Counterpoint

Myth or Reality: Is There a Generation Gap Among Electronic Records Archivists?*

by THOMAS ELTON BROWN

Terry Cook proposed a paradigm for how traditional archivists have confronted electronic records in an article entitled "Easy to Byte, Harder to Chew: The Second Generation of Electronic Records Archives." For the past thirty years, he argued, archivists have confronted two separate and distinct generations of automated records. Initially, the first generation consisted of machine-readable statistical survey data files. These gave way to a second generation with far more complex records created by far more complex technologies. This interpretation sees the archivists administering this first generation as isolated from other archivists concerned with traditional formats. In this view, the focus is on data files as individual dissemination products separate from the context of their creation. Richard Cox seized on Cook’s concept of distinct generations of archivists for a monograph in which he argued that little had been done in the archival management of electronic records until the 1990s. This article evaluates these two studies and challenges their conclusions.

Terry Cook’s article commented on eight publications dealing with electronic records and traditional archives, including some which argued that electronic records archives have witnessed two generations. The first generation dates from the 1970s and early 1980s. The late 1980s were the transition period, and the second generation emerged during the current decade. The article maintains that initially archivists had no “archival models” to follow and hence turned to the other professionals managing computerized information for research purposes—primarily social science data archivists and data librarians. In making this distinction, Cook does not accept data archivists as professional archivists. Because the holdings of most social science data archives were predominantly machine-readable statistical census or survey files, the article argues that these types of data files likewise became the focus of traditional archivists working with electronic records. He suggests that this led them to neglect electronic records more central to the business of creating organizations. Given the scenario where these archivists were isolated from other archivists, and given the independent nature of the survey data files as frequently the result of “one-shot” data collection activities, Cook argues that archivists described the files as discrete bibliographic entities. These descriptions were separate both from their contextual relationship with the creators and from the full corporate documentation.
of the agency of creation. Cook maintains that this concept of archival holdings as individual units dictated archival description as discrete entities similar to books. Finally, to deal with the first generation of electronic records, the archival professionals were jacks-of-all-trades "characterized as performing both the archival functions of appraisal, description and reference...and the technical processes at the actual computer terminal involved in copying, verifying and manipulating the machine-readable records."

Based on experiences of the staff at the United States National Archives and Records Administration (NARA), this description does not apply. The crux of Terry Cook's article is that the first generation of electronic records archives consisted of only statistical and survey machine-readable data files. While NARA did acquire or accession statistical and survey files, the agency also accepted many transfers beyond this narrow genre of records. Prior to 1980, probably the most historically significant collection of electronic records transferred to NARA were those created by the United States military during the course of the Vietnam War. The Vietnam War was the first war fought by computer and, thus, the prosecution of the war generated computerized records. These records were central to the military's war effort and therefore were part and parcel of the core mission of Department of Defense components. During the Vietnam War, unit commanders compiled operational information which was coded and entered into a computer for analysis. Thus during the late 1970s, NARA acquired files which have individual records for each military engagement with either Viet Cong or North Vietnamese forces, and have individual records for each air sortie flown over the combat area, with individual records for each naval action in Southeast Asia, whether naval gunfire or the mining of Haiphong Harbor. Furthermore, NARA acquired electronic records which documented the effectiveness of these military actions, such as evaluating the extent of the pacification effort, the effectiveness of the South Vietnamese military, plotting Communist Base Area locations by geographical coordinates, and estimates on the loss of materiel and personnel on both sides of the conflict. The documentary electronic materials which the Department of Defense sent to NARA in the late 1970s clearly disproves the statement: "Aside from...social science data files, the only business applications being automated [before the mid-1980s] 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."

During the 1970s, U.S. regulatory commissions adapted the computer to their regulatory functions. Since a regulatory agency normally controls a given commercial activity, it routinely acquires information from every establishment engaged in the regulated activity. The regulatory body uses reported information to determine whether an individual business establishment is complying with the various regulations. Agencies computerized such records in order to assist in analyzing each submission and in identifying those reports which deserved closer scrutiny. Two regulatory agencies engaged in such automated support were the Securities and Exchange Commission, which required registration of each offer of securities for sale to the public, and the Federal Home Loan Bank Board, which required monthly, quarterly, and annual reports from federally-chartered Savings and Loans Associations. In the late 1970s, these regulatory agencies began transferring such records to NARA. These computerized databases had been assembled as part of the basic regulatory function—not as some housekeeping function like payroll or accounts receivable.
Beyond these databases at the core of agency missions and functions, the U.S. government greatly expanded the use of the computer in the 1970s beyond number crunching, and records resulting from these expansions came to NARA. For example, NARA acquired files which were precursors of today’s geographic information systems. As Theodore J. Hull of the staff of NARA’s Center for Electronic Records reported at a recent meeting of the Association of American Geographers, “Computer cartographic techniques were still in their infancy [prior to 1980]...and the technical differences between preserving flat files and computer cartographic data files were nil. In fact, one of the largest early transfers of computerized, or electronic, records to the National Archives were 865 data sets from the Nautical Chart Data Base of NOAA [National Oceanic and Atmospheric Administration] in 1976 and 1979.” Similarly, many of the files from the military documenting the Vietnam War contain the UTM (universal transverse mercator), which permitted spatial analysis of the combat and pacification activity.

Text information was also part of the early automation efforts and was also transferred to NARA as electronic records. For example, NARA holds the records of the Watergate Special Prosecution Force (WSPF), whose investigation led to the successful prosecution of several high-ranking officials of the Nixon Administration and the resignation of the President himself. In 1974, the WSPF indexed the textual evidence by witness, names mentioned, dates of testimony and incident, and subjects using a controlled thesaurus. Each index entry contained an abstract of the evidence. The purpose of the index, as discussed in the report of the WSPF, was to manage the textual evidence being assembled by the prosecutor’s staff. The index was a primitive litigation support system whose sophisticated descendants are proliferating throughout law offices, court rooms, and government buildings across the continents.

Another precursor of the technology to come was the Consistency Audit Data Base developed by the Presidential Clemency Board, now held by NARA. President Ford appointed this board to make recommendations about clemency for individuals who, during the course of the Vietnam War, violated the Military Selective Service Act or who were sentenced or discharged for violating certain articles of the Uniform Code of Military Justice. Since clemency was frequently contingent on the performance of alternative service of varying duration, the board established an automated system to ensure consistency in the length of alternative service among those individuals who shared the same aggravating and/or mitigating circumstances. In operation during 1975, the system moved from a post facto audit of the board’s recommendation to proposing an a priori recommendation to the board for the length of alternative service based upon “rules” established by the board in previous cases. As this Consistency Audit Data Base evolved into making recommendations to the Clemency Board, it became a precursor to today’s artificial intelligence systems or expert systems.

Indeed, in 1979, NARA acquired 849 data sets with highly software dependent digital images. These are examples of what Cook describes as a harbinger of things to come—documents “where text, graphics, images, and voice are converted to electronic format”—but which NARA acquired over a decade ago. One should note that the electronic records discussed above reflect over half of the corpus of records which NARA accessioned before 1980. During that time, NARA acquired 118
accessions of electronic records. Of these, seventy (59.3 per cent) were programmatic records or records derived from programmatic operations. While most were programmatic databases or automated case files, records containing non-numerical information formed a portion of the holdings of NARA's electronic records programme from its earliest days.

The Cook article states that “the techniques of the library world were...adopted to describe or catalogue the early machine-readable data files in archives.... [D]ata files were treated as publications, their contextual relations...being either secondary or non-existent compared to highlighting their information content as discrete bibliographic units.” In light of the activities at NARA and within the United States, this statement needs clarification. NARA's descriptive practices described electronic records according to the format in use for all archival records and then described them a second time in a data-file format whose standards are compatible with standards for archival records which have since emerged. Throughout the 1970s and early 1980s, NARA staff described electronic records in the format that the centralized descriptive system prescribed for all series accessioned into the National Archives, namely the “NARS A-1” system. As NARA was designing and implementing “NARS A-1” inclusive of its electronic records, external standards were emerging for describing machine-readable data files through the efforts of social science data archivists and librarians. As a result of a blue ribbon meeting on description of electronic materials being generated by federal agencies, the Office of Federal Statistical Policy (OFSP) proposed in March 1979 descriptive standards for federal machine-readable data files. The standards explicitly recognized that bibliographic control developed for print publications was inadequate for electronic data files; rather, abstracts with clearly defined data elements were also required. As an agency of the U.S. government, NARA began describing its accessioned electronic records, regardless of informational content, according to the proposed standard while continuing “NARS A-1” efforts. The OFSP standard was subsequently compatible with the Machine-Readable Cataloguing (MARC) standard for Machine-Readable Data Files (MRDF), which later became the MARC Computer Data File format. At a later point, beginning in the late 1970s, the Society of American Archivists launched an effort to standardize the description of archival materials through the National Information System Task Force (NISTF). Throughout its life, NISTF had a close working relationship with the Society of American Archivists (SAA) Task Force on Automated Records and Techniques. As a result, the final Archives and Manuscript Collections (AMC) format which emerged from NISTF and the SAA's Committee on Archival Information Exchange (CAIE) was essentially compatible with the MARC-MRDF format. Hence, if an institution had successfully described its electronic records in a MARC-MRDF format, it was a simple crosswalk to the AMC format.

The Archivaria article accurately described the hardware and software which supported these core databases, litigation support systems, geographical information systems, and expert decisions systems. Agencies created them on mainframe computers using batch processing. The reason for this was quite simple: in the 1970s, mainframe technology was the only way to automate the processing of the information. Clearly, if these applications were in use today, the systems staff would use relational or object oriented technology, highly sophisticated geographic information
systems, or other advanced software packages. Yet these technologies are how the Archivaria article characterizes the second generation of electronic records archives. The article implies that the distinguishing feature between the first and second generation is the technology which processed the records and not the function or context of the records. But if one looks to today’s Bureau of the Census, the U.S. major survey agency and the closest entity to a central statistical agency, one sees that it now uses relational database technology to process its one-shot survey and census information. Therefore, the technology used to process the records does not determine the nature of archival records, rather the records themselves determine the nature of archives independent of the technology which created and processed them. Unfortunately, Cook’s article confuses records with the technology which produces the records.

This synopsis of the Archivaria article is admittedly and undoubtedly overstated, for the article qualifies its contentions, admits that the “broad canvass of the history of the first generation...is undoubtedly unfair in many details,” and talks about an “evolution” of the archives over time. Further, the author uses the term “assertions” in describing the conclusions he draws about the nature of electronic records archives during the past two decades. Terry Cook’s fairness contrasts with Richard Cox’s recent monograph, The First Generation of Electronic Records Archivists in the United States: A Study in Professionalization. This study takes the assertions in the Archivaria article and accepts them as fact. The persuasive influence of the Archivaria theme on the Cox monograph can be seen in the title’s reference to the “First Generation of Electronic Records Archivists.” But rather than an “evolution,” the monograph asserts that the second generation represents “a severe break with what has occurred up to this point.” And rather than “assertions,” the monograph claims “analysis”--“an analysis of false starts, wrong approaches, experimentation, poor professional priorities, inadequate leadership, and other problems that have prevented American archivists from embarking on more meaningful research and application to preserve records with archival value in electronic form.” And so where Cook’s article threw out a concept as a context to consider the merits of some recent literature concerned with electronic records in archives, Cox’s book threw down a gauntlet.

Cox focuses on future developments in the training and education of archivists, so that they will be able to manage electronic materials professionally. The monograph first modifies Cook’s assertion that little during the early years has relevance to today’s records systems to the claim that little was done during the early years. As an explanation, the book outlines two mutually-dependent conditions: the lack of progress with electronic records and the lack of effective educational and training programmes. According to Cox, this reflects a profession which has failed to deal effectively with the information age. As evidence for this point, the author analyzed position descriptions for archivists at state archives in the United States and vacancy announcements listed with the Society of American Archivists. The purpose of this exercise was to learn what qualifications state archives were demanding of the archivists in their employ. According to the author, neither of these sources listed the skills and abilities needed to manage electronic records.

Clearly, the archival profession has not come to grips with the electronic age. This, however, does not mean that the profession has not accomplished anything
during the last quarter century during which archivists have managed computerized records. One problem with the book is its conclusions; another is its reliance on evidence which distorts NARA's involvement with electronic records.

To determine whether archivists have the knowledge and skills to manage electronic records, the research design examined position descriptions--only at state archives. On its face, the design excluded NARA. Seemingly as an effort to compensate for this omission, the book references the qualifications included in one position announcement for NARA's Center for Electronic Records. The book asserts that the vacancy announcement only listed the standard qualifications for an archivist within the federal government, essentially a bachelor's degree with work in American history, "while, in fact, the duties were directed solely to working with electronic records." The explanation contains two omissions. First, it did not explain that these are only the minimum qualifications. Secondly, and more significantly, the announcement stipulated that any applicant who met the minimum qualifications would be evaluated upon four additional criteria: (1) ability to apply archival principles to electronic records, (2) knowledge of technical properties of electronic records to accomplish their transfer into the National Archives, (3) ability to communicate orally and in writing, and (4) knowledge of data processing. Clearly, these are skills and abilities required for a position "directed solely to working with electronic records."

The analysis of position announcements relies totally on the Society of American Archivist's newsletter and then only on announcements for archivists. By limiting the research material to just these sources, the publication has lifted these announcements out of their functional context. Job announcements are not discrete entities but rather one step of a lengthy classification, recruitment, and selection process. Since the establishment of the Center for Electronic Records in 1988 through 1994, it has hired thirteen archivists through announcements in the SAA newsletter. Each announcement was for multiple entry level positions at NARA and, as such, did not specifically mention electronic records. However, during the subsequent steps in the evaluation and selection of candidates, the issue of electronic records came to the fore. During the interview of those applicants who were evaluated as highly qualified on the basis of the general qualifications for an entry level position, the applicants were asked about their experience with data processing and electronic records. In addition, when the Center has recruited and hired archivists beyond the entry-level position, it has distributed vacancy announcements through vehicles other than SAA, such as newsletters, electronic bulletin boards, and mailing lists. In this way, NARA has hired five senior archivists for positions in the Center for Electronic Records.

By limiting the sources to position descriptions and vacancy announcements for archivists, Cox's study omits consideration of many professional staff at an archival institution who are not archivists. The assumption that professional staff in an archives were all archivists is implied in Terry Cook's article when he writes about archivists performing both archival and technical processing and about the "increasingly uneasy union of the archivist and the computer technician." Far stronger than an implication, Cox's book states that "state government archives' position descriptions [for all staff] seem to reflect a very traditional set of responsibilities and activities that pre-date the pervasive use of computers." And then the study praises libraries
for the influx of new types of professional and paraprofessional staff, including programmers and systems analysts. This conclusion about the qualifications of archival institutions versus libraries is somewhat suspect, since the analysis intentionally omitted all non-archival positions and any vacancy announcement for a non-archival position. Again, the electronic records programme at NARA serves as a corrective. Since 1973, the programme has included professional ADP personnel whose expertise was technical, not archival. Furthermore, since the early 1970s, NARA has used, and continues to use in its electronic records programme, a professional job category of "archives specialist" which equates roughly to an "information specialist." In the Center for Electronic Records, these staff members have strong technical backgrounds rather than credentials in the use of primary historical sources. A look at the staffing pattern of the Center for Electronic Records shows a rich diversity in knowledge, skills, and abilities. This diversity reveals the flaw in drawing conclusions about the professional staff of archival programmes for electronic records based only on information about the archivists in those programmes.

As the Cox study took the vacancy announcements out of the functional context, and archivists out of the work place context, so too did the study take its sources out of their context. The cited sources are invariably published works from the professional archival literature. This reliance on published sources neglected the rich literature among conference presentations and among the photocopies of reports in filing cabinets in archives on both sides of the 49th parallel. This surely must explain why two giants in the field of electronic records, John McDonald and the late Harold Naugler, have only one work each in the list of cited sources. But the dependence on published works without consultation with primary sources has removed the publications from their context. This leads to misinterpretations. The most blatant example occurs in a discussion of whether electronic records can adapt to traditional archival techniques or whether they demand a new approach. A reference is made to a 1986 paper in which John Mallinson argued that electronic records should be converted to microfilm for preservation purposes. The book correctly states that the conclusion was "in agreement with the recommendations of a National Archives and Records Administration committee" and that it "is within the parameters of traditional archival approaches." As a counter, the book cites an article by Sue Gavrel which argues that electronic records should be preserved in electronic format and states that Gavrel's comments "indicate that there are inherent differences in electronic records that require new approaches and thinking by archivists." A reader would conclude that Mallinson was the traditionalist; Gavrel was the advocate of change, and NARA was sympathetic to the traditional approach. Nothing could be further from the truth. The primary source documentation indicates that Gavrel was arguing on behalf of nearly two decades of preservation practices of those archives which had established a programme for the archival control of electronic records. Furthermore, the Archivist of the United States had two years earlier emphatically rejected Mallinson's recommendation: "We feel that the concept is based upon a particular interpretation of the mission of our agency... Such a policy could weaken the foundation of the agency." Thus the primary sources reveal a misinterpretation of the published articles upon which the monograph relied. As such, it compounds the neglect of "gray literature" and calls into question one of the author's main conclusions, namely, "the professional literature seems to lack consensus on the nature
MYTH OR REALITY: IS THERE A GENERATION GAP AMONG ELECTRONIC RECORDS ARCHIVISTS?

Thus the Cox study has taken vacancy announcements out of their functional context, archivists out of the work place context, and archival publications out the context of the corporate whole of source material on archives and electronic records. The direct result is that the work takes the present out of the context of the past and leads to an unquestioning acceptance of a generational gap in the archival management of electronic records. But if the first and second generation paradigm is not appropriate, what would be an alternative model? Terry Cook was moving in the right direction when he suggested an "evolution," and that the growth has possibly been incremental rather than generational. Archivists should recognize that their profession has made progress in managing electronic records and yet understand that much more needs to be done. To answer unanswered questions, archivists must exploit the past. This means that techniques which successfully dealt with earlier technologies should be applied to emerging ones. For example, what was known about managing structured and fielded "computer-based surveys and electronic census information" has given archivists the tools to manage "large, hierarchial, networked and especially relational databases [which] are becoming the norm in business, universities, and government." These same tools are applicable for certain types of contemporary information systems with structured record formats, such as indexes to paper-based materials and geographic information systems. But these techniques are not applicable to "[c]ompound, 'smart,' or hypermedia documents...merging these individual media in ways that are complex, very software-dependent, and very difficult to recreate." Again, archivists should look to the past. As archivists in earlier years learned from social science data archivists and librarians to manage machine-readable data files, so archivists today must form partnerships with allied professions to develop the techniques which will allow the archival profession to cope with the expanding products of the information revolution.

Notes

* The author presented an earlier version of this paper at the 21st Annual Conference of the International Association for Social Science Information Service and Technology (IASSIST), Quebec, 9-12 May 1995. He is indebted to Peggy Adams, Fynnette Eaton, and Linda Henry, colleagues in NARA's Center for Electronic Records, for their thoughtful comments and suggestions. While the paper draws upon twenty years of experience with electronic records at NARA, the conclusions of the paper are the author's personal opinions and do not reflect the views of his colleagues or represent an official position of his agency.


3 Cook, "Easy to Byte," p. 205.

4 Prior to 1 April 1995, the National Archives and Records Administration (NARA) was the National Archives and Records Service (NARS), a service within the General Services Administration. While this article refers to practices prior to 1995, NARA is used throughout for the sake of consistency.


Combat Naval Gunfire Support File (CONGA), 1966-1973; Naval Surveillance Activities File (NAVSA), 1966-1972; [Information on Operations Market Time and Game Warden]; [Electronic record]; Records of the U.S. Joint Chiefs of Staff, RG 218; National Archives, College Park, MD.

Mine Warfare Operations File (MINEA), 1972-1973; [Naval Mine Warfare Against North Vietnam]; [Electronic record]; Records of the U.S. Joint Chiefs of Staff, RG 218; National Archives, College Park, MD.

Hamlet Evaluation System (HES), 1969-1974; [Electronic record]; Records of the Office of the Secretary of Defense, RG 330; National Archives, College Park, MD.

Territorial Reporting System (TFARS), 1972-1974; [Effectiveness and Readiness of the South Vietnam Territorial Forces Units at the District Level]; [Electronic record]; Records of the Office of the Secretary of Defense, RG 330; National Archives, College Park, MD.

Operations Analysis System (OPSANAL); Enemy Base Area File, 1966-1971; [Communist Base Area Locations by Geographical Coordinates and By Administrative Areas]; [Electronic record]; Records of the Office of the Secretary of Defense, RG 330; National Archives, College Park, MD.

Situation Report Army File (SITRA), 1966-1973; [Ground Combat Data]; [Electronic record]; Records of the U.S. Joint Chiefs of Staff, RG 218; National Archives, College Park, MD.

Cook, "Easy to Byte," p. 203. The above citations are only a partial representation of the electronic records which document the Vietnam conflict and which are in the National Archives of the United States. For a more complete discussion of the content and context of these records, see Margaret O'Neill Adams, "Vietnam Records in the National Archives: Electronic Records," Prologue: Quarterly of the National Archives (Spring 1991), pp. 76-84; and Donald Fisher Harrison, "Computers, Electronic Data, and the Vietnam War," Archivaria 26 (Summer 1988), pp. 18-32.

Registration Offering Statistics (ROS), 1971-1988; [Electronic record]; Records of the Securities and Exchange Commission, RG 266; National Archives, College Park, MD.


Watergate Special Prosecution Force (WSPF): File 73, Records from the Senate Select Committee; File 96: Records Collected by the WSPF; File 96: Abstracts of Documents Concerning the 1969 Gift of Pre-Presidential Papers; [Electronic record]; Records of the Watergate Special Prosecution Force, RG 460; National Archives, College Park, MD.

Case Dispositions, Computer Programs, and Demographic Survey of the Presidential Clemency Board (Audit Data Base; Applicants File, and 13 Programme Files); [Electronic record]; Records of Organizations in the Executive Office of the President, RG 429; National Archives, College Park, MD.

House Select Committee on Assassinations Computer Enhancements of Assassination Photographs, 1976-1979; [Electronic record]; Records of the U.S. House of Representatives, RG 233; National Archives, College Park, MD.


Ibid., p. 204.


The data elements of the MRDF and AMC formats continue today in NARA's internal descriptive standards for all records. More interesting, the latest incarnation of the MRDF and AMC formats appears in the mandatory core elements required by the Government Information Locator Service (GILS). As will be reported later this year, NARA archivists were heavily involved in the development of the GILS core elements and ensured compatibility with the archival descriptive standards which emerged in the late 1970s and 1980s. Because of the evolution over the past twenty years from archival
descriptive standards, the GILS records provide an opportunity for electronic records management by archivists and for access to current records by researchers.

25 This influence is all the more telling when one realizes that, according to a promotional flyer from the publisher, the original title of the monograph was *Electronic Archives: The Challenges to Archival Education and Research*.


27 Ten years ago at the conclusion of a history of SAA's early involvement with the computer, I concluded, "[I]t would be premature to propose reconstituting SAA as the Society of Atari Archivists." I am no more optimistic today as I look at the technical naïveté which leads some to embrace simplistic solutions to complex problems. Thomas Elton Brown, "The Society of American Archivists Confronts the Computer," *American Archivist* 47 (Fall 1984), p. 382.

28 Cox, *The First Generation*, p. 82.


31 Ibid., pp. 41-42.


34 The terminology comes from Cook, "Easier to Byte" p. 206.