Review Article

Easy To Byte, Harder To Chew: The Second Generation of Electronic Records Archives

by **TERRY COOK***

Archival Management of Electronic Records. DAVID BEARMAN, ed. Pittsburgh: Archives and Museum Informatics Technical Report No. 13, 1991. 56 p. ISSN 1042-1459.

The Archives of the Future: Archival Strategies for the Treatment of Electronic Databases. NATIONAL ACADEMY OF PUBLIC ADMINISTRATION. Washington: Report to the National Archives and Records Administration, 1991. xii, 40 p. and appendixes.

Conceptual Problems Posed by Electronic Records: A RAMP Study. KATHARINE GAVREL. Paris: UNESCO, International Council on Archives, 1990. 49 p. PGI-90/WS/12.

Keeping Data: Papers from a Workshop on Appraising Computer-Based Records. BARBARA REED and DAVID ROBERTS, eds. Sydney: The Australian Council of Archives and The Australian Society of Archivists Incorporated, 1991. iv, 122 p. ISBN 0 947219 03 X.

Management of Electronic Records: Issues and Guidelines. THE ADVISORY COMMITTEE FOR THE CO-ORDINATION OF INFORMATION SYSTEMS (ACCIS). New York: United Nations, 1990. x, 189 p. ISBN 92-1-100348-2.

Managing Electronic Records. NATIONAL ARCHIVES AND RECORDS ADMINISTRATION. Washington: National Archives and Records Administration, 1990. vii, 32 p. and appendixes.

Research Issues in Electronic Records. NATIONAL HISTORICAL PUBLICATIONS AND RECORDS COMMISSION. St. Paul: Minnesota Historical Society for the NHPRC, 1991. vi, 37 p.

Taking a Byte Out of History: The Archival Preservation of Federal Computer Records. UNITED STATES CONGRESS, HOUSE COMMITTEE ON GOVERNMENT OPERATIONS. Washington: United States Government Printing Office, 1990. v, 30 p.

The foregoing array of publications appearing in the past two years signals the arrival of the second generation of electronic records archives. To many archivists, this may be disturbing news indeed, for most archives have yet to develop programmes to cope with the first generation of such records. In Canada and the United States, for example, aside from the two national archives, only a handful of states have active electronic records archival programmes and only three provinces at last count have positioned themselves to deal with computerized records; none of the latter has as yet acquired any archival versions of such records. At the September 1991 Archives Course held in Ottawa, of well over twenty representative archivists from across the country, only three had any experience dealing with electronic records (all from the National Archives of Canada), three more had access to a computer at work, and one more had a personal computer at home.

This professional lag occurs at a time when virtually all archivists acknowledge that computer records are increasingly pervasive among the records-creating organizations for which they are responsible, that such records are increasingly complex, and that the old safety-net assurance that there are paper backups is increasingly untrue or even irrelevant. To lack of knowledge or opportunity concerning electronic records, then, is added the guilt of knowing that a large part of our documentary heritage is disappearing without much archival intervention.

Archivists who missed the first generation of electronic records activity in archives may despair that the profession's leaders, as represented by the authors under review, are streaking into the second lap of the race to save that heritage, while they remain frozen at the starting gate. Such despair may be misplaced, for it is easier (it will be argued) to join the electronic records race the second time round than it was initially. Indeed, the eight publications under review represent a watershed of opportunity for the novice in electronic records rather than a cascade of hopelessness. A review of the evolution of electronic records programmes in archives — in effect, an outline of the differences between the first and second generations — will support this assertion, as well as place the volumes under review in context.

The archival pioneers of the first generation of computerized records in the 1970s and early 1980s had, along with the exhilaration of being trail-blazers for the profession, the pain of false starts and blind alleys; the electronic records programmes in both national archives, for example, virtually disappeared in the mid-1980s before being reorganized into "second generation" entities. For the pioneers, 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 away from the era of mainframe computers and punched cards in the 1950s and 1960s 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 such social science data files, the only business applications being automated were administrative, such as payroll, inventory, shipping, receiving, accounts receivable and so on, and records produced by these functions 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 pioneering generation. These circumstances in turn had numerous implications for the development of the first generation of archival practice in electronic records.[1]

Statistical and survey files were primarily forms or questionnaires which, for ease of tabulation, 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.[2] 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 in and less control over computerized information. Such information was largely viewed as "data," not "records." Archivists did not, therefore, have allies among 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 — usually sporadic and personal — were only with people in data administration areas, and this too reinforced the evolving focus of appraisal and acquisition on electronic records having informational value.

Surveys, statistical files and census information tended to be "one-shot" data. A need was identified, a survey (of teenagers' smoking habits, perhaps) 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 to or deleted from the data file. 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.[3] 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 access.

Such "archival" machine-readable files were first retained in documentation centres or data libraries located in universities, or within the 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 data files in archives.[4] In effect, data files were treated as publications, 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.[5]

Reference service in the first generation of electronic records archives was fully consistent with the foregoing appraisal and descriptive developments. Researchers were given a version of the flat file copied on to magnetic tape, along with a copy of the related documentation. They were then left to "run" the data file through the statistical software package of their choice, at an outside computer facility having a mainframe capacity. In effect, reference and retrieval service mirrored that of a records centre or library, where the box or book is handed out as a discrete item.

Archivists who pioneered work in this environment had the added burden of very user-**un**friendly computers and processes. Activities such as copying a data file on to a backup tape and obtaining a print-out of the record label or first 100 logical records within the data file for verification purposes often took hours, and required very cumbersome computer language and commands. The pioneers of the first generation did not have the benefit of the world of microcomputers with easy command menus that mask for today's archivist scores and scores of lines of computer code; rather, they had to enter that code themselves and watch "jobs" crash if a semi-colon was mistyped. Yet because the flat-file data structures were relatively simple, those early archivists did this technical computer work themselves. Thus the practice evolved that "data archivists," as they were generally called, were characterized as performing both the archival functions of appraisal, description and reference, as outlined above, and the technical processes at the actual computer terminal involved in copying, verifying and manipulating the machine-readable records. By the mid-1980s, this increasingly uneasy union of the archivist and the computer technician in the same person resulted in a world of arcane procedures, "high-tech" jargon and almost impenetrable practice, especially as data structures had evolved from the fairly simple to the exceedingly complex.

Two very fine, landmark books by Margaret Hedstrom and Harold Naugler in 1984 may justly be seen as the culmination of the pioneering generation.[6] They indicate clearly to any sensitive reader the debt which the archival profession owes to the first-generation pioneers who, starting with almost nothing, produced within the confines of the prevalent data structures and computer technologies of the period, a lasting legacy of theory and practice. Both Hedstrom and Naugler also deal with, or at least foreshadow, the new, more complex electronic records of the second generation, the need for more archival contextual approaches, and the growing importance of evidential concerns in appraisal. Both monographs are in that sense transitional, summing up the best of what went before and anticipating the new generation characterized by the eight works under review. And that merely underlines that the foregoing broad canvass of the history of the first generation of electronic records archives is undoubtedly unfair in many details, while I hope accurate in capturing the overall landscape. It goes almost without saying that this generational distinction concerns archival mindsets and programmes, **not** individual archivists; such first generation pioneers as Margaret Hedstrom, Charles Dollar and John McDonald remain at the cutting edge of second-generation electronic records archival activity.

What, then, are the characteristics of the second generation of electronic records archives, and why have they come about? Taking the "why" question first, there have been fundamental changes in information technology, and thus in the nature of the computerized information to which archivists must now respond. While some computer-based surveys and electronic census information remain valuable, large hierarchical, networked, and especially relational databases are becoming the norm in business, universities and government. Here, electronic information is stored in many internal tables, entities or structures, that have meaning only inasmuch as they are related to each other. From an archival perspective, and that of future users, running traditional statistical software such as SAS or SPSS against relational databases may be impossible, and, at the very least, will fail to unearth their richness; instead, sophisticated database management system (DBMS) software is required to establish or, better,

re-establish the linkages that give the data meaning. If data files coming from DBMS environments may not be so 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 may be added, revised or deleted almost every second, 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 having defined data elements and values, there is the unstructured realm of the automated office, where text, graphics, images and voice are converted into electronic format. Compound, "smart," or hypermedia documents take this a step further, merging these individual "media" in ways that are complex, very software-dependent, and very difficult to re-create. 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?

These developments are having a fundamental impact on the archival world, as "traditional" archivists see "their" records being automated. Computerized records are no longer just "machine-readable" statistical files kept by data archivists far off in a corner — and, moreover, so the thinking erroneously went, were there not published summaries of most of the important statistics in any case? Now it is letters, memoranda, policy summaries, operational case files, crucial financial spreadsheets, vital interpretive graphic material, even maps, photographs and sound recordings, that were being "automated" — that is, being converted into the digital bits that make up electronic records. It was no longer a matter of a few records being rendered "machine-readable," but all traditional media being rendered "electronic," the media lines being blurred thereby, and the paper backups either disappearing or not even being produced.

As more and more traditional archivists necessarily got deeply involved in electronic records as an extension of their own media interests, an involvement mirrored in the records and information management communities, they also brought their "traditional" archival principles to the world of computerized records. The library-oriented, discrete-item approach of the first generation seemed both inadequate for the new types of electronic media and inappropriate for archivists trained to think in terms of context, provenance and evidential value. Indeed, David Bearman, the most visionary of archivists dealing with electronic records, has stressed that the conceptual "power of the principle of provenance" not only holds the key to archival success in dealing with computerized records, but also uniquely positions archivists to help records creators cope with their vanishing corporate memory.[7] 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, decison-making, accountability, policy and legislation.

Such evidential and functional context is the province of the archivist. Re-creating the interrelationships of electronic records should be no different for the archivist, **at a conceptual and theoretical level**, than unravelling the interconnections of many series of related registers, indexes, incoming dockets, outgoing letterbooks, subject files, coded maps, photograph albums, blueprints and postcards from a nineteenth-century office. In fact, the ephemeral nature of the new electronic media, by **reinforcing** the contextual heart of the profession, may well make even the "traditional" archivist more archival.[8]

This archival reorientation towards electronic records has been made easier (and caused by) a growing amount of legislative and policy changes affecting information management in many jurisdictions, which in turn naturally reflect developments in information technology. The general thrust of these changes is to treat information as a corporate resource, to include electronic records as a full partner within information management, and to clarify the regulations governing the ongoing management, control, scheduling and disposal, including archival retention, of all recorded information. The movement towards standards (especially "open systems") for data interchange to ensure transportability of electronic information among different hardware and software environments, as well as increased data storage capacity at greatly reduced costs, are other developments conducive to easier archival participation in the new world of computer records. Naturally, these changes also support the acquisition of electronic archival records from many disparate systems among creating departments. As a result of pressure from "corporate" or institution-wide functional business models and more global information policies, as well as technological advances, more holistic conceptions of electronic information are now possible. These are expressed in the form of metadata, automated data dictionaries, and corporate data models. This establishes a *de facto* contextual agenda among the creators and contemporary users of these new records that is very compatible with archival presuppositions concerning provenance, interrelationships, context, order and evidential transactions.

The second-generation archival reorientation has also been accompanied by (or resulted in) the growing recognition that archivists and computer specialists are professional partners, not the same people, and that to assume that one can learn the other role "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 data files, and manipulating these software-dependent electronic records into special subsets for researchers, all requires a computer expertise vastly different from and far more complex than making tape copies or running cross-tabulations on SAS or SPSS. To believe that archivists 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 practical 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" having a defined set of requirements to which the computer specialist responds.

Rather self-evidently, given the volumes under review, the second generation of electronic records archives has also witnessed an explosion of research reports, studies and publications to which archivists can refer, a luxury which first-generation electronic records archivists did not enjoy. For all these reasons, it is easier for an archivist without experience in the electronic records field to enter the race in its second lap without having to repeat the discovery work of the pioneers. In so doing, archivists will find in the second generation a congenial environment based on archival principles, on evidential interrelationships rather than discrete items of information, and on contextual and provenance paradigms rather than technology- or user-driven approaches. This certainly does not mean that the issues, both theoretical and practical, are less difficult. In many ways, the problems are still being defined, the issues are extraordinarily complex, and the solutions are tentative at best, unseen at worst. The second generation, too, will have its pioneers. If it is easier to "byte" into this new electronic loaf, it will not necessarily be easier to chew!

The publications under review present numerous access points to the second generation of electronic records archives. My intention is to summarize the general approach of each work in order that readers, novice or experienced, may decide for themselves where to start, according to their needs. There is little point in conducting a detailed comparative critique of the eight works, for the differences in scope and purpose are so great as to be comparing apples with oranges, or perhaps more accurately apples with camels! Nevertheless, in surveying these works, generally in order from the more practical to the more theoretical, I hope some of the foregoing assertions concerning the evolution of electronic records archives over the past two decades may become a little clearer, for many are based on the issues raised in these publications.

Taking a Byte out of History was produced by the Committee on Government Operations of the United States House of Representatives. That a Congressional Committee is concerned with, as the subtitle runs, "The Archival Preservation of Federal Computer Records," is an encouraging sign. The most recent in a large literature drawing attention to the vanishing electronic heritage,[9] *Taking a Byte* notes that 75 per cent of all government transactions will be in electronic form by the year 2000. Unless government agencies and the National Archives and Records Administration (NARA) take decisive action, computer records will be like Egyptian hieroglyphics before the Rosetta Stone: seen, but not understood. Stating that "if a computer record cannot be read, then for all practical purposes, the record no longer exists," the Committee warns that there will be no electronic Rosetta Stone for this medium. In the face of proliferating personal computers, incompatible software, and ever-more-complex data structures and uses, the Committee believes that keeping electronic records readable, and therefore in existence, requires long-range research and planning by NARA into appropriate hardware and software for electronic records-keeping systems, the development and implementation of standards for computer records to minimize hardware and software differences, and the establishment of records-keeping and long-term preservation standards for major electronic records systems at the outset, as a mandatory system-design consideration. The report also explicitly criticizes archivists for relying on first-generation techniques for current electronic records. The practices suitable for the archival files from the era of the mainframe computer, centralized batch-processing and magnetic tape storage, cannot cope with

automated office records, complex relational databases, and the expected needs of future users to access, link and manipulate archival data in a myriad of ways. All in all, *Taking a Byte* is a good introduction to electronic records archival issues, the changing technological landscape, major emerging problems, and possible solutions.

Managing Electronic Records is a free publication (while supplies last) of the National Archives and Records Administration. Part of its "Instructional Guide Series" aimed at providing advice to officials of government agencies on the management of particular types of records, it must be seen for the primer which it is. The booklet begins with brief sections on roles, responsibilities and a general overview of records disposition. It then deals succinctly with inventorying electronic records, applying general records schedules, preparing schedules for operational records not covered by general schedules, identifying potentially archival records, and transferring such records to NARA, as well as offering advice on the maintenance and use of electronic records. A useful appendix suggests the appropriate disposal of different categories of electronic records at certain stages in their life cycle, including master files, system update files, input/source files, extracted data files, reformat and print files, and backup files. The primary flaw of *Managing Electronic Records* is that it does not address the corporate management of electronic records, as its title promises, but rather focuses on their scheduling and disposal. On that basis, however, it is a brief, helpful companion to anyone beginning work in electronic records, whether as a records manager or as an archivist.

Much more concerned with the current management of computer-generated records, and providing a great deal more detail and sophistication, is the United Nations' publication *Management of Electronic Records: Issues and Guidelines*. Revealing the steady hand of David Bearman as principal author among a long list of official participants, this study addresses the records management issues concerning electronic records for the United Nations and its many subsidiaries: the World Bank, International Monetary Fund, World Health Organization, UNESCO, International Labour Organization, and so on. Throughout all these agencies, information technology has been widely introduced, especially individual computer workstations, in order to improve operational efficiency and productivity. In so doing, however, issues concerning the management of the information produced by this technology were generally overlooked. Electronic records exactly like their paper predecessors are needed by the United Nations (or indeed any corporate body) not just for such increased productivity, but also for "management accountability, operational continuity, legal evidence, disaster recovery, and 'institutional memory'." Such information management issues have not been squarely faced within the United Nations, and this report seeks to change that situation. The UN situation is common to most jurisdictions, where information technology is rapidly introduced, but information management (and archival retention) policy concerning the new technology lags far behind. It is in archivists' best interest to close that gap.

The focus of *Management of Electronic Records* is on the records creator's corporate resonsibility for managing its current information properly. The book is thus a particularly excellent guide for those readers with records management responsibilities, for the volume contains much sensible advice, step-by-step guidelines, and analysis of policy issues and implementation options. Yet archivists as much as records managers

will find in it a wealth of theoretical and practical issues of interest to them, for the guide also discusses the nature of the electronic record, provenance, problems of evidence and accountability, and policy formulation. It also contains the best exploration in print of the implications for long-term preservation of electronic mail in office networks, and the role of standards in integrated systems management. There is a particularly useful appendix analysing the types of electronic records, and the best available glossary of terms used in electronic records archives and records management. **Overall, this is the best single volume of the eight under review.**

Both the United Nations and NARA publications approach the management of electronic records in ways quite familiar to archivists and records managers dealing with paper records: both works outline the need for corporate control of all recorded information and its orderly disposal, including archival preservation. The National Academy of Public Administration's (NAPA) report, *The Archives of the Future: Archival Strategies for the Treatment of Electronic Databases*, takes a very different tack. Indeed, such debates over varying strategic approaches to appraising and acquiring electronic records are very much a characteristic of second-generation archival activity in this field.

Rather than recommend the orderly disposal of all information in all media from a corporate or business perspective, NAPA decided to focus on identifying the major databases across the United States federal government, "particularly those 'Fortune 500' or so that have significant historical and research value." After inventorying some 9,000 major databases in government agencies, NAPA turned to panels of "subject matter experts" (historians, political scientists, geographers, etc.) to develop appraisal criteria for approximately 900 databases identified as having the greatest research potential. These criteria, which are reproduced in the report, cover electronic records in five areas: diplomatic and foreign affairs, fiscal and economic, military and defence, natural resources and science, and social and judicial. There are some useful suggestions among these criteria, but they are inconsistent. Some follow traditional archival appraisal criteria for electronic records (data integrity, uniqueness, manipulability, extent of data coverage, documentation availability, etc.). Other criteria underline the importance of certain **generic** characteristics of electronic records which archivists should consider protecting because they greatly facilitate research (in studies of diplomacy, for example, researchers will require standardized electronic text markings to recover which words and phrases were underlined or otherwise emphasized, transmission records if a document was sent electronically, and evidence of the use of aggregated data compared to raw microdata, etc.). More problematically, other criteria outline the subject areas for which researchers will need electronic records preserved (terrorism, arms control, foreign aid, etc.).

There are admittedly advantages to the NAPA approach. First, it is a strategic, active approach to identifying and appraising major databases in large corporate settings, rather than the failed, passive approaches of the past. The NAPA approach also has the advantage of ensuring that very scarce archival resources for appraising, acquiring and processing large databases are distributed evenly across agencies, rather than spent on a first come-first served basis, as archivists schedule databases in a contextual vacuum, one by one, until, presumably, the money runs out. In addition to the focus on strategic database appraisal, this report also contains many other useful observations concerning

electronic records, including recommendations for future research and advice for strategic positioning of the profession to deal with electronic records issues in the years ahead. The work also contains lists of all the assessed databases, complete with the appraisal results for each.

However, there are fundamental problems with the whole design of the NAPA project. Its isolation of 900 key databases by subject should cause many archivists instinctively to recoil. It wrenches records from their natural context of creation and divorces them from closely related records. It separates the electronic records in these databases from the business activities of the creator, the need for operational corporate memory, and the functional framework of the creator's information universe. More than media myopia, it imposes on appraisal the **external** criteria of anticipated use rather than criteria designed to reflect the **internal** functionality behind creation (i.e., provenance). This report wrongly suggests that no archival strategy will "work without the active involvement of the historical and research communities," so that their "needs" are "made known and understood" and given "reasonable accommodation." This approach frankly returns archivists to the 1970s stance, against which Gerald Ham long ago protested, of being "at best nothing more than a weathervane moved by the changing winds of historiography."[10] Why pick the five NAPA subject areas — why not women, or urban life, or communications as major "subject" foci? In short, the NAPA study is provocative, but fundamentally flawed.

Keeping Data: Appraising Computer-Based Records is the closest thing to a textbook among the works under review. This seems intentional, for the title imitates the well-received text, *Keeping Archives*, which was also published by the Australian Society of Archivists. Although the proceedings of a conference and the work of thirteen authors, the papers are arranged in a logical order, and whatever duplication occurs does not detract from the overall unity of the work. There are four sections. The first places computers and information systems in context by providing helpful introductions to computer terms and functions, standards, system development, and the concept of documentation. The second outlines methods, giving case studies and examples, for collecting the information about information systems that is required for sound archival appraisal, in both the public and the private sector. The third examines methods of appraisal in general, and includes several extensive and relevant case studies. The fourth, entitled "Reflections," contains three brief essays on the challenges posed by electronic records to traditional archival theory. As with *Keeping Archives*, *Keeping Data* is generously illustrated with sample forms, diagrams and work flow-charts that complement the text; other archivists will doubtless find it useful to adapt some of these to their own practice

The Australian appraisal strategy is designed to apply to existing databases and to those still in the system-design phase. By evaluating information, not records *per se*, the approach closely follows Harold Naugler's RAMP study, cited earlier, and the system overview approach pioneered by John McDonald at the National Archives of Canada (although neither are acknowledged). Information is gathered on the content and processes associated with the system, content and technical analyses takes place, data flows are considered (although how this differs from technical appraisal is not made clear), and then appraisal decisions are made. Unlike *Taking a Byte* or the NAPA study, little consideration is given (despite much discussion of surveys of computer

environments) to the strategic problem of macro-appraisal, of how to appraise the many databases in large corporations and governments in order to isolate those few that should be subjected to the more detailed appraisal methodology outlined in *Keeping Data*.

The concluding theoretical "Reflections" section is the only disappointing part of the book. While raising several important points, these are rarely analysed in depth. After predictable musings on the impact of electronic records on the old debate of archivist *vs.* records manager and life cycle *vs.* continuum, Frank Upward believes that he "is getting fanciful" in concluding, almost as an aside, that future archivists will focus "on appraising the records creator before the records ..." He adds that electronic records "will rip apart organisational walls." But neither assertion is pursued — and neither is fanciful, as both are the subject of keen debate in North America. Yet taken together, these assertions undermine much of the thinking of the preceding chapters of *Keeping Data*, to say nothing of being implicitly contradictory: records creators exist within organizations, do they not? Surely the key theoretical issue, long debated by social thinkers, is the relative importance of function and structure in creating any social dynamism, including records-keeping activity and the resulting records. Despite declaring Jenkinson "triumphant" in the title while noting provocatively "a willingness to abandon the record as sacrosanct," Michael Saclier's piece is not theoretical at all, but rather an advocatory piece urging archivists to action. Glenda Acland saves "Reflections" from being an entirely weak ending to an otherwise fine book. Her "Archivist — Keeper, Undertaker or Auditor" explores how Peter Scott's Australian series approach to archives might relate to electronic records, although, as her title implies, she is more concerned with the strategic positioning ("upstream" or "downstream") of the archivist in Information Resource Management.

David Bearman's *Archival Management of Electronic Records* contains six reworked papers from the 1990 conference of the National Association of Government Archivists and Records Administrators. All authored by experienced electronic records archivists, these essays provide readers with a good entrée to the conceptual problems and choices posed for the profession by the appraisal and custody of electronic records (there are three essays on each). Less procedural or introductory than some of the works mentioned above, *Archival Management of Electronic Records* is strategic in focus and argumentative in style. Alan Kowlowitz urges archivists appraising electronic records to rise out of the trenches and stop "appraising in a vacuum." Otherwise, their work threatens to dissolve into little more than "a fruitless intellectual exercise," whereby only very few electronic records of continuing value are being preserved and made accessible to researchers. To prevent this, archivists must join other information professionals in establishing a multi-jurisdictional information policy, comprehensive records and information management programmes for electronic records, and the technical capacity to take archival custody of at least certain types of electronic records, so that appraisal may be more (pun intended) than a paper exercise. Michael Miller agrees with Kowlowitz, but broadens the discussion by urging that electronic records, stop being treated as "special" cases by archivists — in their thinking and their own internal organizational structures. He asserts that traditional appraisal theory should be seen as a great strength in coping with the new multiple-media information age. To the question in his title, "Is the Past Prologue?" he answers "yes," for the key questions archivists must ask of electronic records are those which they ask (or should ask) of

more traditional recording media. In outlining the new approach to appraisal in Canada at the federal government level and the revitalized electronic records programme at the National Archives of Canada, my own article in the volume supports Miller's assertion that the power of provenance, of contextual, holistic conceptualizations of the information universe, holds the strategic key to archivists' success in dealing with electronic records.

On the custodial side, Ken Thibodeau argues in "To Be or Not To Be: Archives for Electronic Records" that it is both possible and desirable to get electronic records into an archival environment. After explaining carefully the evolution from flat file environments to those of the relational database systems now common, Thibodeau argues that electronic records, designated as archival, can be managed by an archival institution in a cost-effective manner. He suggests that if the records are left under the creating institution's control, they are really at risk because the agency's priorities will never include updating archival electronic records to keep them readable in the face of countless, inevitable technological and scientific changes in the future.

David Bearman strongly disagrees, asserting in "An Indefensible Bastion: Archives as a Repository in the Electronic Age" that archives cannot begin to cope with transferring, controlling and making available the volume or complexity of modern electronic records. Except as a last resort, archives should not acquire electronic records at all. Rather they should identify those records which have continuing or archival value, work with agencies to ensure their long-term preservation, and serve as an information entrepôt that places in context for its clients the myriad of databases maintained "out there," in order to direct researchers to the correct sources of data which they may require. Between these two poles, Margaret Hedstrom attempts to mediate, and does so by first criticizing the usual assumption made by archivists that electronic records somehow impede traditional archival objectives. This "bleak perspective" overlooks the many advantages electronic records offer to archives, to researchers, to archival principles and to the enhanced responsibilities of the archivist. Central to these is the possible shift in focus away from preserving records, or even data, to preserving "a wide range of functionality from an active records system." The argument should not be whether archivists are Thibodeau's custodians or Bearman's regulators, but which types of records-creating environments and which types of electronic information require which approaches in order best to meet the needs of users, both in the parent institution and on the outside. To this end, Hedstrom advances nine criteria that will be very helpful in making the required choices. Her attempt to equip the archivist with a portfolio of varied responses is certainly, in my view, not just a compromise, but the best direction for electronic records archives in the next five years.

Katharine Gavrel's RAMP study, *Conceptual Problems Posed by Electronic Records*, contains a wide range of useful observations concerning the second-generation world. One of only five Canadian authors represented in the ICA's RAMP series of nearly 100 publications, Gavrel reflects herein her work as a pioneer in electronic records. After reviewing the evolution of machine-readable records over the past twenty years in terms of practices and procedures, and noting how traditional archival approaches were both adopted and modified by this process, she surveys the recent changes in information technology that are profoundly affecting how records are created, and therefore the

problems which electronic records archivists face in the second generation. In the short theoretical chapter dealing with *respect des fonds* and provenance, Gavrel pits Michel Duchein's fonds against Peter Scott's series in light of constantly changing bureaucratic structures and easily copied electronic records. Unfortunately, the discussion is brief and derivative, and adds nothing to what is known about the concept of *fonds d'archives*. The chapter on appraisal, however, is stronger. Not wanting to repeat the general guidelines for appraising electronic records provided in Naugler's RAMP study, Gavrel focuses on the special problems of electronic records in automated office systems, geographical information systems, and inter-organizational databases. These are well defined (with the exception of the geographical information systems), examples are given, and the resulting archival challenges are raised. The chapter on arrangement and description mainly concentrates on how electronic records were processed in the old world of statistical files, and how new standards will make data more transferrable. While this discussion does not explore the relationship of electronic records, first- or second-generation, with those traditional tools of archival arrangement and description, the inventory and the series, it does address the multimedia nature of electronic records and the challenge this poses for the concepts of record group and fonds. The chapter on conservation, storage and use of electronic records, which is Gavrel's specialty, is much stronger, offering good advice and updated information on these topics. In summary, at the level of archival concepts and theory, this RAMP study teases with good insights, but, from a broader archival perspective, does not satisfy. It certainly raises many of the issues posed by the evolution of records from physical to logical structures, and provides a useful basis for the subsequent analysis that archivists will have to conduct in order to develop concrete solutions or theoretical reformulations.

Finally, archivists get a look into the future and, who knows, maybe an early peek into the third generation of electronic records archives. The National Historical Publications and Records Commission (NHPRC) gathered forty-six experts in early 1991 to spend two days developing a research agenda for electronic records. In effect, the NHPRC was being provided with advice on which types of projects it should fund in the next few years. *Research Issues in Electronic Records* is the result (also free of charge while supplies last).[11] Each issue is identified, its purpose defined, background delineated, possible approaches to solutions suggested, and the benefits of tackling the issue outlined. Ten central issues were identified, the first three being seen as priorities upon which the others depend. Those three ask which functions, players and information are required to effectively manage electronic records in order to meet archival requirements; what are the technological, conceptual and economic implications of capturing and retaining data, descriptive (content) information, and contextual information in electronic form through a variety of applications; and how software-dependent data objects can be retained for future use. Briefly, the seven other areas of potential research are as follows: how can data dictionaries, information resource directory systems, and metadata be used to support records management and archival retention programmes for electronic records; what archival requirements should be addressed at the system design stage, and why; what policy frameworks are required to support archival concerns about electronic records; what functions and activities should be present in electronic records archival programmes, and how should these be evaluated; what incentives can make creators and users of electronic records support records management concerns in this area; what barriers have prevented archivists from

developing and implementing electronic records programmes; and, finally, what do archivists need to know about electronic records.

What they need to know, at least in part, is avaliable in these eight publications under review. As I hope is clear, all have merit. Each book will seem "the best" to certain archivists, depending on their experience and the immediate challenge before them in the electronic records field. But no reviewer should get away so easily, so here are my recommendations. For someone with no background or practical experience in electronic records, the Australian quasi-textbook, *Keeping Data*, is the best place to start. For those in the field who have mastered first-generation approaches or finished *Keeping Data*, and who are seeking the latest on information technology and strategic directions, the best work is the United Nations' publication, *Management of Electronic Records*. And for beginners and seasoned veterans alike wishing to speculate on the conceptual implications of electronic records for archival programmes, the best sourcebook is the group of essays which David Bearman has assembled, *Archival Management of Electronic Records*.

Facing such a banquet with eight courses on the table, there is no excuse for any archivist to hesitate biting into this electronic food. Archivists now need the nourishment so provided in order to cope with the archival challenges of the present and immediate future. Without it, they and their profession will gradually starve. Without it, moreover, they will certainly be unable to move the profession forward by themselves contributing to the too meagre professional literature that combines mainstream archival thinking with electronic records, or that joins the growing awareness of a corporate business context to the archival vision of preserving functionality through provenance, or that explores the unique contribution which archival theory can make to human understanding in the Information Age. And such an inability to contribute would be, as one might say while sipping tea at the end of the banquet, a pity.

Notes

* I want to thank Ed Dahl, Eldon Frost, Candace Loewen and especially John McDonald, who read the first draft of this essay and gave me many helpful suggestions. While most of these have been incorporated, I alone am responsible for the interpretations that follow.

1 In addition to my own observations over sixteen years at the National Archives of Canada and the scattered information found in some of the eight works under review, the following "historical" analysis is informed, at least in part, by Catherine Bailey, "Archival Theory and Electronic Records," *Archivaria* 29 (Winter 1989-90), pp. 180-96; Katharine Gavrel and Walter Meyer zu Erpen, *Machine Readable Archives Division*, National Archives of Canada, General Guide Series, (Ottawa, 1984); and Michael L. Miller's workshop on electronic records given at the Mid-Atlantic Regional Archives Conference in November 1990.

2 The classic statement is Charles M. Dollar, "The Appraisal of Machine Readable Records," *The American Archivist* 41 (October 1978), pp. 423-30.

3 A flat file has been defined as one the structure of which does not support hierarchical relationships. It is thus a two-dimensional, self-contained arrangement of data elements and logical records. Because relationships with other hierarchies or networks of data or files are not important, it can be converted or migrated relatively easily to standard software for use.

4 The seminal work is Sue A. Dodd, *Cataloging Machine-Readable Data Files: An Interpretive Manual* (Chicago, 1982). It betrays little of the provenancial and contextual focus of archival description, and effectively expresses the first-generation descriptive practices in electronic records.

5 This is stark, but I think generally true. In the documentation packages, there certainly was contextual information regarding the creating agency, the methodology of data collection, and links to the parent programme. But the overwhelming emphasis and public presentation was content-based rather than contextually oriented.

6 Margaret L. Hedstrom, *Archives & Manuscripts: Machine-Readable Records* (Chicago, 1984); Harold Naugler, *The Archival Appraisal of Machine-Readable Records: A RAMP Study With Guidelines* (Paris, 1984).

7 See David Bearman and Richard Lytle, "The Power of the Principle of Provenance," *Archivaria* 21 (Winter 1985-86), pp. 14-27; David Bearman, "Multisensory Data and Its Management," in Cynthia Durance, ed., *Management of Recorded Information: Converging Disciplines* (Munich, 1990); and David Bearman, *Archival Methods* (Pittsburgh, 1989). This is an approach I strongly endorse as well: see Terry Cook, *The Archival Appraisal of Records Containing Personal Information: A RAMP Study with Guidelines* (Paris, 1991); and "Mind Over Matter: Towards A New Theory of Archival Appraisal," to appear in the ACA festschrift for Hugh Taylor, forthcoming 1992.

8 For more on this, see Terry Cook, "From Information to Knowledge: An Intellectual Paradigm for Archives," *Archivaria* 19 (Winter 1984-85), pp. 28-49; and Terry Cook, "Rites of Passage: The Archivist and the Information Age," *Archivaria* 31 (Winter 1990-91), pp. 171-76. It is also the theme, in part, of Michael Miller's essay referred to below and of a great deal of David Bearman's work.

9 The best known is perhaps Committee on the Records of Government, *Report* (Washington, 1985), with its stark opening line: "The United States is in danger of losing its memory." This report still bears reading.

10 Gerald Ham, "The Archival Edge," in Maygene F. Daniels and Timothy Walch, eds., *A Modern Archives Reader* (Washington, 1984), p. 329 (article first published in 1975).

11 A precursor to this project that archivists should not ignore is National Historical Publications and Records Commission, *Electronic Records Issues: A Report to the Commission*, Commission Reports and Papers No. 4 (Washington, 1990). In nine brief pages, this pamphlet neatly outlines the key issues which archivists, their institutions and sponsoring agencies must face in electronic records: "the most significant and difficult challenge currently confronting the archival community."