on the two specific examples of electronic records that I used (TNO in the text and CCARCS in note #6, respectively); Paul also commented on an early draft. Terry Cook offered constant encouragement and suggestions. The opinions expressed herein do not represent those of the National Archives of Canada.

1 Letter, Jay Atherton to Catherine Bailey, 28 December 1989, p. 1.

2 Ibid., p. 1.

3 Ibid., p. 2.

4 For a brief overview of the new Select Committee on Electronic Records, see "New ACA Initiative for Electronic Records," *ACA Bulletin*, 16, no. 2 (November 1991), pp. 4-5.

5 Terry Cook, "Easy To Byte, Harder to Chew: The Second Generation of Electronic Records Archives," *Archivaria* 33 (Winter 1991-92), pp. 202-16.

6 An excellent example of this occurs in the Canadian Civil Aviation Registration Computer System (CCARCS), found in the holdings of the National Archives Government Archives Division Record Group 12 (Records of Transport Canada). Acquired in 1990, this database is a mainframe computer application which supports the registration, inspection and operation of civil aircraft in Canada, and replaces an extensive paper civil aircraft registry. At the present time, it is estimated that there are between 35,000 and 42,000 civil aircraft registered in Canada. The National Archives holds a copy of the master data file and receives updates on an annual basis. This acquisition is significant because it is the first electronic data file acquisition to be treated as an accession unit and integral part of the record group inventory, not as a datafile to be described as a discrete item. It has led to the development of integrated accessioning procedures for electronic and textual records, and may well be one of the forerunners of an integrated system of description for *all* government records at the National Archives. A more extensive description of CCARCS is found in *Machine Readable Records Bulletin* 7, no. 1 (1990).