

## The Electronic Records Archival Programme at the National Archives of Canada: Evolution and Critical Factors of Success

by Terry Cook and Eldon Frost

At the beginning of the 1970's, a few leading archives around the world began to respond to the changing nature of information signalled by the widespread introduction of computers into government and business after the Second World War. The Public (now National) Archives of Canada was one of the first to do so. In 1974, the Machine Readable Archives Division (MRA) was established to address this newest archival medium. The mission of the Division was to acquire the archival electronic records of Canadian federal government institutions, as well as those of national significance from the private sector. From the start, however, the primary focus was on government records (a seven-to-one ratio), as was befitting the National Archives' role as the sole archives of the Government of Canada.

This article will explore the evolution of the electronic archives programme at the National Archives of Canada, focusing especially on those factors which helped and hindered its early progress, on the critical changes implemented over the past few years in order to meet the realities of the 1990's, and on the challenges ahead.<sup>1</sup>

For these pioneering archivists in MRA, there were no archival models to follow and often little understanding of their endeavours by "regular" or "traditional" archivists. Thus isolated, the "first generation" of electronic records archivists turned to others using computerized records for advice and inspiration: statisticians, sociologists, other social scientists, and librarians. Not coincidentally, the computer records which first drew the pioneers' attention from the era of mainframes and punch cards in the 1950's and 1960's were the same statistical or survey files used or collected by these allied professions. An equally important limiting factor was the state of information technology at the time. Aside from social science datafiles, the only business applications being automated were administrative, such as payroll, inventory, shipping, receiving, accounts receivable, and so on, and these had little or no archival value. Information technology, other professional alliances, and an undefined archival framework for automated records combined to set the focus for the

---

<sup>1</sup> This article has its origins in and is a blend of two earlier pieces: Terry Cook's essay prepared for the CART meeting, which in large part extracted, with some revision and expansion, a portion of his "Easy to Byte, Harder to Chew: The Second Generation of Electronic Records Archives," *Archivaria* 33 (Winter 1991-92), pp.202-16; and Eldon Frost's "Strategies for Acquiring and Accessing Electronic Archival Records: The National Archives Perspective," a paper delivered to the annual conference of the Canadian Historical Association, Ottawa, June 1993.

pioneering generation. These circumstances in turn had numerous implications for the development of the first generation of archival practice in electronic records.

Statistical and survey files were primarily forms or questionnaires which, for tabulation ease, had been made "machine-readable." Thus the information they contained was central to the sociologists and statisticians using and often creating them, while the context surrounding their creation was of secondary importance. Similarly, for the first generation of pioneering electronic records archivists, "informational" value was emphasized in their theoretical commentaries and appraisal practices, whereas contextual "evidential" value was less important. Moreover, other electronic records more central to the business of the creating organizations were difficult to appraise and acquire, for unlike their paper counterparts, the records management community until recently had little interest and less control over computerized information. Such information was largely viewed as "data," not "records." Archivists did not, therefore, have allies amongst those who controlled the corporate information systems and who were accountable, in theory at least, for the orderly scheduling and disposal of such records. Rather, archivists' contacts were only sporadic and personal with people in data administration areas, and this too reinforced the evolving focus of appraisal and acquisition on electronic records with informational value.

Surveys, statistical files, and census information tended to be "one-shot" data. A need was identified, a survey was designed, the completed questionnaires were automated, the results tabulated, the report written, and the project was over. There was no cumulative, longitudinal dimension to such records and, once collected and tabulated, it was highly unlikely that data would be added or deleted from the datafile. Each such file, in effect, existed independently. Each was fairly simple in structure and, as a result, could be converted into a so-called "flat file" that was relatively software independent. And each flat file, with sufficient documentation (record layouts, codebooks, data values, etc.), could readily be reconstructed to "run" using such standard social science statistical software packages as SAS and SPSS, for both data verification and researcher use.

Such "archival" machine-readable files were first retained in documentation centres or data libraries located in universities or within creating agencies themselves. Given this situation, as well as the isolated, independent nature of these early datafiles, the techniques of the library world were not surprisingly also adopted to describe or catalogue these early machine-readable datafiles in archives. In effect, datafiles were treated as publications, with their contextual relationship to creators, inventories, fonds, series, and related system information being either secondary or non-existent compared to highlighting their informational content as discrete bibliographic units.

By the late 1980s, this early approach was no longer deemed satisfactory for the National Archives of Canada. It imposed too many constraints on running an archival electronic records programme. In fact, a contextual, provenance-centred, evidential reorientation was seen as necessary to replace the social science-library focus of the first generation. As central agencies and departments in the Canadian government increasingly began to concentrate, as a matter of public policy, on information as content and resource rather than as

medium, as logical construct rather than physical artifact,<sup>2</sup> the National Archives also moved towards information integration for the two key media found in all government institutions. Accordingly, the two divisions most concerned with government records -- the Federal Archives Division (paper textual records) and the Machine Readable Archives Division (machine readable datafiles) -- were merged in December 1986 to form the new Government Archives Division.

From this change emerged the present structure of the electronic records programme at the National Archives of Canada.<sup>3</sup> The mandate of the Government Archives Division reflects those of its two predecessors, with the two exceptions that acquiring electronic data from private or non-government sources now falls within the mandate of the Manuscript Division and that managing electronic mapping and architectural data is assigned to a programme in the Cartographic and Audio-Visual Archives Division. Assisting these archival operations are two support areas: a tape library in the Conservation Branch which winds/copies and stores all electronic media and a unit of computer specialists in the Informatics and Records Services Branch whose technical expertise, as will be explained below, complements the archivists' work. Finally, an Information Management Standards and Practices Division has, as part of its mandate, research into and promulgation of improved practices and standards for electronic records-keeping throughout government departments. With this reorganization at the National Archives of Canada, a "second generation" of electronic records archivists and archival practices began to emerge, and by the early 1990's was firmly in place for appraisal and acquisition work, increasingly active in processing and descriptive initiatives, and starting as well to consider new reference paradigms.

Analyzing the characteristics of this second generation of electronic archives, and the reasons it has come to the fore, will perhaps indicate the pressures and constraints, as well as the exciting challenges and awaiting opportunities for archivists in other archival institutions facing similar reorientations of their older machine-readable records programmes, or, more likely, contemplating starting a new electronic records programme from scratch in the 1990's.

Taking the "why" question first, there have been fundamental changes in information technology and thus in the nature of computerized information to which archivists must now respond. While computer-based surveys and electronic census information obviously remain valuable, large hierarchical, networked, and especially relational databases are becoming the norm in the Canadian government. From an archival perspective and that of future users, running traditional statistical software such as SAS or SPSS against relational data may be impossible and, at the very least, will fail to unearth their richness; instead,

---

2 See, for example, *Management of Government Information Holdings*, a policy issued by the Administrative Policy Branch, Treasury Board, Government of Canada, 1 August 1989. See as well the phrasing in the *National Archives of Canada Act* (35-36 Eliz. II, c.1).

3 This reflects the situation of July 1983. Since then the organization has been re-organized.

sophisticated database management system (DBMS) software is required to establish or, better, re-establish the functional linkages that give the data meaning within certain applications in its office of origin. If datafiles coming from DBMS environments may not be as software independent as the smaller statistical files were in the first generation, neither are they any longer "one-shot" in nature: archivally valuable computerized data in large social and economic programmes' systems are often added, revised, or deleted almost every second, with evidence of such transactional instructions stored in metadata, while scientific and environmental data are longitudinal, cumulative, and extremely complex and expensive to maintain. Outside the world of structured data in such databases, wherein at least there are logical records with defined data elements and values, there is the unstructured realm of the automated office, where such media as text, graphics, images, and voice are now widely converted into electronic, often virtual formats across government. Indeed, these various unstructured "media" are often combined into a single record using intricate and highly software-dependent linkages. Combining computerized information in these new formats with a telecommunications revolution affecting the transmission of electronic records threatens decision-making accountability and corporate memory: if an electronic document has no physical existence, but rather is a virtual composite of disparate information appearing but fleetingly on a terminal, how does the institution, let alone the archivist, preserve evidence of significant transactions, especially as they relate to important decisions regarding programme activity. Where is the "evidence" or accountability of the transaction? Where is the context? What is provenance?<sup>4</sup>

Thinking through these developments has had a fundamental impact on the National Archives of Canada, as "traditional" archivists saw "their" records being automated. Computerized records were no longer just "machine-readable" statistical files kept off in a

---

4 For an excellent, generalized analysis, see Charles M. Dollar, *Archival Theory and Information Technologies* (Marcerata, Italy, 1992). The theoretical implications of these changes in electronic records for archival appraisal and arrangement/description, based on and largely arising from National Archives of Canada experience, have also been suggested by Terry Cook in, respectively, "Mind Over Matter: Towards a New Theory of Appraisal", in Barbara L. Craig, ed., *The Archival Imagination: Essays in Honour of Hugh A. Taylor* (Ottawa, 1992), pp.38-70; and "The Concept of the Archival Fonds: Theory, Description, and Provenance in the Post-Custodial Era," in Terry Eastwood, ed., *The Archival Fonds: From Theory to Practice* (Ottawa, 1992), pp.31-86. An important expansion of these views arising from National Archives experience, is Richard Brown, "Records Acquisition Strategy and Its Theoretical Foundation: The Case for a Concept of Archival Hermeneutics," *Archivaria* 33 (Winter 1991-92), pp.34-56. For a related study based on the new approach and its impact, also arising from National Archives experience in these issues, see Candace Loewen, "From Human Neglect to Planetary Survival: New Approaches to the Appraisal of Environmental Records," *Ibid.*, pp.87-103.

corner somewhere. Now it was letters, memoranda, policy summaries, operational case files, crucial financial spreadsheets, vital interpretive graphical material, even maps and photographs and sound, that were being automated. It was no longer a few records being made "machine-readable," but all traditional media being made "electronic," the media lines being blurred thereby, and the paper backups disappearing or not even being produced.

As more and more traditional archivists perforce got deeply involved in electronic records as an extension of their own media interests, especially following the reorganization merger of late 1986, they also brought their "traditional" archival principles to the world of computerized records. (This process mirrored, importantly, a similar transformation in the records and information management communities.) The library-oriented, discrete item approach of the first generation seemed both inadequate for the new types of electronic media and metadata, and inappropriate for archivists schooled to think in terms of context, provenance, and evidential value, in terms of records and integrated "knowledge" rather than data and disconnected "information." Indeed, several archival commentators stressed that the conceptual "power of provenance" not only holds the key to archival success in dealing with computerized records, but uniquely positions archivists to help records creators cope with their vanishing corporate memory.<sup>5</sup> A world of relational databases, of complex software linkages, of electronic accountability trails in office systems, of hypermedia documents, of multi-layered geographical information systems, is, in short, a world of relationships, of interconnections, of context. It is above all a world requiring "evidence" -- evidence of record creation, use, alteration, merging, deletion, transmission -- or at a higher level, a world reflecting business functions, interrelated programme delivery, decision-making, accountability, policy, and legislation.

Turning then from the theoretical and conceptual issues outlined above, the practical constraints against addressing these issues were very real, and led to an internal questioning of the older machine-readable records programme. Appraisal decisions, which according to archival principle ought to be made in context, with full understanding of the importance

---

5 For such sweeping statements anticipating the second generation of electronic records archives, see Terry Cook, "From Information to Knowledge: An Intellectual Paradigm for Archives," *Archivaria* 19 (Winter 1984-85), pp. 28-49; and David A. Bearman and Richard Lytle, "The Power of the Principle of Provenance," *Archivaria* 21 (Winter 1985-86), pp. 14-27. An excellent overview of this theoretical reorientation of the archival profession, a reorientation that in many ways made possible and underlines the new approach to electronic records archiving, is Tom Nesmith, "Introduction: Archival Studies in English-speaking Canada and the North American Rediscovery of Provenance," in Tom Nesmith, ed., *Canadian Archival Studies and the Rediscovery of Provenance* (Metuchen, N.J., 1993), pp. 1-28.

and inter-relationships of functions and activities of programmes and full knowledge of all records in all media forms, were usually made without such knowledge. Acquisitions of electronic records were largely isolated and episodic, in that no adequate system was operating to appraise government records systematically and to ensure that those identified as archival were transferred regularly to the National Archives.<sup>6</sup> Because of exclusive dependence on a service bureau for computing power, options for acquiring the increasing variety and complexities of electronic records were limited. Descriptive practices for electronic records were also brought into question. Should such records not be described in the same way as other archival records, that is, by reference to the context within which they were created? Concerted effort by the archival community to arrive at common descriptive practices across Canada gave an added impetus to this question. Finally, questions were raised about the ability of the National Archives to make electronic records available to the researching public. Was it always to operate more or less as a tape library? Could archivists, and eventually researchers, never use such records in an archival setting? Was the National Archives only to be a storehouse for the records of the future? Or should it, as some have suggested, be a clearing-house for information about records, and leave some -- maybe all -- of the records with the creators?<sup>7</sup>

In light of such questioning, certain concrete, practical steps have been taken within the National Archives to put its archival programme for electronic records on a firmer footing. In relation to acquisition, the National Archives of Canada Act is a very important instrument, in that it requires the National Archivist to give approval before government records in any medium can be destroyed. To end the fragmentary approach to appraisal and acquisition, an overall plan for the systematic appraisal of all records in all formats created by the Government of Canada has been formulated, which in turn led to individual appraisal plans with each government institution being negotiated and signed at a senior level. The impact of this planned approach to appraisal and acquisition on electronic records was significant. Archivists no longer get stuck down in the trenches facing often five hundred systems in a single large agency and trying to appraise them one-by-one in a near-hopeless attempt to stitch together disparate systems and applications. Rather, they adopt a top-down perspective related to corporate functions, programmes, and activities and focus **first** on the mandates and functions of the agency and its interactions with citizens and societal groups. Some such functions are increasingly inter-agency or even inter-jurisdictional in nature. **Only** once the significance of these broad functions are clearly assessed -- which requires careful research by archivists -- can archivists start pin-pointing key systems and records **in all media** for actual appraisal and possible acquisition by the Archives or for long-term retention by the agency under the archives' auspices.

---

6 For an extended critique based on National Archives experience in the Government of Canada, see Eldon Frost, "A Weak Link in the Chain: Records Scheduling as a Source of Archival Acquisition", *Archivaria* 33 (Winter 1991-92), pp. 78-86.

7 On this latter point, see the debate by Ken Thibodeau, David Bearman, and Margaret Hedstrom in David Bearman, ed., *Archival Management of Electronic Records*, Archives and Museum Informatics Technical Report No. 13 (Pittsburgh, 1991), pp. 1-30.

This new approach began in November 1990, and plans have been signed for initial appraisal work with some 35 institutions. It is evident in the appraisals completed to date -- some 60 in the first two and one half years for whole programmes of the federal government -- that this work is now being carried out in a useful manner. Effort is made in negotiating the plans to concentrate first on those functions of government which are most important and, by this means, records of highest archival value are being appraised in the initial stages. This allows the National Archivist to give approval for the disposal of other records having no archival value -- about 95 per cent total -- with more dispatch. Appraisal is carried out with a holistic knowledge of the significant functions and activities of whole large programmes, and it has been possible thereby to appraise all record holdings in all media of those programmes in context.<sup>8</sup> Archivists believe that better appraisal decisions are the result. Some significant electronic record holdings of Revenue Canada-Tax, the Privy Council Office, Treasury Board Secretariat, Environment Canada, the Public Service Commission, the Canadian Radio-Television and Telecommunication Commission, Indian and Northern Affairs, Central Mortgage and Housing Corporation, and Energy, Mines and Resources have been identified and protected. Some of these records have been transferred to the National Archives, and some are beginning to become available for research (availability is in accordance with the federal Access to Information Act and the Privacy Act).

Options have been expanded in the way electronic records are treated in the National Archives. The very expensive external service bureau has been abandoned in favour of an in-house computer facility which runs on an interconnected microcomputer and file server and permits a wider variety of tasks to be undertaken. This facility supports the appraisal of electronic records, in that records can be brought into the Archives for assessment and trial manipulation. It also permits records to be copied both from a variety of electronic media (diskettes, discs, magnetic tape, and cartridges) or via telecommunication linkages, and then verified and processed. A parallel system has been established for the acquisition and viewing of spatial data, such as that produced through Geographical Information Systems and computer-assisted mapping and design applications. These facilities have been designed to permit incremental expansion and enhancement. This is important as the pace of technological change is likely to continue to outstrip efforts at standardization of the format and structure of electronic records.

A final important practical step taken by the National Archives has been the hiring of staff and the creation of a special unit dedicated to provide computer systems support to archivists in appraising, processing, describing, and making available electronic archival records. The older model of the "data archivist" who did both traditional archival functions and computer processing at the terminal screen has been abandoned. The second generation archival reorientation recognizes that archivists and computer specialists are professional partners, not the same person, and that to assume that one can learn the other role

---

<sup>8</sup> This approach is outlined in "An Appraisal Methodology: Guidelines for Performing an Archival Appraisal, Government Archives Division, National Archives of Canada", by Terry Cook, 31 December 1991. The National Archivist has approved this methodology for comprehensive appraisal.

"on the job" overlooks the vital complexity of both. Migrating electronic records from relational databases or office systems to an archival environment, reconfiguring and remounting them on in-house archival DBMS software, linking data from different datafiles, and manipulating these software-dependent electronic records into special subsets for researchers all requires a computer expertise vastly different and far more complex than making tape copies or running cross-tabulations and data verifications on SAS or SPSS. To believe that an archivist could acquire the extensive technical expertise to do this is unrealistic, even though they must train themselves to be comfortable with the conceptual issues, if not the "how-to" mechanics, of technical appraisal. Conversely, to imagine that the computer scientist, without graduate education in history and archives, can become an archivist is even more unrealistic. The archivist must decide what data to save, how to describe it, how to manipulate it, how to make it available, and why; the computer specialist determines how to implement these decisions made by the archivist. The archivist is, in effect, a special kind of "user" with a defined set of requirements to which the computer specialist responds.

Readers should not infer from what has been said that the National Archives thinks it has solved all problems of dealing with electronic archival records. Hardly. Many challenges remain and these, in effect, form the managerial and research agenda for the months and years ahead. The results of thinking through these following issues will certainly lead to new initiatives and programme enhancements, and may in some instances lead to a fuller reorientation and, who knows, perhaps a "third generation" of electronic records archives.

As with other electronic records archives, the National Archives of Canada must deal with those computer records which still cannot be acquired because they are completely dependent on extraordinarily complex software. There are others systems with records of undisputed archival or permanent value -- weather and ozone monitoring data -- so huge and expensive that no archives in the world could begin to acquire them. The National Archives is only beginning to acquire and gain experience in electronic mail systems, and the very imperfect control of such records in government institutions gravely threatens their archival survival. Here the National Archives must work with others to define better what a "record"

is in an e-mail environment and to articulate functional requirements which will allow significant transactions to be documented as archival evidence.<sup>9</sup> Imaging systems are currently being implemented at an alarming rate across government without benefit of standards, making them incompatible with one another and across technological generations. Some of these situations, as noted before, challenge the traditional role of archivists as custodians of records and instead suggest that they may become instead controllers of information about records. On this last point, the National Archives is now finalizing a document stating the circumstances and conditions whereby archival electronic records will be left "out" with their creating institutions and thus not acquired by the National Archives.<sup>10</sup> Whether this will work in reality, in terms of obtaining the necessary cooperation and resources needed of the creating institution, will only be answered by practical experience.

Many issues associated with making electronic records available within archives also remain unresolved.<sup>11</sup> One will be the level and type of description appropriate for electronic records and their metadata. The applicability of the Canadian Rules for Archival Description to electronic records is still being investigated by a national working group devoted to the subject. Certainly experience to date suggests that neither the "first generation" approach of cataloguing machine-readable datafiles as discrete bibliographic items nor the traditional archival record group is applicable to multiple-provenance and multiple-layered records from complex information systems, but beyond that point the direction is not at all clear. The standard way of making electronic records available to researchers must also be changed. Archives have usually only provided copies of software-independent datafiles, with very limited facilities available for viewing records on archives' premises. Should archives operate computing facilities in support of researchers? Should linkages of different archival records obtained from different sources be encouraged in our archives -- with all the ethical concerns that entails concerning personal privacy? What use may be made of network facilities or optical media packaging to distribute data? If data are to be left permanently in agencies, should the archives have communications linkages with the agency's computer(s) and the capacity to query these databases on behalf of staff and researchers? It is expected that, in an era of declining budgets, answers to these questions will be dictated as much by financial as technological considerations.

---

9 Here the National Archives will follow the lead set by David Bearman in several recent speeches and manuscripts: see for example his "Electronic Mail: Is it a record? How can it be managed?" presentation to the Society of Canadian Office Automation Professionals, 31 March 1993; or "Record Systems as the Locus of Provenance: Implications for Automation of Archival Control and Management of Electronic Records", presentation to the Ontario Association of Archivists, 13 May 1993. The issue of records vs. data, of evidence vs. information, is carefully analyzed in Dollar, *Archival Theory*, chapter three, pp. 45-51, especially in relation to provenance.

10 "Leaving Archival Electronic Records in Institutions: Monitoring Arrangements", a draft (5th) policy document written by Terry Cook, 1 June 1993, for the comment and approval of the National Archivist.

11 Within the context of the National Archives of Canada, a number of these were identified and action steps planned to address them in "Government Archives Division: Issues Relating to Electronic Records", a policy paper by Eldon Frost, 10 July, 1992.

In this regard, after thinking through the theoretical and conceptual issues, the response of the Archives is essentially pragmatic. Some electronic records of permanent value should not be acquired by the Archives for reasons of cost or of technological barriers or the cumulative nature of the record. In these cases, the Archives is negotiating agreements with creating institutions whereby the permanency of the record is guaranteed, and the Archives will have a monitoring rather than a custodial role for the records. In other cases, the Archives is equipping itself to actually acquire, preserve, and make available an expanding variety of electronic records. It is expected that this mix of approaches will serve its research community best.

While not presuming at this point to have answers to all these questions, the National Archives of Canada has adopted approaches which it believes suit present and future circumstances. First and foremost, in legislation, policy, theoretical formulations, strategic approaches, and operational practice, it has firmly declared that the preservation of records in electronic format is essential to an understanding of Canada's past and to the on-going continuity of government operations. Secondly, the electronic records programme has been fully integrated with that for records in paper form emanating from the same source; archival principles of provenance and contextual analysis in both appraisal and description have been adopted for electronic records in place of an early media-exclusive focus. Such a sound theoretical base is essential for all archival programmes. Thirdly, archivists have as a central focus of their work a research agenda in the contextual information about records creators -- their history, functions, programmes, activities, structures, records-creating systems and processes, and the records themselves -- that provides the substance to all the procedures and processes in appraisal, description, and reference. Finally, the implementation at the National Archives of Canada of a "second-generation" electronic records archives removes the archivist and his or her records manager ally from their traditional, reactive, *ad hoc*, servant relationship with records creators and substitutes instead an active, planned, strategic, functions-oriented, research-based, and archival approach to archival appraisal and subsequent work.

Without such an approach, we are convinced archivists in the information age are doomed to an antiquarian, curatorial irrelevancy. With such an approach, we feel archivists are positioned as never before to inject an archival agenda into public affairs and to preserve a rich corporate and societal memory.