Report for Recommendation 90-5

# Federal Agency Electronic Records Management and Archives

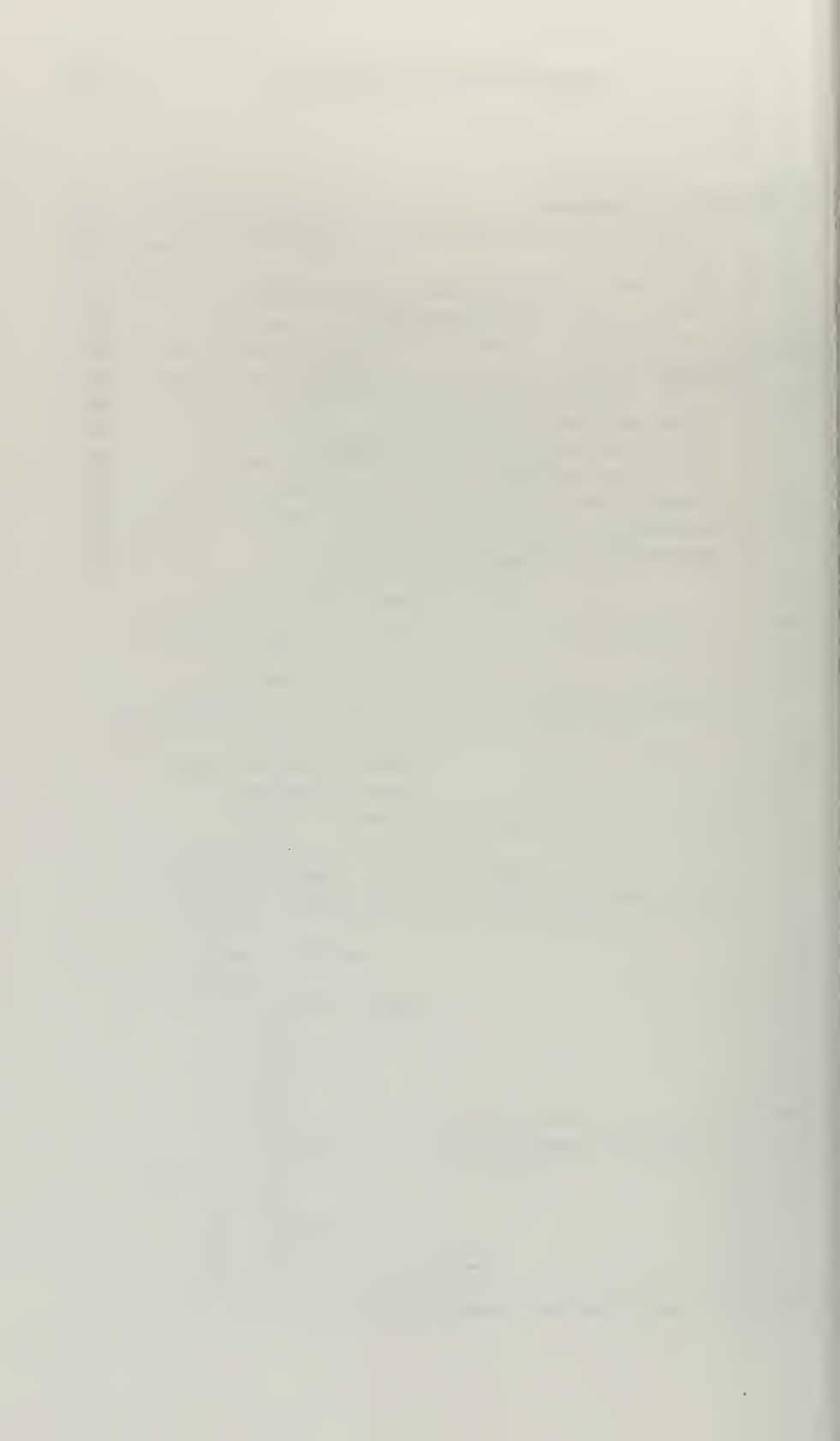**Henry J. Perritt, Jr.**
School of Law
VillanovaUniversity

# EXECUTIVE SUMMARY

A growing proportion of the universe of federal agency information is being created, processed, stored and retrieved in digital electronic form. The law of records management and preservation runs the risk of becoming increasingly obsolete unless it adapts to cover electronic information formats. So far, the policy setting and oversight agencies for records management and archives have had some time to work out appropriate long run approaches because of the feasibility of keeping paper documentation of official decisions. Now, however, the requirement to keep paper copies of electronic documents is increasingly an unsatisfactory solution because of the burdensomeness, the artificiality, and storage requirements. The use of electronic technologies makes it easier to implement appropriately designed records management and archives systems. On the other hand, it is important that the design of these systems not impede efficient use of existing and future technologies, and equally important that they not expect more of individual users of such technologies than is reasonable. Designing agency information systems to provide adequate records management capability implicates a tension between centralized and decentralized automation strategies that is pervasive in office automation. A good electronic records management program permits substantial amounts of user autonomy with respect to software selection and use patterns.

In the long run, the only significant threat to an historical record based on electronic formats is technological obsolescence--the possibility that future computer hardware and software would be unable to read files created with today's hardware and software. Mass electronic storage is not the problem; the problem is ensuring access to electronic information over periods exceeding thirty to forty years. Most people agree that the ultimate solution to the technological obsolescence threat is the development of appropriate information format standards for interchanging information. Such standards improve the integrity of the national archives, and also can improve the utility of information systems in everyday use.

This report reviews the legal environment, identifies the major questions presented by the shift to digital electronic information technologies, and develops a set of recommendations intended to shape conformity between an appropriate electronic records legal environment and emerging technologies.

# I. Introduction

## A. Viewpoint

More and more records of federal agencies are created, used, and stored in electronic rather than paper form. As this occurs, legal requirements and management efforts designed around paper records are progressively less satisfactory to ensure an adequate legal and historical record of government decisionmaking. Administrative Conference Recommendation 88-10[1] and the accompanying report address public access and dissemination of electronic information. The 88-10 effort considered statutes and policies that contemplate agencies releasing information they already have to the public.[2]

---

[1] Federal Agency Use of Computers in Acquiring and Releasing Information (Recommendation 88-10), 54 Fed. Reg. 5207, 5209 (Feb. 2, 1989), to be codified at 1 CFR §305.88-10 [hereinafter "ACUS Recommendation 88-10"]; *See* H. Perritt, Electronic Acquisition and Release of Federal Agency Information (Oct. 1, 1988) (report); Perritt, *Electronic Acquisition and Release of Federal Agency Information: An analysis of ACUS recommendation*, 41 Admin. L. Rev. 253 (1989); Perritt, *Federal Electronic Information Policy*, 63 Temple L. Rev. 201 (1990).

[2] Information release can occur at three levels: access, disclosure, and dissemination. *See* ACUS Recommendation 88-10, glossary; Office of Management and Budget Circular A-130, 50 Fed.Reg. 52730 (Dec. 24, 1985). Circular A-130 distinguishes between "access" (§6(f)) and "dissemination" (§6(g)), and an appendix explains the distinction essentially the same terms as those used Recommendation 88-10. 50 Fed.Reg. at 52745 (Appendix IV to Circular A-130). *Access* is the lowest level, representing the most passive form of release. The agency must release information upon request but takes no affirmative steps to release information in the absence of a request. Paper information subject to access is kept in regular agency files and indexed and packaged for routine agency use; not for public availability. The ACUS recommendations use the term electronic access to refer to the lowest level of electronic release; the ability to obtain agency information through ad-hoc demands. Electronic access usually is accomplished by releasing bulk information on tapes or disks, in the format used by the agency. Information covered by the miscellaneous records provisions of the Freedom of Information Act, 5 U.S.C. §552(a)(3) (1982), but not by other release obligations is an example.

*Disclosure* is an intermediate level of release, involving some affirmative effort by the agency to make the information easily available to the general public. Regulatory dockets, SEC filings, and indices of adjudicatory decisions all are examples of information that is disclosed. Most agencies meet disclosure requirements by providing public reference rooms. Electronic disclosure involves making terminals and suitable retrieval software available in public reference rooms, and possibly at other fixed locations. The ACUS recommendations use the term electronic disclosure to refer to an intermediate level of electronic release; making information available electronically to the public at one or only a few places.

The recommendation and report also considered agency acquisition of information in electronic form. The 88-10 effort did not consider internal agency use of information in electronic form; nor did it consider the circumstances under which agencies should create electronic information systems or the characteristics of such systems.

This report complements the 88-10 effort. It focuses on internal agency electronic records management, and the preservation of electronic records for historical purposes.

The basic legal framework, explained in Part II, contemplates that agencies must retain certain records of public decisions so that there will be a historical record. As electronic technologies become the predominant means through which decisions are recorded, faithful application of the policies embodied in the records statutes requires assessment of whether electronic records should be treated the same or must be treated differently from paper records. This report also develops the idea, shared widely among information systems professionals, that the best way to ensure retention of electronic records having archival value is to design archival features into information systems in the first place. Accordingly, the report necessarily considers electronic information system design.

## B.  Relevance to Private Sector

It is erroneous to suppose that electronic records management and electronic archives management present issues solely of concern to creators and managers of government information. The private sector should be concerned with electronic records management issues for three reasons overlapping with

The highest level of information release involves the most agency activity: *dissemination*. This involves a high degree of affirmative agency action actually to publish the information and to distribute it. The National Library of Medicine and the Federal Register are examples of this activity. The ACUS recommendations use the term electronic dissemination to refer to the highest level of electronic release; using electronic means to make information widely available to the public at places where it is used. Electronic dissemination involves making available dialup links or disks containing data structures and software for easy retrieval on small computers. Electronic publishing is the same thing as electronic dissemination.

The stages of added value, discussed in §IV(B)(2) of this report, elaborate on these three levels of value.

It is difficult technologically to draw clear lines among the three levels once information is computerized, (*See* ACUS Report §VI(C) explaining why) but the distinctions nevertheless are useful in evaluating policy options. Paragraph C of ACUS Recommendation 88-10 makes use of these three levels of release.

the subjects discussed in this report. First, the government imposes certain recordkeeping requirements on the private sector that may or may not be met by electronic recordkeeping strategies. Second, private entities want to maintain records for their own archiving and litigation support functions. Third, some of the same issues regarding reliability of official records kept in electronic form also are raised in connection with private electronic contracting methods.[3]

The federal government imposes diverse records keeping requirements on private sector entities, as aspects of economic regulation by the tariff regulatory agencies, as aspects of environmental and health and safety regulation and food and drug regulation, and regulations imposed on defense contractors.[4] OSHA requires that hazardous substance exposure record for employees be maintained for 50 or more years.[5] The same technologies that permit electronic archiving of government information also permit electronic maintenance of private sector information subject to governmental requirements. Most of the same policy issues concern both private and public information, and the same procedures that make sense for government records managers also may make sense for private sector records managers operating under governmental requirements.

Purely private incentives also raise electronic archiving issues for the private sector. Private sector entities have an incentive to maintain records in connection with potential litigation, such as medical records, OSHA-related hazardous substances records, and banking records.

A growing number of private and public entities are using electronic messages to make and administer contracts and to engage in other legally significant transactions. One important requirement for these electronic

---

[3] A growing number of private companies are using electronic messages to conduct commercial business affairs, for example electronic invoices and purchase orders exchanged directly between computers. These electronic contracting activities use a family of format standards developed by ANSI accredited standards organization X12 and commonly known as Electronic Data Interchange ("EDI.")

[4] See *Guide to Recordkeeping Requirements in the Code of Federal Regulations* (published by Office of the Federal Register, National Archives and Records Administration, as a guide to recordkeeping requirements that the federal government imposes on private identities).

[5] 29 CFR §1910.1001 (1988) (requiring employers to keep records of exposure to certain substances like asbestos for period of employment plus thirty years); 29 CFR §1910.1029 (1988) (requiring employers to keep records of employee exposure to coke oven emissions for 40 years, or period of employment plus 20 years, whichever is longer); 29 CFR §1910.20 (1988) (requiring employers to employee exposure and medical records for 30 years); 29 CFR §1904.2 (1988) (requiring employers to keep records of reportable occupational injuries and illnesses for five years). See also 29 CFR §1951.7 (1988) (permitting state agencies to keep certain OSHA mandated records in microform).

contracting systems is that reliable records of the transactions be maintained in the event of a dispute over whether the transaction occurred, whether a contract was formed, or over the terms of the contract.[6]

The issues regarding the rules of evidence and reliability discussed in Part IV(H) of this report are material to reliability concerns as to private sector electronic contracting.

## C. Development of Report

### 1. Entities consulted

The author consulted the following interests:[7]
Office of Management and Budget
Staff of the House Committee on Government Operations
National Archives and Records Administration
Electronic information industry
Public interest groups
ASC X12 members and officers
Individual agencies

### 2. Acknowledgments

The author particularly appreciates assistance from the following people:
Richard Berry, Chief of the Information Services Division of the World Bank
Tora Bickson, Rand Corporation
Richard J. Ciamacca, research assistant, Villanova University School of Law, Class of 1991
Charles M. Dollar, Assistant Director, Archival Research and Evaluation Division, NARA
Robert Gellman, Majority Counsel, Subcommittee on Information, House Government Operations Committee
Patti A. Goldman, Public Citizen Litigation Group
David Johnson, Wilmer, Cutler & Pickering
Sarah T. Kadec, Consultant

---

[6]Electronic Messaging Services Task Force, *The Commercial Use of Electronic Data Interchange--A Report*, 45 Bus. Law. 1645 (1990).

[7]Individuals who provided information or otherwise assisted in the preparation of this report are listed separately in the acknowledgment section of this part.

John McDonald, Canadian National Archives
Ronald L. Plesser, Piper and Marbury
Timothy Sprehe, OMB Office of Information and Regulatory Affairs
Kenneth Thibodeau, Director, Center for Electronic Records, NARA
Mary Ann Wallace, Director, Agency Services Division, National
     Archives and Records Administration ("NARA")
Peter Weiss, OMB Office of Federal Procurement Policy

# II. Current Legal Environment

## A. Institutional Responsibility

### 1. National Archives and Records Administration ("NARA")

The National Archives and Records Administration (NARA) was created by the National Archives and Records Administration Act of 1984[8] as an independent executive agency of the United States Government. NARA succeeds the National Archives Establishment, which was created by act of June 19, 1934.[9] The National Archives Establishment was subsequently incorporated into the General Services Administration (GSA) and renamed as the National Archives and Records Service by §104 of the Federal Property and Administrative Services Act of 1949.[10]

Under existing law, the Archivist of the United States has the authority to accept federal records and private materials that the Archivist has determined to have sufficient historical or other value to warrant their continued preservation by the United States Government. The Archivist then causes the transfer of accepted records[11] to the National Archives. As a general matter, the National Archives and Records Administration handles records according to agency instructions, subject to FOIA obligations imposed on NARA.[12]

---

[8] 98 Stat.2280, 2295 (Oct.19,1984), amending 44 U.S.C. §2102 and following sections.

[9] 48 Stat. 1122-1124 (Jun. 19, 1934).

[10] 63 Stat. 377-403 (Jun. 30, 1949). *See generally* U.S. Government Manual, 1988-89, p. 612.

[11] A common term for the acceptance and transfer process is "accession."

[12] *See* 36 CFR §1250.16(b) (if NARA received request for records that are primary responsibility of another agency, NARA refers the request to the agency); 36 CFR §1228.198(a)

Only permanent federal records defined as such on an SF115 (request for records disposition authority) and approved by the Archivist, are transferred to the National Archives.[13] Records already deposited into the National Archives may be designated for transfer to educational institutions or associations.[14] The Archivist is authorized to promulgate regulations establishing procedures for preserving records.[15]

## 2. Agencies

Federal agencies work with the Archivist to determine what records will be deposited into the National Archives. Each agency has a records manager who is responsible for (1) inventorying all records in the agency, (2) proposing disposition instructions for approval by NARA, and (3) creating a comprehensive agency records disposition schedule. This schedule is submitted to NARA for approval.[16]

## B. Statutes and Executive Orders

## 1. Disposal of Records Act

The Disposal of Records Act[17] authorized agencies to submit schedules of records[18] not currently needed "that appear to have no permanent value or historical interest"[19] to the Archivist, authorized the Archivist to report to a joint Congressional committee lists of records to be destroyed, and authorized agency heads to destroy records, as to which the joint committee had no

---

(NARA will observe restrictions lawfully imposed on the use of records transferred from other agencies "to the extent that they do not violate 5 U.S.C. §552)."

[13]36 CFR §1228.190.

[14]44 U.S.C. §2107 (1989). This provision has been used only when institutions already had collections of federal records that were approved for deposit, and NARA allowed the records to remain where they were as "satellite archives."

[15]44 U.S.C. §3302 (1988).

[16]36 CFR §§1228.12 and 1228.2 (1988).

[17]53 Stat. 1219-1221 (Aug. 5, 1939).

[18]Records was subsequently defined to include all "books, papers, maps, photographs, or other documentary materials, regardless of physical form or characteristics . . . ." 57 Stat. 380-383 (July 7, 1943). 44 U.S.C. §3301 now defines records to include machine readable materials.

[19]53 Stat. 1219 enacting clause.

objection.[20] Legislation enacted in 1940[21] authorized the destruction of records that have been microfilmed in accordance with technical standards developed by the National Bureau of Standards, and made the microfilm records admissible in evidence.[22] Legislation originally enacted in 1943[23] authorizes the Archivist to promulgate regulations establishing procedures for preserving[24] and disposing of records,[25] and requires agencies to submit records disposition schedules to the Archives, in accordance with Archives regulations.[26]

## 2. Executive Order 9784

Executive Order 9784[27] required agency heads to establish programs for effective management and disposition of agency records.

## 3. Federal Property and Administrative Services Act

The Federal Property and Administrative Services Act of 1949[28] transferred the National Archives Establishment to the General Services Administration,[29] and authorized the Administrator of General Services to obtain reports from federal agencies on records management and disposal practices, to "promote, in cooperation with the executive agencies, improved records management practices and controls . . . including the central storage or disposition of records not needed by . . . agencies for their current use" and to report to the Congress and to the Director of the Bureau of the Budget on the subject.[30]

---

[20] Id. §5.

[21] 54 Stat. 958-959 (Sep. 24, 1940).

[22] §2, 54 Stat. 958. 44 U.S.C. §3312 now makes "certified reproductions" admissible in evidence.

[23] 57 Stat. 380-383 (July 7, 1943), 44 U.S.C. §3302 (1988).

[24] Preservation was handled implicitly by authorizing the disposition of records not having "sufficient administrative, legal, research, or other value to warrant their further preservation." §3, 57 Stat. 380 (July 7, 1943).

[25] §2 (authorizing National Archives Council); 44 U.S.C. §3302.

[26] §3; 44 U.S.C. §___. Congressional involvement was deleted from the Act in 1970.

[27] Sep. 25, 1946.

[28] 63 Stat. 377-403 (June 30, 1949).

[29] Id. §104(a).

[30] Id. §104(c).

## 4. Federal Records Act

The Federal Records Act[31] authorized the Administrator of General Services to promote standards for improved records management[32] and to "establish standards for the selective retention of records of continuing value."[33] It required agency heads to develop records management programs, to preserve records "containing adequate and proper documentation of the organization, functions, policies, decisions, procedures, and essential transactions of the agency,"[34] and to establish safeguards against removal or loss of records under regulations promulgated by the Administrator of General Services.[35] The Act authorized the Administrator to accept for deposit with the National Archives agency records determined by the Archivist "to have sufficient historical or other value to warrant their continued preservation . . . ."[36]

## 5. Federal Records Management Amendments of 1976

The Federal Records Management Amendments of 1976[37] reorganized and codified the records management statutory provisions,[38] to clarify the objectives of records management to include:

-- accurate and complete documentation of government policies and transactions

-- control of quantity and quality of government records

-- simplification of records management activities, emphasizing the prevention of unnecessary paperwork

-- judicious preservation and disposal of records[39]

Legislation enacted in 1978 made application of general records schedules to agencies mandatory.[40]

---

[31] 64 Stat. 583-591 (Sep. 5, 1950), adding §§501-511 Federal Property and Administrative Services Act of 1949.

[32] §505(a), 64 Stat. 583.

[33] §505(b), 64 Stat. 583.

[34] §506(a), 64 Stat. 583.

[35] Id. §506(e).

[36] Id. §507.

[37] 90 Stat. 2723-2727 (Oct. 21, 1976).

[38] Id., revising 44 U.S.C. §§2901-2907, 3103.

[39] 44 U.S.C. §2902.

[40] 92 Stat. 1063 (Oct. 10, 1978), amending 44 U.S.C. §§1503 and 3303a(b).

## 6. National Archives and Records Administration Act

The National Archives and Records Administration Act of 1984[41] removed the National Archives from the General Services Administration and set it up as an independent establishment in the executive branch called "The National Archives and Records Administration."[42] The Act authorized the Archivist to issue regulations, and required agency heads to "issue such orders and directives as such agency head deems necessary to carry out such regulations,"[43] thus preserving the divided responsibility between the National Archives and individual agencies.

## 7. Presidential Records Act

The Presidential Records Act of 1978[44] establishes governmental ownership in presidential records, and obligates the President to implement records management controls to "assure that the activities, deliberations, decisions, and policies that reflect the performance of his constitutional, statutory or other official or ceremonial duties are adequately documented."[45] The Act permits the President to dispose of presidential records no longer having "administrative, historical, information, or evidentiary value," but only if the Archivist of the United States has no objection, or over the objection of the Archivist if the President submits a schedule of records to be destroyed 60 days before the destruction.[46] Upon the conclusion of a Presidency, the Act makes the Archivist the custodian of the presidential records of the departing President, and obligates the Archivist to make the records available to the public, subject to restrictions on access imposed by the President under specific categories set forth in the Act.[47] The categories of authorized restrictions closely track FOIA exemptions.[48]

---

[41] 98 Stat. 2280-2295 (Oct. 19, 1984).

[42] 44 U.S.C. §2102.

[43] 44 U.S.C. §2104 (a).

[44] 92 Stat. 2523 (Nov. 4, 1978), codified as amended by 98 Stat. 2287 (Oct. 19, 1984) to 44 U.S.C. §§2201-2207 (1982 & 1982 Supp. V 1987).

[45] 44 U.S.C. §2203 (a).

[46] 44 U.S.C. §§2203(c), 2203(d).

[47] 44 U.S.C. §§2203(f)(1) (making Archivist custodian and imposing obligation of public access), 2204 (authorizing Presidential restrictions under enumerated categories). *See* Exec. Order 12667, 54 Fed.Reg. 3403 (Jan. 18, 1989) (providing for coordination among Archivist, present and former President on release of records of former President).

[48] Compare 5 U.S.C. §552(b)(1) -(6) with 44 U.S.C. §2204(a)(1)-(6).

The Act's provisions were written to follow the guidelines established for Congressional regulation of Presidential records established by the Supreme Court in *Nixon v. Administrator of General Services*.[49]

## 8. FOIA

The Freedom of Information Act ("FOIA")[50] does not require preservation of agency records.    The FOIA requires that the public be given access to records but does not require them to be maintained, preserved, or created in the first place. The FOIA arguably can be used indirectly to require the preservation of electronic files,[51] however. If an FOIA requester requests access to a large number of electronic files, a reasonable inference is to require the agency possessing the records to preserve them until its FOIA access obligations are determined. Those files that are exempt from public access under the FOIA could be destroyed.

The requester would argue that the Freedom of Information Act would be defeated by permitting an agency to  destroy records that eventually would be subject to access under the Act.   Under this argument, an injunction against destruction or erasure would be necessary to permit FOIA rights of access eventually to be realized.  This is essentially the rationale adopted by the D.C. Circuit in *American Friends Service Committee v. Webster*.[52]

The absence of explicit linkage between the FOIA and records statutes creates problems for sound records management. Agencies are concerned that declaring something to be a record for records management purposes will increase the likelihood of access under the FOIA.  They tend, therefore, to be underinclusive in their records management practices.  Ironically, the FOIA definition of agency record has been interpreted more broadly than NARA interprets the records definition under the records statutes.

## 9. Paperwork Reduction Act

The current text of the Paperwork Reduction Act makes records management a part of information resources management and gives OMB authority to review compliance with records management laws and

---

[49]433 U.S. 425 (1977).  *See* H.R.Rep. 95-1487, 95th Cong., 2d Sess. at 6, reprinted in 1978 U.S. Code Cong. & Admin. News 5732, 5737 (discussing case and justifying provisions of bill that became Presidential Records Act).

[50]5 U.S.C. §552 (1988).

[51]Or paper files.

[52]720 F.2d 29, 69 (D.C.Cir. 1983) (affirming district court order compelling development of adequate records disposal plan and enjoining destruction of records until plan developed; finding standing and reviewability). *See* NARA Report at 20 (suggested language).

regulations.[53]  A bill to extend the Paperwork Reduction Act was developed in the current Congress, but failed to pass before the Congress adjourned. A December 1989 version of H.R. 3695 amends the federal records act to authorize the Archivist of the United States to issue regulations establishing standards for interpreting the term "records."[54]

## 10.  Privacy Act

The Privacy Act[55] has little effect on electronic records management practices, except that it has marginally improved agency practices by creating an additional incentive to designate certain electronic information systems as containing records, thus making NARA aware of them.

## C.  Regulations, Schedules, and Guidelines: Existing and Proposed

## 1.  OMB/OIRA

OMB has directed that agencies should incorporate records management and archival considerations in the design, development, and implementation of electronic information collection systems in accordance with the Federal Records Act.[56]

## 2.  GSA

The General Services Administration ("GSA") has issued regulations identical to those issued by NARA, discussed in the following section.[57]

---

[53]44 U.S.C. §§3504, 3513.

[54]Typescript bill (November 17, 1989).  H.R. 3695 §205(a) (amending 44 U.S.C. §3302 authorizing new paragraph 4 to paragraph 3 to authorize promulgation of standards for interpreting the definition of the terms 'records' and standards to be incorporated into recordkeeping requirements by heads of agencies).

[55]5 U.S.C. §552a (1988).

[56]52 Fed.Reg. 29454 (Aug. 7, 1987) (paragraph h), referring to 44 U.S.C. §§29, 31, and 33.

[57]55 Fed.Reg. 19221 (May 8, 1990) (to be published at FIRMR 201-45.2).

## 3. NARA

### a) Regulations

NARA revised its regulations addressing electronic records on May 8, 1990.[58]

On May 8, 1990, NARA published a final rule revising its regulations on electronic records.[59] The rule established procedures for management of electronic records, for selection and maintenance of electronic storage media and for compliance with legal requirements for disposition of electronic records. Identical regulations were issued simultaneously by the General Services Administration.

The new NARA rules, like longstanding statutory and regulatory records management policies, impose on individual agencies the primary responsibilities for development and implementation of electronic records management plans. Under the regulations, the head of each Federal agency must ensure that his or her agency manages its electronic records so as to satisfy certain requirements:

--  assigning responsibility for an agencywide program for the management of all records created, received, maintained, used, or stored on electronic media, and notifying NARA of the name and title of the person assigned the responsibility.

--  integrating the management of electronic records with other records and information resources management programs of the agency.

--  incorporating electronic records managemento objectives, responsibilities, and authorities in pertinent agency directives and disseminating them throughout the agency as appropriate.

--  establishing procedures for addressing records management requirements, including recordkeeping requirements and disposition, before approving new electronic records systems or enhancements to existing systems.

--  ensuring that adequate training is provided for users of electronic records systems in the operation, care, and handling of the equipment, software, and media used in the system.

--  developing and maintaining up-to-date documentation[60] about all electronic records systems adequate for specifying technical characteristics for reading or processing records, identifying

---

[58] 55 Fed.Reg. 19216 (May 8, 1990) (revising 36 CFR Part 1234).

[59] 55 Fed.Reg. 19216 (May 8, 1990) (revising 36 CFR Part 1234).

[60] See 36 CFR §1234.20(b) (as revised May 8, 1990).

defined inputs and outputs, defining the contents of files and records, determining restrictions on access and use; defining conditions and rules for adding changing or deleting system information and ensuring the timely, authorized disposition of the records.

-- specifying the location, manner, and media in which electronic records will be maintained to meet operational and archival requirements, and maintaining inventories of electronic records systems to facilitate disposition.

-- developing and securing NARA approval of records disposition schedules, and ensuring implementation of their provisions.

-- specifying the methods of implementing controls over national security-classified sensitive proprietary, and Privacy Act records stored and used electronically.

-- establishing procedures to ensure that the requirements of this part are applied to those electronic records that are created or maintained by contractors.

-- reviewing electronic records systems periodically for conformance to legal requirements and to determine if the records have been properly identified and described, and whether the schedule descriptions and retention periods reflect the current informational content and use.

Revised 36 CFR §1234.22 specifically addressed creation and use of text documents. It requires that systems maintaining the official file copy of text documents on electronic media to meet the following requirements:

-- identify electronic versions of documents appropriately[61]

-- provide a method for all authorized users of the system to retrieve desired documents, such as an indexing or text search system;

-- provide an appropriate level of security to ensure integrity of the documents;

-- provide a standard interchange format when necessary to permit the exchange of documents on electronic media between agency computers using different software/operating systems and the conversion or migration of documents on electronic media from one system to another; and

---

[61]Appropriate identifying information for each document maintained on the electronic media may include: office of origin, file code, key words for retrieval, addressee (if any), signator, author, date, authorized disposition (coded or otherwise), and security classification (if applicable). Agencies shall ensure that records maintained in such systems can be correlated with related records on paper, microform, or other media. §1234.22 (b) (as revised May 8, 1990).

--      provide for the disposition of the documents including, when necessary, the requirements for transferring permanent records to NARA.

The NARA regulations show strong concern about the permanency of record media. Revised §1234.28 (Selection and Maintenance of Electronic Records Storage Media) requires consideration of the following factors in selecting a storage medium or converting from one medium to another:

--      the authorized life of the records, as determined during the scheduling process;
--      the maintenance necessary to retain the records;
--      the cost of storing and retrieving the records;
--      the records density;
--      the access time to retrieve stored records;
--      the "portability" of the medium to permit the medium to be used on equipment offered by multiple manufacturers and to permit transfer the information from one medium to another (such as from optical disk to magnetic tape); and
--      whether the medium meets current applicable Federal Information Processing Standards.

Agencies must ensure that all authorized users can identify and retrieve information stored on diskettes, removable disks, or tapes by establishing or adopting procedures for external labeling.[62]

Agencies must ensure that information is not lost because of changing technology or deterioration by converting storage media to provide compatibility with the agency's current hardware and software. Before conversion to a different medium, agencies must determine that the authorized disposition of the electronic records can be implemented after conversion. NARA declined to amend §1234.30(b) (retention of electronic records) to require transfer of a copy of the software and operating system in order for the records to be read at the Archives. Transfer of the software and operating system, whenever necessary, can be specified by NARA when approving a records schedule.

NARA had received comments from 18 federal agencies, two private organizations, and one member of Congress on its draft regulations. Most commenters supported the concepts embodied in the regulations, while suggesting less detail. Because the detailed requirements apply only to unscheduled or permanent electronic records, NARA noted in its preamble that agencies could reduce the burdens of compliance by scheduling records as soon as possible, thus reducing the universe of records subject to the detailed

---

[62]36 CFR §1234.28(d) (as revised May 8, 1990).

requirements. Furthermore, NARA also urged agencies to transfer permanent electronic records to NARA as soon as possible after their creation, further reducing burdens on agencies.

NARA rejected a comment recommending a new paragraph related to electronic signatures, on the grounds that it was premature to issue regulatory guidance on this subject. NARA noted objections to its proposed ban on floppy disks for permanent storage, and modified the language to give agencies more discretion,[63] while expressing continued concern about careless handling leading to loss of data stored on floppy disks.

### b) Schedule 20

NARA has created general records schedules which govern the disposition of certain records common to many agencies.[64] General Records Schedule 20 deals with the disposition of Machine-Readable Records.[65] Instructions as to how to transfer and handle machine-readable or other magnetic media are contained in 36 CFR §1228.188 and 1234.4 (1988). These records include magnetic tape, disk packs, magnetic drums, punch cards and optical disks. The general schedule classifies records for disposition as (1) master files, (2) processing files and (3) documentation files.

Master files are databases containing statistical, scientific and other information. Databases include economic, social and political data as well as natural resources data, emergency operations and national security information. Other master files include indexes, summary files, backup files, technical reformat files and housekeeping files that facilitate use of the substantive content of master files. If information is created or received by an agency of the government in the performance of its duties under federal law, the material must be maintained in accordance with the instructions of the Archivist. Similar types of information created under contract and in the possession of the contractor must be transferred to the National Archives.

Documentation files are generally maintained to facilitate the use of processing files. The files are maintained if the corresponding master file is to be maintained. These types of files include data systems specifications, user guides and information retrieval.

---

[63] "(c) Agencies should avoid the use of floppy disks for the exclusive long-term storage of permanent or unscheduled electronic records" contrasts with more mandatory language in other sections of the regulation.

[64] 36 CFR §1228.22 (1988).

[65] §20.

Processing files are intermediate files used in creation of master files and generally do not qualify for long term retention. These files include work, input and valid transaction files.

### c) NARA Bulletins

NARA Bulletin 87-5 offers guidance to agencies on management of electronic records, noting that whether information is maintained in electronic form or not makes no difference in the operation of records management obligations, but it may heighten the risk of data loss, while creating the potential for more transparent and easier extraction of records of historical value for eventual transfer to NARA. The bulletin recommends including records management objectives as a part of information system design, and offers practical suggestions for agency users of electronic technologies to minimize data loss and facilitate continued accessibility of information as technologies change. It contains a chart comparing various media, and notes that information must be written to magnetic tape or paper before it is transferred to NARA.

NARA Bulletin Number 89-2[66] reminds agency heads of their legal responsibilities to ensure the security and integrity of federal records and to safeguard against unauthorized disposition, including guidance for personal papers.

NARA Bulletin Number 88-8[67] rejects optical disks as media for NARA excision of permanent records "because of the hardware and software dependence of current optical disk systems and the absence of standards that ensure affordability of data to one system to another." Paragraph 6 of the Bulletin reports that industry standards for optical disk systems are now under development. The Bulletin also explicitly permits agencies to use optical disk media for storage and retrieval of permanent records while the records remain in an agency's legal custody, although no permanent records may be destroyed after copying onto an optical disk without NARA's approval.

NARA Bulletin 88-5[68] provides guidance to agencies regarding acquisition and authorized disposition of data created or maintained on behalf of the government by contractors. The Bulletin suggests that agencies write contracts to require the delivery "of all pertinent documentation of how the contractor carried out the program, which may include transaction or case files, handbooks, directives, procedural statements, and other information created by the contractor," to the agency. The Bulletin also suggests that agencies obtain

---

[66]November 14, 1988, scheduled to expire November 14, 1990.

[67]September 19, 1988, scheduled to expire September 30, 1990.

[68](May 20, 1988), scheduled to expire May 31, 1990.

contractor-created data that may have "reuse value to the government." In the category are background data to statistical analysis reports that represent the official position of the agency or that are required by statute, production elements needed for reproduction of audio visual products including negative and magnetic soundtracks, original drawings, and other research data.

## D. Judicial Decisions

### 1. Kissinger v. Reporters Committee

In *Kissinger v. Reporters Committee*,[69] the Supreme Court held that there is no private right of action to enforce the Presidential Records Act or, by analogy, the Federal Records Act.

### 2. AFSC v. Webster

*American Friends Service Committee v. Webster*[70] affirmed in part an injunction against destruction of FBI records until the agency formulated a records retention and disposal plan meeting the requirements established by NARA.[71] The court of appeals held that the federal courts have jurisdiction to review agency action for compliance with statutory standards on records disposal and preservation,[72] and that "private researchers and private parties whose rights may have been affected by government actions" have standing to seek review of agency destruction of records in violation of statutory provisions.[73] It distinguished *Kissinger* because of its different context, especially relying on footnote 5 in that case,[74] which expressly left open the possibility of APA review.

---

[69] 445 U.S. 136 (1980).

[70] 720 F.2d 29 (D.C.Cir. 1983).

[71] NARA then was the National Archives and Records Service ("NARS") within the General Services Administration.

[72] 720 F.2d at 45.

[73] 720 F.2d at 57.

[74] 720 F.2d at 52, citing Kissinger v. Reporters Committee, 445 U.S. at 150 n.5.

## 3. Armstrong v. Bush

*Armstrong v. Bush*[75] involved a suit by public interest groups to enjoin erasure of electronic mail files (created and stored on the "PROFS" system) at the end of the Reagan administration. The district court denied the government's motions to dismiss and for summary judgment. It concluded that judicial enforcement of the Federal Records Act and Presidential Records Act under the Administrative Procedure Act does not offend separation of powers. While the court was bound by the Supreme Court's determination in *Kissinger v. Reporters Committee*,[76] that there is no private right of action to enforce the records acts the *Armstrong* court, like the *AFSC* court, nevertheless concluded that these statutes impose nondiscretionary duties, the performance of which can be reviewed under the general provisions of the Administrative Procedure Act. In reaching this conclusion, it equated the committed-to-discretion-by-law exception to Administrative Procedure Act reviewability[77] with the political question doctrine. Because the obligations imposed by the Presidential Records Act and Federal Records Act are specific and the standards for determining what records must be preserved quite specific, the court concluded that ministerial duties are involved under these statutes rather than political questions. Therefore, decisions made in connection with the PROFS tapes are reviewable.

The court also concluded that disputed issues of material fact existed necessitating denial of the government's motion for summary judgment. The government argued that any obligations imposed by the Presidential Records Act and the Federal Records Act were met because guidelines applied to the PROFS tapes required that any messages the content of which would constitute Presidential records must be reduced to hard copy. The court concluded that it was potentially irrational or arbitrary and capricious to assume that these requirements were met in all cases and that it was also arbitrary and capricious to assume that any remaining PROFS files were equivalent to telephone messages rather than written records.[78] Thus facts could be proven permitting the decision to be set aside under §706(2)(A) of the Administrative Procedure Act.

The *Armstrong* case is significant in two major respects. First, it identifies a legal theory through which records keeping obligations imposed by the two records statutes can be enforced. *Armstrong* is vulnerable on judicially

---

[75]721 F. Supp. 343 (D.D.C. 1989).

[76]445 U.S. 136 (1980).

[77]5 U.S.C. §706(2) (1982).

[78]More generally, the court suggested that the analogy between electronic mail messages and telephone conversations itself might be arbitrary and capricious.

enforcing records keeping duties on the President, but eventual reversal on this ground still would leave intact its reasoning as applied to most agencies potentially involved in retaining or discarding electronic records below the presidential level.

Second, the case tests the government's argument that electronic records are like telephone conversations and therefore need not be preserved at all.

On December 21, 1989, the district court granted the defendant's motion for certification for interlocutory appeal, and stayed further proceedings in the district court pending resolution an interlocutory appeal by the Court of Appeals. On June 5, 1990, the United States Court of Appeals for the District of Columbia Circuit granted the interlocutory appeal. The issues as presented to the Court of Appeals require it to consider whether *Webster v. AFSC* governs the PROFS case. It could conclude that the involvement of the President in the PROFS case and the Presidential Records Act issues associated therewith distinguish *AFSC* and the PROFS case. It also could conclude that *AFSC* involved the destruction of records to a greater extent than the PROFS case.

On appeal, the plaintiffs emphasized that the FOIA claim covering the records recorded on the tape would require the district court to determine whether PROFS materials are records and if so, whether they are "agency" or "presidential" records, points the Justice Department believes are judicially unreviewable questions.

## III. Agency Practices

The basic approach of the records statutes and regulations is to give programmatic agencies primary responsibility for developing and implementing appropriate records management systems. NARA has policymaking and oversight responsibility, as well as ultimate custodial responsibility for records with historical value. The records management responsibility includes a duty to ensure that the activities of the government are adequately documented, as well as a duty to ensure preservation of records that actually are created.[79]

---

[79]*See generally* 55 Fed.Reg. 27422, 27423 (Jul. 2, 1990) (36 CFR §1220.14, as amended, defining "adequate and proper documentation.")

## A. NARA

NARA already has made some regulatory changes in connection with documentation of electronic records. More detailed requirements are not likely to be developed by NARA in the near term because of its philosophy that (1) archival uses are secondary to operational agency uses of information, and (2) agencies rather than NARA should decide in the first instance what information to create and use and retain, subject to approval by NARA.[80]

NARA has a Center for Electronic Records, representing an organizational focus for electronic records programs that has been dispersed more widely in the Canadian initiatives discussed in §III(F). NARA has actively has developed new policy guidelines for electronic records, manifested by the regulation and the training and education activities for agency personnel discussed in §III(C)(3).

## B. Department of the Interior Bureau of Land Management (BLM)

The Bureau of Land Management (BLM) within the Department of Interior is developing an Automated Land and Mineral Records System ("ALMRS"), part of a larger Land Information System ("LIS.") LIS will integrate BLM's land and natural resource records, using a public land survey system to manage land resource and status information at the legal land parcel level. The LIS will integrate spatial information from a geographic coordinate database with cultural and natural resource information.[81] The result will be an automated resource data system and land conveyance and land and mineral use authorization information administered in ALMRS. BLM and NARA both expect ALMRS to be a model for federal agencies as they develop similar systems for managing, protecting, and providing disposition for records superseded by the system, the records created by the system, and the records that document system implementation. ALMRS will provide experience in identifying record types in which flat ASCII files do not work. A major goal of ALMRS is to permit graphical information as well as alphanumeric data to be managed.

---

[80]*See* 55 Fed.Reg. 27422, 27425 (Jul. 2, 1990) (amended 36 CFR §1222.32(d) requires approval of Archivist for final disposition).

[81]ALMRS has three major databases, based on case recordation, land description, and status.

A BLM/NARA work group has developed system capability specifications to be contained in an eventual RFP for the system itself.    The draft specifications require records retention capability to:
-- designate and implement retention periods for any data set, record, or file
-- delete designated temporary records only at specified times according to approved records retention and disposition schedules
-- transfer unneeded records to storage media and facilities.
Records preservation capability to:
-- preserve permanent records, both documents and data
-- transfer designated permanent records to the National Archives
-- accommodate record format established for the National Archives.
Record integrity capability to:
-- restrict totally the editing of a specific file, legal transaction, or document at a given time
-- restrict or block editing of specific data except by specific password assignment
-- leave an audit trail of entry into databases tracking when alterations were made and by whom.
During 1990, BLM is transferring manually maintained information to electronic formats and working through a pilot program in its New Mexico offices to classify records, to provide disposition standards for each category, to establish recordkeeping requirements to ensure that legal, administrative, and archival needs are met, to test the transfer of permanent records to the National Archives, to develop recommendations for standards, policies, and procedures for BLM to use during the transition from the manual to an electronic system, and to develop government-wide guidance that might be issued by the National Archives.[82]    The New Mexico test is planned for completion by the beginning of FY91, at which point test results will be evaluated by the work group.    The work group will decide whether NARA wants certain types of records, such as master title plats, historical data, and serial register pages.

The original target for contract award was August 1990.    The contractor initially would develop a detailed design in a pilot project.    The total cost of ALMRS is expected to be $240,000,000. LIS is expected to be fully operational by 1994, with the ALMRS portion implemented with commercial off the shelf software as soon as possible, and administrative, office automation, and resource data features implemented by September 1994.

Initial requests were received from McKinley County, New Mexico and Meridian Oil to interact with ALMRS electronically, the county wanting

---

[82]January 31, 1990 NAPA draft at 6.

public reading room electronic access, and the oil company wanting electronic access to its own applications for drilling permits.[83]

## C.  World Bank

The World Bank has 6,000 people using EMail.  Presently, it commits to paper matters of an official nature, but gradually is moving toward an environment in which EMail is used for official communications.  The tendency is for EMail to be used for most communications until someone says, "I want your signature," and then the author of the document prints a hard copy, signs it, and sends it.  There is a growing recognition that EMail constitutes an important part of the historical record, the record of the debate in the policy formulation process.

The World Bank is committed to work out a set of "cues" or templates for EMail, including specifications for informal notes, informal memoranda, official letters and letters to and from clients.  In addition, there is an effort to develop naming conventions so that retrieval of documents can be more precise. The World Bank is fortunate to have good programmer resources for its DEC All-in-One system.  Other agencies, however, may be better off to use communications and word processing environments that are more accessible to off the shelf indexing and library or archives retrieval software.

## D.  Forest Service

The Rand Corporation has studied information technology in the U.S. Forest Service.[84]  The study concluded that the Forest Service has successfully internalized information technology, enabling it to become the first federal agency to employ electronic documents as its official record. Generally, an early commitment to information technology paid off in user satisfaction with applications for word processing, electronic mail, and data analysis. Nevertheless, there are problems integrating local and centralized systems and problems with excessive data in central database systems.[85]  Electronic mail

---

[83] *See* NARA BLM document at B-10 to B-11.

[84] *See* C. Stasz, T. Bickson, J. Eveland, and B. Mittman, Information Technology in the U.S. Forest Service (1990) (Rand Corporation R-3908-U.S.D.A.S.F.) [*hereinafter* "Rand Report"].

[85] Rand Report at vi-ix.

was identified as the most valuable feature of electronic information technology, surpassing even word processing.[86]    The only problems with electronic mail, which was made more flexible to permit communicating across chains of command,[87] had to do with "junk mail."

Problems exist with database management, including unpredictable updates, "inability" to find needed data, and too many duplicated files.[88] Generally, Rand recommended that the forest service provide more flexibility for effective use of microcomputers and off-the-shelf software.[89]

## E.  Nuclear Regulatory Commission Licensing Support System

The Nuclear Regulatory Commission's ("NRC") Licensing Support System ("LSS") is intended to manage the information related to the licensing of receipt and disposal of high level radioactive waste, which is expected to be one of the largest administrative litigation matters ever to come before the NRC. The LSS was reviewed in the report accompanying ACUS Recommendation 88-10, and is notable also in connection with the electronic records management and archives issues addressed in this report.    To a considerable extent the LSS rule[90] sidesteps the direct question whether its electronic records and electronic records practices satisfy the requirements of the records acts and the NARA regulations. For example, proposed §2.1013 establishes procedures for the electronic submission of pleadings, and provides for the electronic transmission of Board and Commission issuances and orders, as well as for on-line access to the LSS during the hearing. It requires submission of official paper copies of matters made part of the hearing record and entered in the docket by the Secretary of the Commission, but leaves the Secretary with discretion to determine if the record copy for records act and archives purposes is the paper copy or the electronic version.    As another example, the LSS avoids obtaining the "record" version of some documents, instead entering copies into the page-image and parallel ASCII components of the system, leaving the "record" copy under the control of the submitter.

---

[86]Rand Report at 32.

[87]Rand Report at 32.

[88]*See* Rand Report at 41-43.

[89]Rand Report at 54-56.

[90]53 Fed. Reg. 44411 (Nov. 3, 1988) (proposed rule developed by rule negotiation committee), proposing amendments to 10 CFR §§2.1-2.1117 (1989).

The LSS is a good example of an electronic records system that has been built to duplicate the "official" paper information system. If the system works as it is supposed to, however, no one in her right mind would seek to retrieve the paper counterparts of the electronic files; thus LSS may serve to build confidence that, over time, requirements for paper versions of information safely can be eliminated.

## F. Canada

The National Archives of Canada has invested considerable energy in developing electronic records management and archives initiatives. For twenty years, the National Archives has supported a simple machine readable archives program. The program's initial implementation is relatively simple. Archivists appraise machine readable records for archival value, after which the records are copied onto archival quality tapes. The data are verified against documentation. The tapes are rewound each year, and the data are recopied every five years to fresh tapes. Catalogues and bulletins advertise the availability of machine readable information on tapes, and the data are maintained in a form and format which minimizes conversion costs over the long-term.[91] John McDonald, Director, Automated Information System, National Archives of Canada, has written extensively on electronic records management issues from an archivist's perspective,[92] and has generally encouraged frequent interaction and sharing among professionals concerned with archives and electronic records management.[93]

---

[91]J. McDonald, Building the "Front End" to a Machine Readable Archives Program: The Experience of the National Archives of Canada at 1 (Nov. 1989) (unpublished paper) [hereinafter "Building the Front End"].

[92]See J. McDonald, Data and Document Interchange Standards: A View from the National Archives of Canada (Oct. 1987) (paper presented to The Society of American Archivists, New York) [hereinafter McDonald 1987 Archivists Paper]; J. McDonald, Records Management and Data Management: Closing the Gap (Oct. 1988) (paper presented to The Society of American Archivists, Atlanta) [hereinafter "McDonald 1988 Archivist Paper"]; J. McDonald, The Information Resource Director System (IRDS) and the National Archives of Canada (Oct. 1989) (paper presented to The Society of American Archivists, St. Louis) [hereinafter McDonald 1989 Archivists Paper]; J. McDonald, Building the "Front End" to a Machine Readable Archives Program: The Experience of the National Archives of Canada, (Nov. 1989) (unpublished paper) [hereinafter "Building the Front End"]; J. McDonald, The National Archives of Canada and Office Systems: A Status Report (Nov. 1989) [hereinafter "Canadian 1989 Status Report"].

[93]See Building the Front End at 6 (encouraging information sharing).

The files maintained by the Canadian archives in electronic form have been relatively small (a few thousand cases stored on single tapes), and typically contain survey and other statistical data. The standard acquisition practice has been to convert data to rectangular format, recorded in EBCDIC at a density of 6250 BPI, in IBM readable formats.[94]

The Canadians are actively involved in a number of initiatives to improve the sophistication of their approaches to electronic records. For example, the Canadian OMB, "The Treasury Board," has issued a government wide directive confirming the adoption of the OSI reference model as a Canadian government standard. A preference for OSI based systems and products will become mandatory in procurement requests in the early 1990's.[95] As a follow up, the Treasury Board is developing an Application Portability Profile ("APP"), based on work by NIST.[96]

The National Archives of Canada envisions a possible applications portability approach based on a UNIX environment such as POSIX, permitting the receipt and/or conversion of electronic records to international exchange formats such as ODA/ODIF, SGML, and GKS.[97] In addition, the National Archives of Canada believes an information resource dictionary system, ("IRDS") is of particular importance because it permits an archives organization to maintain intellectual and physical control over electronic and nonelectronic holdings.[98] The National Archives has a working group on office system standards bringing together 25 government agency representatives involved in the design, installation and use of integrated office systems. A major project of the working group is to test ODA/ODIF by exchanging information among the Treasury Board and some 4 or 5 other agencies to demonstrate ODA converters, applications portability and other related questions. The working group also plans to develop a refined set of functional specifications for managing information in an office systems environment.[99]

A report prepared in early 1990 proposed functional requirements for managing information in networked agency office systems.[100]    The requirements focused on capabilities to be provided in application software:

---

[94]McDonald 1988 Archivists Paper at 1.

[95]Canadian 1989 Status Report at 2.

[96]Canadian 1989 Status Report at 2.

[97]Canadian 1989 Status Report at 2. The United States has adopted aspects of the POSIX standard. FIPS Pub. 151, 55 Fed.Reg. 11424 (March 28, 1990).

[98]Canadian 1989 Status Report at 2-3.

[99]Canadian 1989 Status Report at 4.

[100]National Archives of Canada, Managing Information in Office Automation Systems: Final report of the FOREMOST project (1990) [hereinafter "FOREMOST Report"].

-- Collecting required records and safely storing them without altering content.
-- Implementing a properly organized formal records system.
-- Maintaining compatibility with software likely to be used in the office environment.
-- Allowing retention and disposition of records as specified by the owning organization.
-- Providing full text search capability.
-- Allowing a records manager to grant or revoke access privileges for records.

The automated records management system ("ARMS") offers the capability for managing electronic and nonelectronic records, including electronic mail records and word processing records.[101] The program envisions converting every document sent to the records system to ODIF standards.[102]

The report envisions a life cycle of electronic documents beginning with creation of a document in a user's "personal work space, followed by transfer of the document to another user or to the records system, at which point a document profile would be created, assessment of the utility of blocks of files to the National Archives, and actual transfer of documents to the archives in media that can be read by the archives and in formats provided by the ODIF standard.[103]

## G. Optical Storage Initiatives

Optical storage, despite the hardware and software dependency motivating NARA's reluctance to accept optical storage formats, is an attractive technology to agencies confronted with large volumes of paper records. Optical storage offers great information density. Transferring paper records to optical media reduces storage requirements. The Internal Revenue Service, the Veterans Administration, the Social Security Administration all are actively using optical storage systems. The Department of Army also is transferring service records to optical disk.

---

[101] FOREMOST Report at 1-10.
[102] FOREMOST Report at 21.
[103] FOREMOST Report at 37.

## H. State Archives

State archivists face many of the same problems as NARA. A conference of government archivists in June, 1989[104] concluded that the common problems are (1) finding an appropriate definition of electronic records, (2) failure of information policy makers to consider archives issues, (3) inadequacy of laws and regulations covering electronic records either because of inadequate coverage or conflicting provisions, (4) lack of legislative interest in information management issues, and (5) decentralized standards for hardware and software.[105] No model electronic records program has emerged at the state level because state archivists tend to look to each other for solutions and no one has taken the lead.[106] The conference concluded that attention should be given to the following principles:

1. Archivists may use approaches for electronic records that have not worked well with paper records.
2. Archival principles like provenance[107] and original order apply to entire electronic information systems, and not to electronic records.
3. Effective access to electronic records may militate in favor of allowing the records to stay with their creators rather than being accessioned to archival repositories.[108]

In addition, the most appropriate role for archivists may change. Archivists may provide leadership in developing appropriate office information standards and in developing archival profile standards to be used in designing electronic information systems.[109] Archives may provide a public service by providing a directory of where various information systems are located and how the public can use them.[110] Archivists also may be able to help design better information systems, because of their particular perspectives.[111]

The National Historical Publications and Records Commission is funding electronic records projects in Wisconsin, Kentucky, New York and Florida,

---

[104]School of Library and Information Science, University of Pittsburgh, Archival Administration in the Electronic Information Age: An Advanced Institute for Government Archivists (June 4-16, 1989) (summary of proceedings) [hereinafter "Pitt Proceedings"].

[105]Pitt proceedings at 4.

[106]Pitt proceedings at 5.

[107]Origin, source.

[108]Pitt proceedings at 6.

[109]Pitt proceedings at 8.

[110]Pitt proceedings at 8.

[111]See Pitt proceedings at 8.

and several state archives have produced white papers concerning various aspects of managing electronic records.[112]    The Commission made certain recommendations in its 1990 report for the types of research activities that should be supported.[113]

## I.  Vision of the Long Term

In the long run, wider use of electronic information technology will change administrative agency structures and application of traditional administrative law concepts.[114] Agencies will use electronic information technologies for adjudication, rulemaking, internal management and to deliver important services.   In electronic adjudication, hearing and pretrial procedures all would be electronic.  Recordkeeping and filing obligations imposed in aid of enforcement[115] also would be electronic, as in EDGAR and the IRS Electronic Filing Project.  The rare exception would be a contested matter in which the credibility of witnesses is at issue. Even there, videotaped depositions could be taken in advance, digitized and presented to the administrative law judge through hypermedia electronic filing and database techniques.   There would be a much more complete and much more accessible record of all adjudications than with paper processes. Because of the basic similarity of the adjudication process regardless of agency, certain format and software standards would emerge that would make it easier for adjudication records to be transferred among agencies and the National Archives.   Appellate review of agency adjudications would be facilitated by development of judicial capacity to handle the standardized formats.

Using electronic information technology for rulemaking has even greater potential to change the way in which government operates because rulemaking as a quasi legislative process is concerned with policy making and is supposed to involve whatever public is concerned with a particular  policy.  Electronic

---

[112]Commission Report 4 at 7.

[113]Commission Report 4 at 8-9.

[114]The phenomenon is not limited to administrative agencies. *See generally* Cash & Konsynski, *IS redraws competitive boundaries*, Harv. Bus. Rev. Mar.-Apr. 1985, at 134 (penetration of information systems into internal business processes); B. Konsynski & F. McFarlan, Information Partnership: Scale Without Ownership (Harv. Bus. Sch. Case N1-191-023) (information technology allows enterprises to cooperate in new ways).

[115]Recordkeeping and filing are grouped with adjudication because both involve individual compliance with preexisting rules, and produce raw materials from which adjudicatory enforcement proceedings may spring.

information technology greatly facilitates timely public involvement and reduces burdens on agencies affording that involvement. The notice and comment process could be made largely electronic, with notices being posted on electronic bulletin boards, and comments being submitted via dialup telephone links. The electronic record of a particular rulemaking decision would be all of the entries in the appropriate table of a rulemaking database. The record on petition for judicial review would be a set of electronic pointers to appropriate places in the table.[116] A reviewing court simply would access the appropriate table through the pointers. The rulemaking process might become less discrete and more in the nature of a dialogue. In the long run, there would be a tendency for the interactive capability to blur the distinction between incomplete and complete rulemaking decisions. The technology permits a dialogue between regulator and regulatee, reducing the need for communications between the two to occur by means of formal written documents in the form of petitions, comments, and final rules.

Internal management practices would build on concepts developed for the Forest Service. Management of agency resources[117] in the electronic vision would work primarily by relaxing the need for official paper documents to record and communicate instructions and decisions to lower levels in the organization. Rather than signing a new delegation of authority, a memorandum making a change in organization structure, or a directive reallocating enforcement resources, an agency head would post a notice directly from her desktop workstation electronically changing a database record. The state of the organization, of delegations of authority and of resource allocations would be defined officially by the state of the database.

Electronic information technology also can be used to deliver some governmental services: dissemination of public information, and electronic transfers of money in connection with public welfare and subsidy programs. Electronic food stamps, direct deposit of social security checks, and electronic management of Medicare and Medicaid benefits all are examples. Such electronic information technology applications use well-proven electronic funds transfer techniques, combining them with EDI standards in the case of Medicare reimbursement. The advantages are faster availability of funds, and

---

[116]The component of the database associated with a particular rule could be thought of as a record. As one considers adding more and more to the file for a particular rule, however, it may be more appropriate to think of the file for a particular rule as a table containing many records, in which case the regulatory agenda for a particular agency would be a database composed of many tables.

[117]One could view management of agency resources as involving weak forms of rulemaking and adjudication, not involving the same degree of legal formality aimed at ensuring political accountability, but involving the same types of decisionmaking processes.

improved ability to audit.   The problems to be solved are those of electronic contracting and electronic funds transfer generally: potentially increased risk of forgery.[118]

# IV.  Policy and Legal Issues

## A.  Goals

It is appropriate to define the principal purposes of records management and archives.  The following purposes seem relatively noncontroversial:
-- to maintain an institutional policy memory
-- to enhance the body of knowledge in a factual and scientific sense
-- to maintain official records for possible use as evidence in legal proceedings[119]
-- to maintain the nation's history
-- (for records management) to promote efficiency and effectiveness of agency operations
-- (for archives) to increase the return on investment from information in government records which has long term value.

The last goal for archives applications would be enhanced by expanding access to archived information at reasonable costs.

Other conceivable purposes would engender more controversy, such as a goal of providing expanded electronic public access to public information on demand.

## B.  Some Basic Concepts

## 1.  Classes of Agency Information

Four different types of agency information present somewhat different records management and archives problems: policy documents, scientific and

---

[118]Forgery, broadly conceived, involves misrepresenting the author of a message, and also involves unauthorized alteration in the contents of a message.

[119]The NARA guidelines for documents to be preserved primarily reflect this goal.

technical information like satellite-collected data, agency administrative records, and social data like census data.

Electronic policy documents only recently have presented records management problems. Because such policy documents must get distributed on paper, and because of uncertainty about electronic signatures, such policy documents virtually always are printed on paper before they are signed by the highest level decisionmaker. The paper medium is the archival record. But this practice will change, as decisionmakers increasingly have access to documents in electronic form, and as electronic transactions become a regular feature of life. Soon some agency officials may make choices by electronically checking a box on an electronic option paper sent by electronic mail. The Forest Service System[120] is another good example of changing practices regarding decisional documents. In the Forest Service System, directives from the agency head to the field offices do not necessarily exist in paper form at any time. As electronic mail becomes more pervasive, and as electronic publishing of government information becomes more common, the paper version of many official documents may disappear. As electronic mail is used more widely by a greater variety of agency decision makers below the top level, the part of public decisionmaking covered by paper documents shrinks. The PROFS System at issue in the *Armstrong* case[121] is a good example.

There also is a need for better records management plans to capture prefinal communications to preserve the content of the decision process. Such predecisional electronic transactions likely are protected by FOIA exemption 5,[122] but they are an important part of the historical record.

Scientific and technical data present very different problems. One problem is the large amount of data collected, particularly the quantities of data collected by satellite. New technology such as weather satellites enormously increase the amount of information that is collected and used by agencies like NOAA. The other problem relates to the interpretation and analysis of such information. An unskilled person cannot use satellite weather information in its raw forms even if presented on paper media. Even a meteorologist cannot use the raw information without software and careful documentation compatible with the data.

To some extent, scientific and technical data share a general problem with electronic records: a need for compatible hardware and software to retrieve the information. In another respect, however, the interpretation and analysis problem with scientific and technical data is different. Scientific and technical data is specialized and its consumers always will be specialists in particular

---

[120]*See* §III(D).

[121]*See* §II(D)(3).

[122]5 U.S.C. §552(b)(5) (1982).

fields such as meteorology and climatology. The specialized nature of this information justifies handling archives of such information separately from the general archives. The issue has been avoided in a direct way because NOAA, NASA, and DOD have not yet offered files containing large quantities of satellite data to NARA in an official way. There is, however, concern about NARA's ability to handle large quantities of satellite data with current resources.

Administrative information includes purely internal information like information pertaining to agency personnel and contracting. It also includes mission-related information such as claims files for the Social Security Administration ("SSA") and the Veterans Administration ("VA"), and tax records of the Internal Revenue Service ("IRS.")[123] Agencies have their own incentives to maintain administrative information for extended periods of time, independent of obligations imposed by the records statutes. These mission-related incentives relate to a potential for litigation over contested claims, or contested personnel actions. Agencies aggressively are automating the processing of such information simply to meet increasing burdens of claims. Agencies also have incentives to apply new archival technologies to increase storage density for large volumes of information. The SSA, VA and IRS have been aggressive in exploring optical storage technologies.

Social information, like census information, has received the greatest attention in terms of electronic records management. Archived census information was the first type of archival information made available to the public, and the private sector invests much effort to ensure the utility of electronic records formats for academic researchers and others. There is, for example, a specialized trade association for users of census information.

A basic crosscutting distinction must be made, under present technology, creation and use patterns, between database information and word processing documents. Word processing documents are fixed in form, though they may exist in different versions. Databases are inherently dynamic. A document is more likely to have decisional significance than a snapshot of a database, "taken" at a time when no one happened to be retrieving information from the database. Moreover, the database itself is not what people consume or ultimately use. A database is analogous to a dictionary, out of which users extract pieces of information to assemble into documents with significant information value.

Is it more important to preserve the entire database, or the retrieval transactions? A convergence of text document and database paradigms is considered in §IV(F)(2).

---

[123]NARA distinguishes "administrative" or "housekeeping" information from mission related or program information.

## 2. Value in Information

An important part of what makes information useful is value that has been added to the information. Nine distinct types of value[124] can be added to information, to make up an information product bundle. The nine types represent a kind of Chinese menu from which information product suppliers pick and choose, combining different amounts and subtypes[125] of the different types into products with features the supplier believes will attract significant demand.[126] The nine types are presented in the order in which they would be added to textual information in print formats.[127] The basic model is adaptable to nontextual information, and is not dependent on any particular sequence of value.

1. Selecting and arranging content. Drafting language for a textual document and the selection of entities and attributes for a database are examples. This type can be called <u>authoring</u>.

2. Adding typographic features such as paragraph, section, and page breaks, headings, tables and summaries, writing an outline or designing a database schema.[128] This type can be called <u>chunking</u>.

---

[124]Both economics and law deal with added value. Added value is the product of economic activity. Law usually protects added value as property. "Value" to an economist is a holistic quality, reflected in price. This report uses the term "type of value" to refer to particular dimensions of value, or product attributes. The types of value can be thought of as the addition to value resulting from a particular process in the production activity. Alternatively, they can be thought of as product features that provide utility to consumers. There are parallels between the type-of-value idea and the hedonic dimension idea.

[125]A variety of subprocesses exists, choices among which can have important economic consequences. For example, types one and two embody drafting, editing, and revising subprocesses that have significantly different costs depending on how the author, secretary, and editor use paper, word processing, and dictating technologies. See H. Perritt, How to Practice Law with Computers, Chapter 8 (1988 and 1990 Supplement) (comparing three different cases employing different technologies).

[126]The nine types work best with text information that is factual in character, as opposed to artistic or functional. See U.S. Congress, Office of Technology Assessment, *Intellectual Property Rights in an Age of Electronics and Information* 65-66 (1986) (three content categories).

[127]The nine types encompass the information processing activities likely to be associated with publishing or exchanging electronic information. Other types could be described, which relate to higher level analysis of information, involving the kinds of functions performed by expert systems. See H. Perritt, How to Practice Law With Computers Chap. 9 (1988 & Supp. 1990). The nine types circle back on themselves. The consumer of one electronic information product may be the author of another information product, representing a human transition between the ninth type of one information product to the first type of another product.

[128]Most, but not all of these features are text <u>compressors</u>--short expressions that summarize larger amounts of text to facilitate browsing. See R. Taylor Value Added Processes in Information Systems 59-60 (1986).

3.   Adding internal finding aids like headers and footers, cross references, tables of contents, indexes, hypermedia pointers or graphical representations of content.[129]   This type can be called internal tags and pointers.

4.   Adding external finding aids like substantive references and citations in footnotes or reference lists, generating multiple-document indexes, or adding hypermedia pointers.[130] This type can be called external tags and pointers.

5.   Printing or displaying a single image.[131] This type can be called presentation.

6.   Making multiple copies of the image.   This type can be called duplication.

7.   Distributing the copies to consumers of the information. This type can be called distribution.

8.   Advertising and promoting the resulting electronic information product, and accounting, billing and collecting prices charged for the use of information. This type can be called marketing.

9.   Assuring users and consumers of the integrity of the information product.   This type includes the kind of quality control function traditionally performed by publishers,[132] relating to whether the content of a publication is likely to interest a particular consumer, whether it is likely to be accurate, and whether substantial value has been added at types two (chunking), three (internal tags and pointers) and four (external tags and pointers), so the consumer can expect high utility. In electronic information formats, type nine value also includes ensuring against corruption and ensuring authenticity. This type can be called integrity assurance.

---

[129]Headnotes in reported judicial opinions are a kind of type 3 value.

[130]Type four is closely related to abstracting and indexing. See R. Taylor at 7 (describing value added by abstracting and indexing operation). Abstracting and indexing creates the hooks and tags by which chunks of information can be linked, through type four value to other chunks of information.

[131]The image may or may not be the same size as the basic "chunk" of information. A database record retrieved may fit on one screen, and a record is the basic chunk in database technology.   Conversely, PC word processing software displays about half a page of single-spaced text, and that is smaller than the basic chunk of a textual document--which is the document itself, or perhaps a page, a section, or a paragraph of the document. See R. Taylor at 11 (discussing chunk concept). Attorney David Johnson argues strongly that the paragraph is the natural chunk.

[132]Readers of magazines and newspapers and purchasers of books and readers of journal articles select material based in part on expectations about the quality of information likely to be published by certain well-known publishers.

Type one (authorship) involves the raw ideas fixed in some tangible medium.[133] A rough draft, or a tape containing dictation, has type one value but little else. Paragraph breaks, outlines, pagination, subtitles, headings, represent type two (chunking) value. The difference between type two and type three (internal tags and pointers), both of which deal with internal structure and retrieval aids, is that type two is sequential or linear, while type three is random access or nonlinear.[134] Random access capability may improve retrieval efficiency. The difference between type three (internal tags and pointers) and type four (external tags and pointers), both of which involve random access, is that type three involves intra document features while type four involves inter document features. Type four includes documentary cross-references or bibliographies pointing to other parts of relevant literature, as in legal footnotes citing cases and statutes external to the citing document.

The nine types of added value can be clustered for easier understanding and exposition. Types two (chunking), three (internal tags and pointers), and four (external tags and pointers) all have to do with representation, organization[135] and retrieval of information.[136] These can be thought of as representational and retrieval factors.[137] Type five (presentation), six (duplication), and seven (distribution) types of value all have to do with delivering information, and are in some sense presentational.[138]

---

[133]To borrow a concept from copyright law.

[134]Headlines, other types of prominent topical headings and other text compressors have characteristics of type two and type three (internal tags and pointers). They fit into type two because one must scan the material sequentially to find them. They are random access or nonlinear and thus fit into type three because one can read only the headlines without having to read the text between them.

[135]Types one, two, three and four embrace the "organizing" activity suggested by Kenneth Boulding as one of two knowledge-producing economic activities. The other activity suggested by Boulding is printing. See Boulding, *The Economics of Knowledge and the Knowledge of Economics*, LVI Amer. Econ. Rev. 1, 5 (1966; papers and proceedings of 78th ann. meeting; Richard T. Ely Lecture) (capital is knowledge imposed on material world by an organizing process, followed by a process akin to three-dimensional printing).

[136]Attorney David R. Johnson points out that one can distinguish value added to information depending on whether it changes content or whether it affects the way in which and the ease with which a reader can find the information.

[137]*See generally*, H. Perritt, How to Practice Law With Computers, ch. 9 (1988 & Supp. 1990) (explaining pervasiveness of knowledge representation issue in automated legal reasoning as an application of Artificial Intelligence techniques).

[138]Presentational factors are not entirely independent of representational factors, however. See Text, ConText, and Hypertext 36 (ed. E. Barrett 1988) (visual production of text on computer screen is "crippled" without accompanying typographic features); id. at 297-299 (screen design for online information products).

Type eight (marketing) and type nine (integrity assurance) value have to do with meta-information: information about information.[139]

This same nine-type value-added concept framework is as useful for evaluating the electronic archives issues as it is for evaluating electronic dissemination issues.[140]  Agencies are likely to strip value from electronic records submitted to the archives for different reasons than they may strip value from information released to the public.[141]  In both cases, however, it is appropriate to understand the nature of the value lost.

The National Archives can provide for all nine types of added value, or none.  For example, the information transferred to the National Archives could include value added at stages one through four,[142] and the National Archives itself could add value at stages five through nine as desired by National Archives users.[143]

Alternatively, agencies could provide only the raw content (stage one value) of the information to the National Archives.   Then, the National Archives system could provide for adding value at stage five, but nothing else. Stage five value (image presentation) is necessary for a person to use information and therefore stage five value is an essential part of any information system including a National Archives system.

Standards are a necessary prerequisite to transferring value added at stages one through four.[144]  Standards also are necessary if an archival system is to

---

[139]Professor Stigler observed that the structure of markets, the role of professions and of other aspects of economic organization frequently are explained best by the need to reduce the cost of information. See Stigler, *The Economics of Information*, 66 J. Pol. Econ. 213 (1961). Stigler is talking about information about noninformational products.   When the product is information, the kind of information considered by Stigler can be labeled meta-information.

[140]*See* Perritt, *Federal Electronic Information Policy*, 63 Temple L. Rev. 201 (1990).

[141]They would strip value from information transferred to the National Archives in order to transfer a standard format, like ASCII.  They might strip information offered to the public in order to avoid competing with private vendors of value-added products, as discussed in the report supporting ACUS Recommendation 88-10.

[142]Pagination, typefaces, index, tables of contents, external references.

[143]For example, presentation on screen or paper, making copies, distributing the presentation images to users.

[144]For example, the added value represented by an index cannot be meaningfully transferred unless the transfer protocol preserves pagination or other references in the index.   This an example of the connection between standards and the transfer of stage three value.   Footnotes cannot be transferred meaningfully unless there is some kind of standard to link the content of the footnote to the footnote reference.  This is an example of the connection between standards and the transfer of stage four value.

add value at stages six and seven.[145]  The stage one to four standards relate to the interface between an agency and the Archives. Standards applicable to stages six and seven relate to the interface between the archives and the user.

## C.  Technical Issues with Legal Implications Raised by Electronic Media

A 1990 report of the National Historical Publications and Records Commission[146] identified the following problems and issues for electronic records:

-- System dependencies
-- Storage media
-- Easy changes and easy loss of information
-- Difficulty in defining a record
-- Difficulty in distinguishing originals from copies[147]

The Commission concluded that the following solution alternatives are worthy of attention:

-- System design to provide for electronic records management and archival needs from the beginning
-- Development of appropriate standards to reduce fragmented system dependency
-- Interdisciplinary projects.[148]

The following sections in this subpart elaborate on the issues raised by the new technologies, generally agreeing with the Commission on the nature of the problems posed.

## 1.  Records disposal

A major problem with paper records is storing them.  If everything is saved, storage requirements become too great. Moreover, effective access for legal, policy development, or historical reasons becomes much more difficult

---

[145]Standards translating attributes in a text file to video display attributes are necessary to display fonts and attributes such as bold facing and underlining.  Standards also are necessary for any kind of data communication, and thus are necessary for adding stage seven value.

[146]National Historical Publications and Records Commission, National Archives and Records Administration, Electronic Records Issues: A Report to the Commission (No. 4 March, 1990) [hereinafter "Commission Report No. 4"].

[147]Commission Report 4 at 3-5.

[148]Commission Report 4 at 6.

with larger inventories of paper records, because finding the desired record requires one to review many more records. Thus, an important objective of a sound records management policy governing paper records is to ensure that a substantial portion of records without historical value are disposed of.

Electronic technologies mitigate this need to some extent. The space required to store information in electronic form is several orders of magnitude less than the space required to store paper information. Moreover, electronic retrieval techniques make it less burdensome to search for a particular record through many undesired records.

Nevertheless, there is a practical need for considering records disposal as a part of system design. Electronic storage technologies have finite capacities. Frequently, as occurred with the Forest Service system, insufficient attention is paid to getting information out of a system. When information is not removed from electronic systems, storage requirements soon exceed design assumptions.

## 2. Records retention

Records retention requires deciding what records should be retained and then actually keeping those records when other records are destroyed. These two acts, deciding and keeping, can be done at the same time or they can be separated. In an extremely simple hypothetical agency, a decisionmaker could review records, making judgments on the spot as to what records are of historical value and setting those records physically aside for retention while throwing other records in the waste basket.

In the real world, however, the decision about what to retain is made in the abstract, through the articulation of criteria for retention. Then the act of keeping records meeting the criteria is made by someone else (perhaps a computer) at a later point in time.

Electronic technologies raise two generic problems with records retention. First, they make it much easier to destroy records. Second, they also make it easier to apply predeveloped criteria for records retention.

There is no question that text and document management systems are appropriate adjuncts to a sound electronic records management plan. Otherwise dispersal of electronic documents on decentralized computer systems, and cryptic file names make capture of appropriate official records even more difficult in an electronic regime than in a paper regime. Moreover, electronic document management approaches may make it easier to discover and capture official records using electronic techniques than is possible with paper techniques.

Agencies usually assign a relatively low priority to record management activities. It is not likely to be feasible, therefore, to impose additional

burdens on agency employees to code working documents for eventual archival purposes if the coding requirements impose any significant cost, in terms of time or inconvenience. Certainly a post hoc requirement to reformat word processing documents with archival value would impose significant costs. Therefore, the best approach for textual documents either is to design archival coding into the document management system at the outset, or to have an overinclusive schedule that causes most textual documents to be transferred to the Archives on increasingly inexpensive storage media, assuming that free text searching techniques can only improve and, even with the present state of the art, would permit appropriate retrieval precision when someone wants to access the documents.

The ultimate goal should be making records management transparent to agency employees. For example, official correspondence could necessitate a flag that must be sent before the system would recognize it as official correspondence. The flag would cause the document to be saved and processed for eventual archiving.

Designing agency information systems to provide adequate records management capability, transparent or otherwise, implicates a tension between centralized and decentralized automation strategies that is pervasive in office automation.[149] Centralizing information system administration makes it easier to ensure that records management policies are followed. But centralization deprives individual users of the autonomy they have come to expect as a result of the PC revolution. Autonomy in using computing resources is not only an expectation; it is a technological capability that enhances individual productivity.[150]

There are two points to be made about this tension. First, the tension should be recognized. It is unavoidable. It also is not new; the conflicting desires for centralized control and decentralized entrepreneurial energy is a central issue of organization design.[151] Second, a really good electronic records management program permits substantial amounts of user autonomy with respect to software selection and use patterns. A concrete example may be helpful in illustrating this point. Suppose an agency adopts ODA/ODIF as the standard for its textual documents. It might consider three alternative

---

[149]*See* H. Perritt, How to Practice Law with Computers Chapter 8 (1988 & 1990 Supplement).

[150]Individual users having reasonably autonomous control over their own computing resources can develop macros, select their own software, and otherwise be innovative in matters that help them get their own work done better.

[151]*See* A. Chandler, The Visible Hand (1977) (describing centralized and decentralized approaches followed as railroads pioneered large scale organization design, and General Motors and Dupont followed different strategies).

strategies for ensuring that agency personnel comply with the standard. One strategy would be to develop new word processing, document management, and desktop publishing systems that comply with the standard. A second alternative would be to procure commercially available products that comply with the standard and to permit personnel to use no other products except a single selected product, say Microsoft Word. A third alternative would be to identify a preferred product, say Microsoft Word, and allow users to use any other product as long as files can be transferred from the nonpreferred products to the records management system, say in ASCII form or some other interchange medium.[152] Of these alternatives, the first is the worst, and the third is the best. The public interest is not served by sacrificing agency effectiveness and productivity in the interest of long term archives enhancement. Electronic records management practices must serve both goals.

Considering management of electronic records forces planners to think about a choice that regularly confronts designers of document management systems for office automation systems. Should one invest effort at the beginning in formats and indexing for an eventual retrieval, or should one simply save the information in whatever format is best for its original use and burden the individual archive user with search and retrieval effort?[153] A clear example of this choice is between formatted textual databases and free text search.[154]

Historically the National Archives took the latter approach. Information was, for the most part, simply forwarded to the National Archives in whatever formats existed, with only limited effort expended towards indexing. Computer database techniques, however, generally focus on front end investments in tagging aimed at reducing the cost of eventual retrieval. To the extent that front end formatting and tagging imposes avoidable human costs, the protocol for records retention is less likely actually to be followed by document generators and receivers. On the other hand, some features aimed at records management can be built into database systems at a relatively low cost. Database approaches frequently are preferred by agencies, not for any reasons having to do with records management, but simply to facilitate accomplishment of the agency mission.

Document management systems that associate profiles with specific textual documents apply some of the database design approach to free text documents. The reworked Forest Service system, providing virtually immediate dissemination of agency head documents to some 300 field offices has a

---

[152]Microsoft's RTF and IBM's DCA/RFT might be examples.

[153]This is another way of raising the publish-on-demand possibility.

[154]*See* H. Perritt, How to Practice Law with Computers 442-444 (1988).

document field that automatically causes appropriate documents to be designated as official records and referred electronically to the archives.

The philosophy expressed by the NARA guidelines envisions that individual microcomputer users follow records management plans in creating files within word processing, database, and spreadsheet applications and that such individual users refrain from erasing files before ensuring that an appropriate official copy exists. Ensuring compliance with such guidelines is extremely challenging. Decentralization of computer file management associated with the microcomputer revolution obviously increases the challenge.

On the other hand, ensuring compliance with this electronic records management philosophy is no more challenging than ensuring compliance with paper records management philosophies, given the reality that important subsets of agency records exist in filing systems (which may not be all that systematic) under individual control, and not only in centralized agency file rooms.

Indeed, it may be easier to capture appropriate records when the records exist in electronic form than when they exist in paper form. A paper document existing only in the personal files of the author and the addressee may never be known to an agency's official records management system. An electronic equivalent of the same document is likely to exist on a computer network and can be discovered by an appropriately designed electronic protocol for electronic records management.

Electronic technology permits draft management, because a well designed office automation system can track who looked at and commented on a document and when.

Automating capture of drafts and of decisionmaking records is much easier on networked computer systems, as compared with free standing microcomputer systems. On networked systems, the records management functions can be designed into the host or the file server, which keeps most of the files for an entire system of users. On free standing microcomputer systems, in contrast, the only way to implement automated records tracking functions is to rely on the user of the particular microcomputer periodically to transfer documents via floppy disk or other transferable medium or to rely on a management function that periodically copies documents from the hard disk of each microcomputer. Indeed, Local Area Networks ("LANS") force a degree of electronic records management on users and LAN administrators that is not absolutely necessary in standalone computing environments. If users are to be able to find files on a LAN, some attention in designing and running the LAN must be given to electronic records management.[155]

---

[155]McDonald 1988 Archivist Paper at 5.

Well-accepted technologies do not permit capturing marginalia, however, although the Wang Freestyle technology, if it is accepted in the marketplace (or is a part of a government standard), could permit systems to capture voice comments and other marginal input. Alternatively, of course, wide acceptance of an existing comment management system would permit such information to be retained.

Despite these technological capabilities, the goal of capturing draft documents in order to record the process of policy making is, to a considerable extent, unrealistic.[156] Draft documents do not necessarily reflect the policy deliberations, and forcing the retention of drafts may be impracticable.

The author has had considerable experience in negotiating labor agreements and participating in policy formulation at the cabinet and subcabinet level of the United States Government. Both of these processes produce many draft documents. Many of the most important ingredients in the policy development process, however, are oral, and are never reflected in discrete changes between two identifiable draft documents. Moreover, many drafts are never seriously considered, because they were written at lower levels not informed by the most current discussions among decisionmakers, because they were drafted by interested parties and submitted to the decisionmakers, or because they reflected efforts to make more concrete the options at the margin of what policy makers were willing to consider. Not only do such draft documents not capture the main parts of policy deliberations, preserving them may affirmatively mislead as to what was seriously considered.

Requiring that personal computer users keep all drafts of their documents on a central network server is somewhat like requiring office workers to keep no papers in their desks, but only in centralized filing cabinets. The likelihood of compliance is low. Of course, one could design personal computers so that they do not have disk drives, just as one could design desks so they do not have drawers. Such an approach, however, sacrifices worker convenience and productivity in the interests of records management, and this may not be a sensible organizational strategy.

The policy formulation process is as much oral as it is written. The new technologies do not change that reality, unless electronic mail induces

---

[156]In reviewing an earlier draft of this report, NARA staff disagreed with the conclusion that goal of capturing draft documents in order to document the process of policy making is unrealistic. NARA believes that the public has a right to know about the evolution of policy and that it is realistic to demand that agencies make an effort to document the evolution of policies. The records management statutes require agencies to maintain adequate documentation of policy and programs. Drafts may reflect how policy and programs are developed, and therefore be part of adequate documentation. Reportedly, a draft of policy papers for the Japanese-American Internment Program in 1942 revealed important new information about the purpose and the factual basis for the Internment Program.

decisionmakers to deliberate less face to face and over the telephone and more via electronic mail.

Significant investment or distortion of agency practices is not warranted in order to capture draft documents. Moreover, as one archivist put it, "People will doodle, and they will figure out a way to frustrate any totalitarian records management system that won't let them doodle." Other archivists, however, stress the importance to the historical record of draft documents. An appropriate conceptual approach to the draft question may be to distinguish between personal drafts and institutional drafts--those that are circulated within an organization for clearance or revision. Only institutional drafts should be covered by retention schedules.

The NARA guidelines may be too limited in requiring structured and field-based approaches to automatic document retention. It is a mistake to limit thinking about automation retention selection to structured headers. Effort should be invested to develop free text algorithms (probably on an agency-by-agency basis) for selecting documents as candidates for retention.[157] Free text searching technologies available now would permit development of reliable algorithms for individual agencies, based on key words likely to be contained and archival records. Thus, the full text of working documents could be searched at appropriate intervals, with the algorithm deciding what records should be archived. Resources permitting, NARA's Center for Electronic Records will begin working on free text retrieval approaches in FY91.

## 3. Records integrity and accessibility

Records retention in archives does no good unless the archived records can be accessed. Three technical issues with legal implications relate to accessibility. All have to do with integrity, in some sense. Meaningful access to electronic records requires that formats be processable and readable. Paper records that physically survive remain readable as long as one understands the language that they are written in. Electronic records present larger hardware, software and format compatibility problems. Some electronic records deteriorate more rapidly than some paper records.[158] Large quantities of records impede practical retrieval unless records are indexed.

---

[157]There is room for debate about whether the development of such algorithms is practicable. The state of the art in natural language processing is still rather primitive.

[158]FAX messages on special paper probably deteriorate more rapidly than some magnetic media. Some electronic media such as materials for optical recording may deteriorate less rapidly than paper.

## a)  Easy alteration

Magnetic media can be altered more easily and there is less likelihood of detecting the alteration than in paper media.  If one changes something on a piece of paper, the eraser or the white out frequently is evident--at least on the original.  Because the physical changes involved in changing the content of magnetic storage are not visible, the possibility of undetected alteration of magnetic records is appropriately of concern.  It should be noted, however, that information stored on optical media cannot be altered as easily--at least on CDROM media.  Especially for information that is used for evidentiary purposes in legal proceedings, procedural or technology techniques must be used to reduce the risk of undetected alteration.

## b)  Deterioration of media

Environmental control[159] is important to protect paper and bindings.[160] All post-1840 paper is subject to acid deterioration which cannot be halted, except for freezing or helium atmosphere storage. In 1986, the Law Library Journal included an updated report on the life-span of paper.[161]  It concluded that 20th century paper possesses a life-span of 30-80 years, depending on the production quality of the book and of the individual paper (in the case of photocopies or unbound works).[162] Another report indicates that the maximum longevity of 20th century paper is 50 years.[163] The House of Representatives has inquired into the problems presented by deteriorating paper information.[164]

Microfilm. Microfilm exceeds all other recording media in terms of longevity. Estimates range upwards to 300 years for silver microfilm and to 100 years for diazo microfilm. Silver film, however, is more easily destroyed by fungus and

---

[159]Especially temperature and humidity.

[160] 73 Law Library Journal 835-36 (1980).

[161]  78 Law Library Journal 244, 258.

[162]  Id.

[163] 50 College & Research Libraries, No. 5, 577 (Sep.1989).

[164]See House Committee on Government Operations, Establishing a National Policy on Permanent Papers, H.Rep. 101-680 Part 1, 101st Cong. 2d Sess. (report accompanying H.J.Res. 226) (urgently recommending use of acid free permanent papers by federal agencies and private sector; reporting average cost of $50 per book for Library of Congress effort to microfilm books that have deteriorated); To establish a national policy on permanent papers, Hearing before the Government Information, Justice, and Agriculture Subcommittee of the Committee on Government Operations, House of Representatives, 101st Cong. 2d Sess. on H.J.Res. 226 (February 21, 1990) [hereinafter "permanent papers hearings"].

scratches than its diazo counterpart.[165] Microfilm requires particular storage conditions, and appropriate processing and packaging in order to avoid quick deterioration.[166]

Magnetic Tape. Conservative estimates concerning the archival lifespan of magnetic tape range from five to ten years.[167] There are other estimates asserting that the tape may last up to twenty years.[168] Reliability of estimates are questionable, however, because of the wide use of the Arrahenius test,[169] which incorrectly presumes that accelerated testing produces the same or similar results as actual long term exposure.[170]

Like any recording medium, the better the quality of the product, the longer it will last.[171] Additionally, appropriate storage conditions extend the lifespan of magnetic tape. The key to successful storage is the retensioning of the tape at least every two years.[172] This presents a burden for constant retensioning on a records custodian like the National Archives, which may possess millions of tapes.

Magnetic Cartridges. Magnetic cartridges are better for storage than magnetic tape. An accelerated test performed by IBM indicated that the 3480 class tape cartridge ("3480") has a life expectancy of over thirty years.[173]

From an archives institutional perspective, cartridges are better than magnetic tapes because no attention need be given to retensioning of

---

[165] Fungus can destroy film in a matter of days.

[166] 73 Law Library Journal 835-36 (requirements: 70 degrees Fahrenheit and 40% humidity).

[167] Telephone interview between Richard J. Ciamacca, research assistant to author, and Dana Grubb, representative from National Institute of Standards and Technology (Oct. 19, 1989). There is no difference between 7 and 9 track tape with respect to lifespan and storage.

[168] Id.

[169] Id.

[170] Id. The test attempts to create ordinary storage conditions but applies them at an accelerated rate so as to illustrate the long-term effects on storage media. This test has not been shown to produce accurate results. Id. The problems of this test are equally applicable to all recording media which is tested using the Arrahenius system.

[171] See 3M, 3M Black Watch: Blackcoated Computer Tape (1989) (corporation's advertisement pamphlet; Blackwatch computer tape evidences little wear after 2,000 passes while conventional tape is scratched after 500).

[172] Telephone interview with Dana Grubb, representative from National Institute of Standards and Technology (Oct. 19, 1989). Retensioning is simply the rewinding of the tape at approximately 13 ounces. See 3M, 3M Black Watch: Blackcoated Computer Tape (1989) (corporation's advertisement pamphlet).

[173] T. Weir, 3480 Class Tape Cartridge Drives and Archival Data Storage: Technology Assessment Report 6 (National Archives Technical Information Paper No. 1, June, 1988).

cartridges.[174] Also, the casing of the cartridge prevents many of the dangers associated with handling of tapes.[175] For example, there is less danger of dust or fingerprints damaging the tape because it is sealed in the casing, unlike the magnetic tape on exposed reels. While cartridges require ambient storage conditions similar to those of magnetic tape, they can withstand more severe conditions than magnetic tape.[176]

In order to ensure reliability, vendors suggest that cartridges be exercised annually or semiannually.[177] Archivists must consider which type of tape to use in cartridges. Chromium oxide tape is less stable than iron oxide tape, particularly in the long run.[178] "If archivists use a chromium dioxide medium, they must monitor it rigorously for deterioration and insure proper storage and handling."[179]

Diskettes. Neither the American National Standards Institute (ANSI) nor the National Institute of Standards and Technology (NIST) have compiled data on the archival life expectancy of magnetic diskettes, though ANSI currently is conducting tests on diskettes to determine their longevity. Regardless of the type of diskette,[180] the archival life expectancy of a magnetic disk allegedly is forever.[181] Diskettes possess an estimated 1,600 hours of usable life; that equates to over 32 million revolutions.[182] Under the proper storage conditions, a diskette may never lose its stored information due to nonuse. These conditions should be approximately 75 degrees Fahrenheit and 40% relative humidity.[183]

The main problem with diskettes is their susceptibility to damage caused by improper storage or handling.[184]

---

[174] *Id.* at 13 (3480 drive compensates if there is tension problem).

[175] *Id.* at 18.

[176] *See id.* at 8.

[177] *Id.* at 9.

[178] *Id.* at 9. The short-term reliability of chromium oxide cartridges is better than that of iron oxide cartridges. The problem is that there are little data on the long-term chemical stability of chromium oxide. Additionally, there is fear that chromium oxide is more likely to adversely react with the binder, thereby destroying the tape. *Id.*

[179] *Id.* (report indicates that risk of long-term problems with chromium oxide is probably small).

[180] Types: 5.25" high and regular density, and 3.5" high and regular density.

[181] Telephone interview with Susan Michaud, representative for Kodak, Inc. (Nov. 6, 1989).

[182] *Id.*

[183] *Id.* (disks can survive in temperatures ranging from 50-125 degrees).

[184] *See* Diskette Technical Service, 3M Diskette Reference Manual 62-5 (July 15, 1988).

CDROM. There are no confirmed data on the lifespan of CDROMs.[185] The CDROM technology is so recent that there has not been ample opportunity to conduct tests on its reliability and storage life expectancy.[186] NIST agrees with this proposition but notes that there is an industry consensus that information stored on a CDROM is secure for at least 3-5 years, with some estimates reaching as high as 20 years.[187] The relatively short estimates for CDROM at least superficially are inconsistent with the virtually unlimited estimates for magnetic disks and WORM.

WORM. Like CDROMs, there is little available data on WORMs[188] because testing is currently being performed. Kodak, however, offers a 30-year warranty on Kodak WORM products, though it claims that information stored on the WORM can last forever.[189] Like all other computer readable media, the longevity of a WORM depends on the product's quality, on careful storage and handling, and on the amount of use (more use leads to shorter life expectancy).

## c) Storage Formats

Maintaining hardware and software necessary to access electronic records, which over time will embody a diverse universe of storage technologies, is impracticable. Storage formats can be handled by periodic copying of files to new media and formats. This deals with the problem of readability of the records. A separate issue conceptually is preservation of the capability of accessing and processing electronic records as technology changes. One may be able to read the bits constituting a record created by a Radio Shack TRS 80 Model I computer in 1978. But that does not mean one will have available an operating system and the primitive word processing program necessary to make sense of the bits. Similarly, it would be difficult to access a document saved in 1984 by the Perfect Writer word processing software running on a CPM microcomputer and saved on a low density 5.25" floppy diskette.

The NARA guidelines may be too limited in requiring structured and field-based document management. It is appropriate also for electronic records management plans to include free text searching, because that may be the best way to find a document that has not been indexed according to a structured

---

[185]Compact Disk Read Only Memory. Bowker, 1 Legal Publishing Preview No. 6, 77-79 (Sep./Oct. 1989).

[186]*Id.*

[187]Telephone interview between Richard J. Ciamacca, research assistant to author, and Dana Grubb, representative from NIST (Oct. 19, 1989).

[188]WORM is an acronym for Write Once Read Many [times].

[189]Telephone interview with a representative from Kodak (Oct. 19, 1989).

approach, or in which the key words assigned during the indexing process subsequently proved to be inadequate.

Generally, there is concern that any storage medium may become outdated by the quick advances of technology in computer hardware. Already, the federal government is having problems finding computers to read tapes from the 1960s.[190] The Census Bureau managed to transfer data from the 1960 census to newer media and formats,[191] although raw data from the 1960 census were lost before anyone articulated a requirement to preserve data in machine-readable form. The most precise data remaining was the lowest level aggregation. Unfortunately, Census reported in May 1990 that files other than public use files holding data prior to 1989 are practically inaccessible because of system dependencies and ad hoc data compression.

The potential for adaptability should be taken into account when selecting a computer readable medium for long-term storage of information. The Internal Revenue Service requires that corporate taxpayers who submit electronically must maintain not only the electronic records supporting the filing, but also a complete copy of the system that can retrieve the information and the operating system.

Mass electronic storage is not the problem; the problem is ensuring access to electronic information over periods exceeding thirty to forty years. Several years ago, a NARA advisory committee said the only way to assure permanent accessibility was to record information in human readable form. It expressed confidence that conversion between human readable form and electronic forms would be nearly transparent. NARA rejected that conclusion, believing that technologies would develop to the point that accessibility of information would not depend on the availability of the same hardware and software that created it. The current solution, and an acceptable one for the future, to impermanence of electronic information in various media is to recopy or "refresh" electronic information at periodic intervals. One simply recopies magnetic tape or files on magnetic disks. Recopying is a solution not only to impermanence but also to technological obsolescence. When a format standard is becoming obsolete, information can be recopied to a new format standard. Of course this recopying requires resources and effective implementation of a program that ensures the copying and appropriate intervals.

The current staff thinking at NARA is that appropriate standards are the best way to assure infinite accessibility notwithstanding changes and hardware

---

[190]New York Times (National), Oct. 10, 1989, at A20, col. 1.

[191]In 1978, the Census Bureau reported to the National Archives that it had successfully converted files in a Univac II-A tape format to a current format, losing only 0.1% of the physical records.

and software platforms. What is needed is a new richer standard to serve the function now served by flat file ASCII.

## 4. Indexing and retrieval methods

Part of the integrity and access issue involves indexing, searching and retrieval technology questions. These are sufficiently important questions that they warrant separate discussion, in the following section.

## D. Archives Retrieval, Access and Distribution Possibilities

As long as the National Archives involve only paper documents, using the archives necessitates physical access to retrieve information. As a larger proportion of the archives exist in electronic form, a wider variety of access and retrieval means becomes feasible. Technology permits access to electronic documents by persons located long distances from where the documents physically reside, by means of telecommunication links. Automatic indexing techniques can be used to generate inverted indexes from electronic documents permitting users to engage in free text searching. Moreover, electronic documents in limited quantities can be duplicated much faster and more cheaply by copying them onto magnetic or optical media than making xerographic copies of paper documents or paper or photographic copies of microform documents. These accessibility, retrieval and reproduction characteristics of electronic records present opportunities and policy issues not heretofore confronted by the National Archives. In particular, policy makers may be forced to decide if the National Archives should evolve into a kind of central electronic repository for government information, accessible electronically by citizens either directly or through value added intermediaries. Thus, electronic archives issues overlap to a considerable extent with electronic access and dissemination issues involving current agency electronic records.[192]

---

[192]Federal Agency Use of Computers in Acquiring and Releasing Information (Recommendation 88-10), 54 Fed. Reg. 5207, 5209 (Feb. 2, 1989), to be codified at 1 CFR §305.88-10 [hereinafter "ACUS Recommendation 88-10"]; *See* H. Perritt, Electronic Acquisition and Release of Federal Agency Information (Oct. 1, 1988) (report); Perritt, *Electronic Acquisition and Release of Federal Agency Information: An analysis of ACUS recommendation*, 41 Admin. L. Rev. 253 (1989); Perritt, *Federal Electronic Information Policy*, 63 Temple L. Rev. 201 (1990).

## 1. CDROM

CDROM makes information distribution more efficient because much information can be put on a single 3.5" or 5.25" optical disk: 500 megabytes, the equivalent of 275,000 pages of single-spaced, typewritten text.[193] CDROM is a solution to distribution, not storage and access. The relatively high fixed cost of a CDROM disk[194] means that it is justified only when the demand for particular information is expected to be high. It is a promising technique, for example, for distributing NARA materials, and other public information, to public libraries.

The first issue with CDROM is the issue of how actively the government should engage in electronic publishing--one of the core issues addressed in ACUS Recommendation 88-10. For example the Bolder Data Center, formerly a relatively passive archives, began combining and releasing its data on CDROM formats, and the demand for the information soared. This raises the possibility that a more affirmative effort by archives organizations to package information and make it easily useful by client groups, would increase its utility.

A second issue with CDROM is the pervasive format standards issue. Both the relational database and the distributed database concepts raise the possibility that in the future, many databases will be organized with relatively standardized "hooks." Someone wishing to combine the data or to access the data in new ways will be able to access these hooks, even without any effort to combine or organize the data in advance in a macro way.

A third issue with CDROM relates to intellectual property, because existing standards for organizing information (data structures) and for accessing it are proprietary. If one set of hardware and software are used to record the information, it is not clear how the information can be accessed without the same hardware and software being available to a user.

The best solution to this problem undoubtedly is through development of appropriate standards in GOSIP, and compliance with those standards by vendors wanting to serve the government market.

## 2. Dialup Access

Despite the eventual attractiveness of direct electronic links, overnight express of 5.25" diskettes or CDROM disks probably will be more cost effective than dialup access to archival materials for a long time. The major reason for this advantage of physical transfer is bandwidth. 2400 baud or even

---

[193]H. Perritt, How to Practice Law With Computers at 734.

[194]CDROM disks must be "mastered," in order to make multiple copies. Fixed costs for a short run of 500 disks can approach $5,000-$10,000.

9600 baud is not a very efficient way to transfer large amounts of information. The telephone system is not likely to offer higher band widths to ordinary users of information in the near term.  Senator Gore proposed a "national information system," centered on a "National Digital Library," to distribute the 88 million printed and recorded items in the Library of Congress.  James Billington, chief librarian of the Library envisions citizens being able to access Library materials at their local libraries.   The Office of Science and Technology has proposed setting up a National Research and Educational Network, linking 1,000 research sites with the capacity to send 50,000 single spaced typed pages in one second.  OST head Allan Bromley endorsed the concept in a letter to Congress,[195] and NREN is the subject of a conference at Harvard University's Kennedy School in late 1990.[196]

Brian Kahin has observed correctly that network and database technologies blur the distinction between dissemination and access in the context of public information.[197]   Electronic databases made available on networks are "published" in the sense that the custodian of the database takes affirmative steps to make the database available via the network, and the network is available to all its users and subscribers at their places of use like a published paper document is available to them there.  On the other hand, the information content is not actually transferred to the user until the user makes a specific request.  This looks more like access in that the original custodian of the information content keeps it until it specifically is requested.

The collapse of the dichotomy sometimes is labelled "publishing on demand."   The collapse of the dissemination/access dichotomy through publishing on demand has several implications for electronic records management.

It eases the pressure for physical transfer of  accessioned records to the National Archives because requesters can access the records electronically via a network wherever they physically reside.[198] Second, and probably more fundamentally, it blurs significantly the present distinction among publishing government materials, carried out by the GPO and NTIS; providing access to information under the Freedom of Information Act in response to discrete requests; and archiving information.

---

[195] Byte, Dec., 1989 at 18.

[196] The author of this report is a participant in the conference.

[197] Kahin, Toward A Public Information Infrastructure: Information Policy and the Internet, Prepared for the U.S. Congress Office of Technology Assessment at 19 (contract number L3-5445.0, May 15, 1990) ("the breakdown of dichotomies in the network environment.")

[198] This may not be altogether true.  The National Archives provides a safe, permanent place for storing records and ensuring their integrity.  It thus performs a security function as important as its access and retrieval functions.

Publishing on demand--the blurring of the dichotomy between dissemination and access--means that the act of publication has much lower cost associated with it than paper publication. Assuming that appropriate type two, three and four value[199] has been added to the information to permit efficient retrieval on request, the act of publication and the cost of the act of publication is simply the cost of establishing an electronic link between the hardware and software where the information resides and the consumer. Typically this is done by designating specific documents or database elements as available on a network. Because type five, six, and seven value is not added until the information actually is requested, the cost of this kind of publication is much lower than traditional publishing and distribution.

It is tempting to suppose that future archival systems could simply retain all electronic information generated by agencies, establishing electronic links through an appropriate network or networks when the information is suitable for public availability. There are several problems with this vision, however. First, it would be difficult to find storage and processing capacity for the rapidly expanding quantity of electronic data. Moreover, it is not quite as simple as simply establishing an electronic link via a network. To make the information accessible, some kind of index must be generated, either a data dictionary coupled to the database-specific indexes, or inverted files for free text searching of textual documents.

Even if storage and telecommunications links were available, widespread incompatibilities among hardware, operating systems, communications protocols and application software would frustrate effective access to the information. Electronic information in agencies is defined and organized to support the conduct of day-to-day affairs. Third party users are unlikely to know enough about organization structures and responsibilities to be able to look for information in appropriate ways. Effective access requires technical and context documentation. Most real world computer applications are not well-documented, and even when they are, links must be created between computer documentation and agency documentation required under §552 of Title 5 United States Code.

---

[199] *See* §IV(B)(2).

## E.  Impact of Emerging Freedom of Information Act Policies for Electronic Records

Administrative Conference Recommendation 88-10[200] impacts electronic records management in four major ways.  First, Recommendation 88-10 strongly embraces the proposition that electronic information should be treated as agency records for FOIA purposes.  The basis for this conclusion supports treatment of electronic information as records for records management purposes.

Second, it encourages agencies to design FOIA retrieval capabilities into computerized information systems from the outset.  Designing FOIA capability into systems is one aspect of designing records management capabilities into systems from the outset.  The major difference between FOIA and archival needs is that identification of information desired in connection with a FOIA request is made at the time of the request by the consumer of the information.  Designation of information for records management purposes is done in advance, as a part of classifying types of information to be contained in a system, or as a part of the authoring step.

Third, Recommendation 88-10 encourages agencies to experiment with electronic techniques for improving public participation in agency rulemaking and adjudicatory proceedings. As this occurs, more of the public record public input to agency decisions will exist in electronic form. Moreover, the database structures necessary for effective management of public comments also will facilitate management of this information for archival purposes.

Finally, Recommendation 88-10 encourages agencies to use accepted standards for formats and implicitly encourages agencies to participate in improving standards.  This is consistent with the need for better standards for agency electronic records.

## F.  What constitutes a record?

Electronic records management concepts are centered on the concept of a record copy of a document.  It is natural to assume that information comes in packages called documents. But it has not always been so.  Until the thirteenth or fourteenth centuries, information was primarily oral, although much

---

[200]54 Fed.Reg. 5207 (Feb. 2, 1989), to be codified at 1 CFR §305.88-10; Perritt, *Electronic Acquisition and Release of Federal Agency Information: An Analysis of ACUS Recommendations*, 41 Admin. L. Rev. 253, 310 (1989); Perritt, *Federal Electronic Information Policy*, 63 Temple L. Rev. 201 (1990).

valuable information, especially official records of church and state, was inscribed on a variety of media, ranging from clay tablets and papyrus to hand written parchment. Written information became important in the legal reforms initiated shortly after the Norman invasion of England, especially by Henry II. The printing press, once it widely was accepted, transformed social and cultural concepts of information again, leading to new packages such as newspapers and books. Gradually, new typographic forms emerged to facilitate the use of information by a broad range of consumers.

Now cultural and social institutions have grown up around printing press technology, and it is natural to think of information issues only in terms of that technology.

The electronic information revolution is, however, as profound as the printing press revolution in its potential impact on cultural and social patterns for creating and using information. While new information paradigms have not yet emerged with sufficient clarity to be widely accepted, policy makers and designers of long term information policies should be wary of accepting printed information paradigms uncritically.

The appearance of on-demand publishing technologies[201] presents major challenges for any archive policy. If books and newspapers can be produced in many different versions, with press runs on the order of a dozen per version, should one archive all versions, and if not, which version should be archived? A compromise possibility is to archive the most extensive version, or possibly a master copy of all fragments that were included in any version, along with a kind of control document that would correlate document sections with particular versions, showing which section was included in which version.[202] The problem, of course, is the same problem associated with archiving any database: which state of the database is the archival copy?

Electronic records can be thought of as analogs of paper memoranda and reports, or they can be thought of as analogs of telephone conversations. The latter analogy obviates concerns with preserving electronic records because telephone conversations are not preserved as a part of records management and

---

[201]McGraw Hill is phasing in a new textual database and book manufacturing technology that can vary the content of a textbook for press runs of as few as ten copies. *See* McDowell, *Facts to Fit Every Fancy: Custom Textbooks are Here*, New York Times, October 23, 1989, at page D1. This is a major development towards on-demand publishing, representing a convergence of database and print media.

[202]This blends the publishing-on-demand issue with the compound document issues discussed in the next few paragraphs.

archives activities[203] even though the technology exists for recording telephone conversations and saving the recordings.[204]   It may be that some electronic information volatility characteristics impose such challenges to records retention efforts that the telephone conversation analogy is the best place to end up.   On the other hand, the paper memoranda and report analogy is clearly appropriate for electronic information possessing reasonable permanence in content, even though the final content may have evolved through many versions.

A more profound example of the inadequacy of customary information concepts is the impracticability of applying them to compound documents created by electronic technologies.   In the future, the prevalent means of accessing information may be through a kind of master document which really is little more than a set of pointers to textual, graphics, and sound information stored in diverse places. Presently, a single query from a relational database is a precursor of that kind of compound document, although all elements of a database query usually represent similar forms of information.[205] Perhaps the most important long term objective is to develop conceptual frameworks and strategies for dealing with compound documents.

The Patent Office system[206] is an actual example of a compound document system.   USGS maps may be another potential example, because of the desirability of representing maps as modules of small geographic areas.   A large map would be a series of pointers to smaller area modules. Hypertext is another clear example.[207]

---

[203]Of course, it is over reaching to characterize broad classes of electronic records as analogous to telephone conversations.   Moreover, 36 CFR §1222.20 requires that significant decisions reached and commitments made over the telephone should be documented.

[204]Voice recordings are completely sequential in character, so retrieving an item of interest is very time consuming.   In other words, voice recordings have no type two (chunking), three (internal tags and pointers), or four (external tags and pointers) value.

[205]But see dBASE Mac and other products that permit database fields to contain graphical images as well as text or numeric information.

[206]See H. Perritt Electronic Acquisition and Release of Federal Agency Information, Report Prepared for the Administrative Conference of the United States §III(F) (Oct. 1, 1988) (describing patent office system).

[207]Hypertext is a technique for linking different documents or materials with pointers that can be activated by a user. The linked materials may includes graphical images and sound, as well as text. Compound documents are collections of related materials the content of any one of which can change independent of the other materials in the collection.  See H. Perritt, How to Practice Law With Computers at 261, 395 (1988 & 1990 Supp.). Both hypertext and compound documents are analogous to a master contract, all of the terms of which are incorporated by reference, as they may be amended from time to time.

At this point, only questions can be posed to frame the inquiry. What is the relevant information for historical purposes: the state of the database against which pointers were executed? A collection of pointers is one unambiguous representation at a point in time. But pointers alone lack meaning unless one can restore their referents. Should the strategy be to save transactions rather than databases? Then there is no context.

These issues are not unique to federal electronic records management; they will be pervasive as compound documents come into wider use.

One possibility, of course, is that the new information technologies require people to abandon their preoccupation with the state of information at a particular point in time as is represented by a paper document. But if this concept is abandoned, it is not clear how historical developments and change can be recorded.

On the other hand, the new technologies permit more information to be captured about the processes of using information than is possible with paper documentation, which shows only what the information was, not how it was used or by whom.

Some of the techniques developed for managing and archiving audiovisual materials may be useful models for compound documents, although of course, they do not require the management and accessibility of text and sound and video together.

## 1. Proposed taxonomy

The following taxonomy is one way of thinking about electronic formats from a records management perspective, according to their relative content volatility and format complexity:[208]

   a) electronic mail messages

   b) word processing documents [209]

   c) content of electronic databases

   d) compound documents with relatively static components, like maps

---

[208]Format complexity increases as more of the nine types of value, addressed in §IV(B)(2), are added.

[209]Both word processing documents and electronic mail messages have relatively static content and relatively simple formats. Word processing documents are more likely than electronic mail messages to have format features like typeface changes, style sheets, headers and footers and footnotes. Word processing documents also are somewhat more dynamic in that there are more likely to be multiple drafts of a word processing document than of an electronic mail message.

e) compound documents with volatile components, like statistical summaries linked to spreadsheets or databases.

The entries on the taxonomy represent an increasing level of structural complexity. As one ascends the list, the challenge is in deciding what to retain, i.e., what is the "record," and in specifying standards to use for information interchange. Retail electronic information products are difficult to position on the list. Some products, like Census CDROM disks offered to the public, are entirely static, because they are frozen in their physical manifestation. Others, however, like dialup bulletin boards such as USDA's EDI, are inherently volatile.

EMail is a particular problem because users of EMail think of it more like telephone interaction than written exchanges.[210] The problem thus is not developing the technical capability to retain EMail messages for their potential historical value; the problem is in developing user acceptance that EMail messages may be subject to "capture."[211]

Growing use of EMail and development of systems to archive EMail messages have a potential to expand the scope of the historical record of agency decisionmaking significantly, while also presenting the risk of overwhelming archival systems with too much information. The risk can be mitigated by developing automated record schedules that save EMail messages only between certain persons, dealing with certain subjects, or both.

## 2. Shift toward database paradigm

The new technologies have stimulated a gradual and barely perceptible shift from free-form textual information into databases. This undermines traditional assumptions about the nature of databases. The EDGAR system is a strong example. Electronic mail, for example makes textual information look sort of like a database, because certain addressee, author, and subject information is fielded.

Growing use of hypertext and hypermedia concepts will increase the convergence between free-form and database paradigms. A "database view" and a "virtual document" - really is the same concept.[212]

---

[210]See U.N. Report at 8-14.

[211]See U.N. Report at 8-14 (users perceive EMail like telephone conversations, the recording of which is generally unethical if not illegal).

[212]C.Dollar, The Impact of Information Technologies on Archival Principles and Practices: Some Considerations at 7 (delivered at University of Macerata, Macerata, Italy, Sept. 5, 1990) [hereinafter Dollar Macerata Paper].

Will new type databases constitute the only evidence or the best evidence of how government works? For example if senior policy officials manipulate budget options in a Lotus 123 spreadsheet? Is not the spreadsheet itself the best evidence of how policy was made? Saving only the content of spreadsheet cells, without saving the formulas or recording the way the data was displayed to the policymaker in the spreadsheet format may obscure the decisionmaking process.

The question of when to take a snapshot of a dynamic database is not unique to electronics records management or electronic archives. Indeed, knowing what the contents of a database were at a particular time is essentially the question that makes electronic signatures and Statutes of Frauds difficult. How can one prove the state of an electronic record as of a particular instant? In general, using snapshots of dynamic data is not an entirely new problem. The white pages of the telephone book, for example represents such a snapshot.

## G.  Copyrights

One difficulty in moving toward electronic archives is the possibility that intellectual property rights in public data, data structures, or retrieval processes will inhibit appropriate levels of public access. Two underlying intellectual property issues are involved with this concern: first, whether public information can become private intellectual property when it becomes part of a compilation authored by a private entity, and whether intellectual property interests in retrieval processes (software, formats and indexes) may be used to retrieve public information with which they are associated without infringing private sector intellectual property interests.

Opposing these concerns about intellectual property impeding public access are concerns to stimulate private sector investment in making electronic public information more readily accessible. In some cases federal agencies create incentives to stimulate investment in electronic publishing of public data. In such cases, as ACUS Recommendation 88-10, paragraph F recognizes, a sponsoring agency may consider protecting markets for private sector providers prepared to make a commitment to appropriate product features. Such arrangements may permit a private entity to reserve a copyright[213] or

---

[213]Under the Copyright Act of 1976, an information provider under contract to the government is entitled to obtain a copyright in the public information, See H.R.Rep. 94-1476, reprinted in 1976 U.S. Code Cong. Admin. News 5671-5672, (construing limitation on copyright of government works not to prohibit copyright of government-commissioned work produced by a

other proprietary interests in specific value added features,[214] but not in the raw agency information or in value-added features developed by the agency.

These intellectual property interests should not be protected in a way that eliminates public benefits from government-funded development or restricts the public availability of public information. For example, members of the public should be permitted to use value-added features to facilitate disclosure-level information release, much as libraries are permitted to disclose copyrighted works to consumers while not interfering with copyright owner rights, by prohibiting commercial exploitation by a library borrower.

The potential conflict between protection of intellectual property and ensuring public access to archived electronic information is especially acute when new formats are involved. It is important, on the other hand, for agencies to embrace new formats so that the public can have the benefits of the new format. The new formats can be embraced only by respecting intellectual property. Intellectual property protection can impede public availability in three ways:

-- because the existence of the intellectual property prohibits agency release of the information to the public. This is unlikely because copyright interests alone should not exempt information from accessibility under the FOIA.[215]

-- because the agency or the user cannot get the proprietary software needed to use the information, either because it is no longer in print, or because it is too expensive.

-- because the proprietary software needed to use the information is incompatible with the user's or the agency's current hardware or operating systems

## H.  Use of Electronic Records as Evidence

Public records serve an important evidentiary role in resolving disputes among citizens and between citizens and the government. It is important therefore that public records be accepted by dispute resolution tribunals. Changing to electronic formats should not impede the acceptability. There is

---

contractor; recognizing that "denial of copyright protection might, in some cases, hamper the production and publication of important works.")

[214]The value added concept is recognized in copyright law through the eligibility of derivative works for copyright protection.

[215]See Perritt, 63 Temple L. Rev. at 234-40 (explaining treatment of copyright protection under FOIA exemption 3, 5 U.S.C. §552(b)(3)).

no reason to be concerned that well designed electronic records systems will present any new problems with admissibility of electronic records under the rules of evidence.

A major difficulty in appropriate records management strategies is unfamiliarity of the government's lawyers with the technology and with conclusions reached about legal treatment of the technology. For example, agency lawyers repeatedly "reinvent the wheel" regarding admissibility of electronic information and the internal procedure necessary to authenticate electronic information offered as evidence. Similar examples involve electronic signatures.

There is a natural tendency when a new technology raises legal concerns to forget basic principles and to embark on unnecessarily difficult and complex de novo searches for new legal principles to govern the new technologies. This natural tendency should be resisted. A de novo inquiry is not necessary; the areas of concern unique to electronic methods can be addressed best by focusing narrowly on the differences between electronic techniques and conventional techniques.

Reliability of computer records is suspect because of the inherently greater volatility of computer and telecommunications data, compared with written data. But relative reliability of different means of proving legal facts is not a new problem for the law. Historically contract law promoted reliability of contract proof by insisting on written documentation of certain contracts under the Statute of Frauds[216] or of other signature requirements. The same underlying issues are addressed by the parol evidence rule[217] and by rules of evidence pertaining to the inadmissibility of hearsay, subject to exceptions for business records. The standard for the business records exception to the hearsay rule is not necessarily the same as the standards for the Statute of Frauds or the parol evidence rule. Nevertheless, all three legal concepts share a common concern with reliability. All three recognize that some forms of evidence have greater reliability than others.

Document authentication[218] or proof of the inclusion of particular terms in documents requires that the communications making up the transaction be recorded and retrievable in some reasonably permanent way. Statutes of frauds

---

[216]*See generally* Restatement (Second) of Contracts ch. 5, statutory note at 281 (1979) (reviewing state statutes and text of original statute).

[217]The parol evidence rule bars evidence of prior or contemporaneous oral statements to vary the terms of an integrated written contract. See Gatins v. NCR Corp., 180 Ga.App. 595, 349 S.E.2d 818, 820 (1986) (interaction between parol evidence rule and Statute of Frauds barred parol evidence to supply duration term of written employment terms).

[218]Authentication is proving that a document is what it purports to be, that the proffered writing is not a forgery.

require a "writing" as the permanent means of recordation. Statutes of frauds also require signatures, which also are intended to serve an evidentiary purpose.

Writings and signatures serve somewhat different purposes, although their purposes also overlap to a considerable degree. A signature protects against outright disavowal of a document by a party who denies authorship of a writing. A writing memorializes the terms of a document and thus protects against disputes over the terms of a document, admitting that a document exists. Reliability and authentication issues are related to the signature requirement under the Uniform Commercial Code.

Hypertext and compound documents are emerging concepts that represent a high level of incorporation by reference--and major challenges in terms of creating a reliable transaction record.

The Justice Department has not provided recent official guidance on the admissibility of electronically filed federal records as evidence. The most recent effort by the department seems to be November 6, 1986, draft memorandum.[219] The points made in the draft are relatively noncontroversial. The provision of the NARA regulations concerning judicial acceptance of electronic records is a step in the right direction, to give clearer and more official guidance on evidentiary treatment of electronic records.

The often neglected first step is to determine the purpose for which the electronic record would be used: to prove verbal conduct with legal significance,[220] to prove the content of an electronic transaction, such as an electronic contract, or as evidence of whether some other fact exists. Only the third purpose involves the hearsay exclusion. Even in that circumstance, electronic records are admissible under the general rules of evidence applicable to hearsay and document authentication.[221]

---

[219]By George S. Kondos, "Admissibility of Electronically Filed Federal Records as Evidence: a Guideline for Federal Records Managers of Custodians (Justice Management Division, Office of Information Technology, System Policy Staff) (Nov. 6, 1986).

[220]Proving intent, the making of a representation, or a defamatory communication are examples.

[221]*See* Federal Rules of Evidence 1001(1) (defining writings and recordings to include "magnetic impulse, mechanical or electronic recording, or other form of data compilation"), 1001(3) (defining original to include print out or other output readable by sight known to reflect data stored in the computer or similar device accurately), 1001(4) (providing for admissibility of duplicate which includes counterpart produced by mechanical or electronic rerecording which accurately reproduces the original), 1005 (permitting introduction of copy of official records certified by a witness who has compared it with the original), 803(6) (16) (excluding from inadmissibility as hearsay various categories of business records and official records), 901(b) (giving examples of authentication of records); and 44 U.S.C. §3301 (explicitly defining federal records to include machine readable materials regardless of physical form or characteristics).

Despite the applicability of these general legal concepts, however, because electronic records are particularly susceptible to purposeful or accidental alterations, or incorrect processing, authenticating them may be more difficult than for paper records using the same standards of authentication.[222]

## I.   Should the President Be Treated Differently From Agencies?

The only policy reason[223] to treat presidential records differently from other agency records are: (1) presidential records are not covered by the FOIA, and (2) access and dissemination issues are much more sensitive politically.

Presidential libraries[224] present some special problems and opportunities in terms of records decentralization and standard procedures.    Historically, presidential libraries have been dispersed throughout the country, one location for each President. When the libraries contain only paper records, this decentralization represents burdens for users of the information. As more of the contents of presidential libraries becomes electronic, however, telecommunications access can facilitate use of information even if the presidential library itself is physically remote from the user.

In order for this increased electronic accessibility to be a reality, however, uniform standards must be used by all presidents. Because of the discretionary authority of the president and his immediate subordinates, it is more difficult for agencies like NARA and GSA to impose burdensome government-wide procedures and requirements on presidential personnel.  On the other hand, because of the virtually complete turnover of presidential personnel between

---

[222]*See* United States v. Scholle, 558 F.2d 1109 (8th Cir, 1977) (computer storage requires more comprehensive foundation for admissibility, including testimony of procedures for input control such as tests for insuring accuracy and reliability); United States v. Vella, 763 F.2d 86 (5th Cir. 1982) (implying disagreement with *Scholle*; computer data should be treated like any other record of regularly conducted activity); United States v. Rosso, 480 F.2d 1228 (6th Cir. 1973) (authentication of computer records requires establishing reliability and trustworthiness of information put into the computer such as showing the input procedures used, tests for accuracy and reliability, showing that an established business relies on the computer as records in the ordinary course of carrying on activities, subject to opponent cross examination concerning input and accuracy); United States v. Fendley, 522 F.2d 181 (5th Cir. 1975) (preparer of record not required to authenticate).

[223]As §II(B)(7) explains, the existing statutory framework treats presidential records differently from agency records.

[224]Presidential records traditionally are archives in Presidential libraries rather than in National Archives facilities containing agency records.

administrations, continuing organizations like the White House Office of Administration, OMB, and the White House Communications Agency have potentially great influence on technical standardization.

A useful conceptual model is to think of each presidency inheriting a computer system for its records much as the presidency inherits the White House. The computer system would be designed to meet presidential records managements needs. The working records of the presidency would be managed by this computer system, which would be transferred in its entirety to the presidential library, wherever it might be. Standard features of the standard presidential computer system would include a telecommunications interface that simply could be plugged into publicly accessible telephone lines at the presidential library, and security modules that could be set to screen standard document flags indicating public accessibility.

## J.  Standing

One of the current controversies over the law of electronic records management is whether private citizens have standing to sue for enforcement of records management statutes. If they do not have standing, whether agencies comply with obligations to retain records is entirely up to agency heads, the Archivist of the United States, and the Attorney General.

Federal courts have power only over cases or controversies, according to Article III of the United States Constitution.[225] Standing has been described as amorphous, but its conceptual contours are reasonably clear in the administrative law context under a line of cases beginning with *Association of Data Processing Organizations v. Camp*.[226] A plaintiff must have injury in fact, within the zone of interests protected by a statute, the injury must be causally related to challenged agency action, and the judicial remedy sought must be able to relieve the injury.

Some statutes create private rights of action, expressly or impliedly.[227] Other statutes do not create private rights of action but create rights which can be enforced under the Administrative Procedure Act. The Administrative Procedure Act creates a private right of action for review only of agency

---

[225]*See* Marbury v. Madison, 5 U.S. (1 Cranch) 137 (1803).

[226] 397 U.S. 150 (1970).

[227]*See* Cort v. Ash, 422 U.S. 66 (1975); Merrill, Lynch, Pierce, Fenner & Smith, Inc. v. Curran, 456 U.S. 353, 373, 377 (1982) (emphasizing intent component of *Cort*); Omni Capital v. Wolff, 484 U.S. 97 (1987) (applying *Cort* and *Merrill Lynch*).

decisions that cause legal injury.[228] Without injury to interest recognized by another statute or by the common law, there is no right to review under the APA.

Of course Congress can expand standing by expanding legal rights. A perfect example of such expansion is the Freedom of Information Act. The FOIA creates a new right and thereby a new category of legal injury. Anyone showing infringement of this statutory right has standing to sue. There are two ways that records retention plaintiffs can have standing. One way is to demonstrate injury to interests protected by the records statutes, as in *American Friends Service Committee v. Webster*.[229] Another way is to show indirect injury under the FOIA. The problem for the *Armstrong* plaintiffs was that they requested retention of records rather than access to them. Refusal of the retention request was not directly an infringement of rights granted by the Freedom of Information Act, and the *Kissinger* Court held that the records management statutes created no private right of action.

The records statutes almost certainly are aimed at protecting the public interest in having certain governmental records preserved for historical purposes or for accountability purposes, however, as the D.C. Circuit held in *American Friends Service Committee v. Webster*.[230] This is a sufficient right to support reviewability of agency action under the APA.

Independently, it is plausible to suppose that if someone requested records under the Freedom of Information Act and the government promptly destroyed the records, that would injure interests recognized by the FOIA. Or, if an agency anticipates a FOIA request and destroys the documents, it is at least plausible that would be an injury under the FOIA.[231] Those hypotheticals are not too far removed from the theory of *Armstrong*.

Broadening standing so that private citizens could sue to enforce the records statutes would increase the attention given to records management responsibilities, and this would improve matters. The risks, however, are that NARA would loose control to the courts over the definition of records and the development of guidelines for appropriate information management

---

[228]Either "legal injury" (presumably meaning injury as recognized at common law) or injury in the sense of being "adversely affected or aggrieved . . . within the meaning of a relevant statute." 5 U.S.C. §702.

[229]720 F.2d 29, 57 (D.C.Cir. 1983) (finding standing to seek review of agency compliance with records management and disposal statutes under Administrative Procedure Act).

[230]720 F.2d 29, 57 (D.C.Cir. 1983) (finding standing to seek review of agency compliance with records management and disposal statutes under Administrative Procedure Act).

[231]It may be noted that the *Armstrong* plaintiffs' reason for asserting injury to rights protected by the record statutes was that at least some of the electronic records in the PROFS system were presidential records and therefore outside the coverage of the FOIA.

procedures, and that records that should be disposed in a broader perspective will be required to be maintained because of litigation over them.

## K. Standard of Review

The standard of review of agency records management decisions undoubtedly is the arbitrary and capricious standard under §706(2)(A) of the APA.[232] Usually, one can recharacterize arbitrary and capricious action as ultra vires action, which is judicially remediable under §706(2)(C) of the APA.[233] For example, if an agency fails to consider a statutorily mandated factor, such failure is arbitrary and capricious, and the decision reached without consideration of a mandated factor also is in excess of statutory authority or "ultra vires." The ultra vires theory is useful to a challenger however, only when an agency needs statutory authority for its conduct. An agency may have inherent authority to arrange for its records without considering specific factors under an affirmative grant of authority. Agency officials may have the authority to destroy computer files without being given that authority affirmatively by a statute. There are, however, statutory restrictions on records destruction.[234] Therefore, destruction of electronic records without NARA authority can be characterized as ultra vires on the theory that a statutory mandate or prohibition removes any inherent authority to act contrary to the statute.

## L. Standards

Standards are important because they reduce the burden on the archives to maintain multiple retrieval hardware and software systems, and because they facilitate agency design of records management components of information systems.[235] Information in electronic form can be transferred in two basic

---

[232] 5 U.S.C. §706(2)(A).

[233] Such action also may be enjoinable under the doctrine of American School of Magnetic Healing v. McAnnulty, 187 U.S. 94 (1902) (once legal right is shown, federal courts have inherent equitable power to enjoin ultra vires action).

[234] See 44 U.S.C. §3303A and 44 U.S.C. §3314 (prohibiting destruction of federal records without NARA approval).

[235] See generally Protocols Standards and Communication, Inc., The Application of ODA/ODIF Standards 3, 5 (Mar. 1988) (report prepared for the National Archives of Canada)

ways: by exchanging physical media such as tapes or diskettes, or by communication links. Exchange of physical media requires more standards than communication links, but generally offers higher bandwidth.[236]

At the present time, flat file ASCII is an acceptable standard for text information exchange. Increasingly, however, agency information systems accommodate higher levels of typographic format information and compound documents. Soon, losing these value enhancements when information is transferred among agencies are transferred to the National Archives will be unacceptable. More sophisticated text standards already are available in concept. For example, SGML,[237] the subject of an NIST FIPS,[238] can save agencies the costs of developing text structuring formats from scratch. Use of SGML and standard CDROM format and retrieval conventions could increase public access to electronic information in the archives greatly.

Present methods for transferring database information already are inadequate. One can transfer database information in three ways: (1) in the limited form of reports on query results through query languages such as SQL; (2) from active database to active database, which makes transfer independent of proprietary software and hardware; or (3) through flat files accompanied by detailed documentation.[239]

Section IV(D) noted that long term electronic records management philosophies must accommodate changes in the form in which information is recorded and communicated. Especially challenging is adaptation to compound document concepts. Compound documents bear certain resemblances to databases in that they are dynamic and the meaning of a particular entry may depend on the context represented by the content of other entries. A less

---

[hereinafter "Canadian ODA/ODIF Report"] (ability of archives to perform mission depends on ability to access formats maintained by agencies).

[236]Bandwidth is the rate of information transfer. I can transfer 1.6 megabytes of information almost instantaneously by handing you a 3.5" diskette. The same quantity of information would take somewhat more than an hour to transfer on a 2400 baud communications link.

[237]Standard Generalized Markup Language.

[238]FIPS Pub. 152.

[239]See generally National Computer Systems Laboratory National Institute of Standards and Technology, Framework and Policy, Recommendations for the Exchange and Preservation of Electronic Records (March 1989) [hereinafter "Framework"] (by Margaret H. Law and Bruce K. Rosen); Judi Moline, Attachment C to Framework: Recommendations for Document Transfer Standards and their Integration into National Archives Policy (January, 1989) (prepared by the Systems and Software Technology Division of NIST); Wilma M. Osborne, Bruce Rosen and Leonard Gallagher, Attachment D to Framework: Recommendations for Database and Data Dictionary Standards and their Integration into National Archives Policy (prepared by the Information Systems Engineering Division of NIST). See NIST Report §3.2.1.1 at 16-17.

challenging technological development, but nevertheless one that requires more of format standards than a word processing document is the concept of an electronic document (like an electronic contract) structured to permit electronic technologies to be used for legally significant transactions.

The need for a format standard for legal and government information will increase as more legal and government information is communicated and stored in specialized electronic documents[240] structured to permit expert system use.[241] The problem is similar to that relating to databases. Electronic documents are structured, and the meaning of a particular part of the document may not be apparent from the context, but only from knowledge of what the contents of a particular field mean. One can, of course, consider the need for standardizing electronic documents simply as a part of a larger issue of creating format standards for databases. An electronic tax return becomes a part of databases maintained by the IRS, and document generators or diagnostic systems in law firms or agencies would collect the files representing particular client matters or transactions and maintain them in databases.

The difference between electronic documents and databases is that a database usually is composed of many records and is useful in the aggregate. A single electronic document, in contrast, is significant independent of other electronic documents.

Compound documents and electronic legal documents present somewhat different problems for records management. The most difficult problem presented by compound documents is dynamic content. The content of the element incorporated by reference may change even though the cluster of pointers representing the "master document" remains the same. Electronic legal documents are not dynamic in this sense, but they have no meaning independent of a look-up table defining the structure of fields and the meaning of codes in the fields. Thus both compound documents and electronic legal

---

[240]Electronic tax returns filed with the IRS or electronic 10K forms filed with the SEC are electronic documents. The file containing a user's answers to the questions asked by a legal expert system, say a will generator, or a system for evaluating social security disability claims, also is an electronic document.

[241]Standards for legal information, like EDI or SGML, are forms of knowledge representation. One must identify the attributes necessary to represent, say, a contract. Those attributes are the data elements in an EDI transaction set. Even more important, in order to make the information in an EDI transaction set machine processable, one must standardize the values that are acceptable in any particular field. Determining whether a particular contract term should be represented by a particular value from a predefined set in, say, the time-is-of-the-essence field, is a fairly high order of decisionmaking about how to represent legal knowledge. It is analogous in some ways to the quality of judgment exercised by a headline writer in a daily newspaper. The headline writer must be adroit with language in order to fit the headline to the space available, and he or she must be a quick study on what the story to be headlined is really about.

documents raise some of the same problems presented by databases. One practicable solution for electronic legal documents is to retain them in flat file form along with a set of definitions for structure and fields values. A solution to the standards problem for compound documents is much more challenging.

It is not, of course, easy to specify a standard that accommodates future, unforeseeable technologies. But the challenge has been met in the past. Standards for magnetic tape storage developed in the 50's served well for thirty years, and the flat file ASCII standard for text and databases has been adequate until recently. While there are major uncertainties about the best ways to represent, display, transmit, and store graphical images,[242] good federal government standards can drive the market and create economic incentives for vendors to adhere to the standards rather than diverging on proprietary approaches to obtain competitive advantage.

## 1. Standards Activities

The National Institute of Standards and Technology ("NIST") prepared background papers for NARA in the fall of 1989 addressing such things as standards for transfer of textual documents.[243] The purpose of the NIST report was to identify a logical architecture for the representation, transfer, storage, and access, of electronic records to be accessioned by NARA.[244] The NIST report notes that "records that have been adequately preserved and physically maintained may still be lost due to lack of indexing information. To be accessible, electronic information must have record indexing, cross referencing, and descriptive information available.[245]

The NIST study concluded:

1.    Electronic information accessioned by the Archives will not be usable over a long period of time unless format standards are developed and enforced by NARA. Otherwise machine

---

[242]It should be noted that optical storage does not mitigate the need for standards; it increases it.

[243]National Computer Systems Laboratory National Institute of Standards and Technology, Framework and Policy, Recommendations for the Exchange and Preservation of Electronic Records (March 1989) [hereinafter "NIST Report"] (by Margaret H. Law and Bruce K. Rosen); Judi Moline, Attachment C to NIST Report: Recommendations for Document Transfer Standards and their Integration into National Archives Policy (January, 1989) (prepared by the Systems and Software Technology Division of NIST); Wilma M. Osborne, Bruce Rosen and Leonard Gallagher, Attachment D to NIST Report: Recommendations for Database and Data Dictionary Standards and their Integration into National Archives Policy (prepared by the Information Systems Engineering Division of NIST).

[244]See NIST Report §1.3 at 3.

[245]NIST Report §1.2 and 2.

obsolescence will prevent access to information stored in proprietary formats--even "industry standard" formats like MSDOS 5.25" 360 kbyte magnetic disk, or Word Perfect--over time.

2. NARA should use an SGML based standards family for textual information, including SGML, DSSL, SPDL, and ODA/ODIF.[246] Together, these standards define document elements referred to by SGML tags and permit compound documents to be represented, transferred, stored, and retrieved.

3. Databases should be transferred in flat file form, with accompanying electronic documentation of the database schema. NARA should select a standard database to use for retrieval, and should update media and technologies used to maintain the electronic database instructions more frequently than the content of the database transferred. Newly adopted IRDS should be adopted by NARA. NARA should maintain in electronic form the necessary descriptive and indexing information.[247] The report recommends creation and maintenance of a NARA-wide data administration function using data dictionary techniques.[248]

4. NARA should review the suitability of NARA electronic media and format procedures in view of current technology at approximately five year intervals. NARA should consider transferring information from older media to newer media,[249] as technology changes.

NARA responded and commented on the NIST recommendations in June 1990.[250] NARA agreed on the utility of standards, though it was somewhat less optimistic than NIST about the utility of moving quickly to adopt standards for either text or databases. NARA expressed a preference, explaining the basis for its preference, for working with standard setting organizations to refine the evolving standards for text and database transfer, for waiting until existing standards are reflected in commercial products, and for affirmative steps to encourage agency experimentation and demonstration programs for the more promising standards. The specifics of these NARA intentions are treated in the recommendations section of this report.

---

[246]Each of these acronyms is explained in the subsections following this overview of the standards problem.

[247]*See* NIST Report §3.2.1.2 at 18.

[248]*See* NIST Report §3.2.2.1 at 20.

[249]*See* NIST Report §5.3 at 37.

[250]NARA Technical Report No. 8 (June 1990).

The NARA response focused on certain important implementation issues more than the NIST Report. Simply adopting existing standards is not necessarily practicable. For example, a Postscript-like SPDL has a large amount of overhead. In one experiment, the Postscript version of a file was about forty times longer than the original file. In the author's experience, the Postscript version of word processing files regularly is at least five times as long as the word processing file itself. When storage is an important consideration, this kind of overhead is unattractive.

A major impediment to the adoption of standards by commercial software vendors is uncertainty with respect to the market for products embodying the standards. Federal agencies and the archives represent a substantial market. An aggressive approach by the federal establishment can hasten the maturation of standards. A wait and see attitude, on the other hand, actively discourages the maturation of standards.

It well may be enough, for example, to stimulate word processing software vendors like Microsoft and WordPerfect to add ODA/ODIF features to their leading products or to add an SGML conversion feature if the federal establishment were to embrace these standards.[251]

The National Archives of Canada has identified the following international standards of particular interest for government archives and electronic records management purposes: International Standards Organization (ISO) Technical Committees TC46 (documentation), TC97 (information processing), TC154 (documents and data elements in administration, commerce and industry), and TC184 (industrial automation systems). TC97 is divided into the following subcommittees: SC2 (character sets and information coding), SC2 (telecommunications and information exchange between systems), SC15 (labeling and file structure), SC18 (text and office systems), SC20 (data cryptographic techniques), and SC 21 (information retrieval, transfer and management for OSI).[252]

Of somewhat less interest but nevertheless pertinent are certain study groups of the CCITT: SG II (operational aspects of telematic services), SG VII (public data networks), SG VIII (terminal equipment), SG X (languages and methods for telecommunications applications), SG XII (SDN and telephone network switching and signaling), and SGXVIII (digital networks including ISDN). SGVII produced X.25 and X.400 and other X series standards for the transport and session layers of the OSI model. SGVIII is responsible for

---

[251]Substantial trade press attention has been given, for example, to the GOSIP standard, and further elaboration of this standard to embrace SGML as well as ODA/ODIF could be expected to encourage private software development. See generally, Technical Information Paper No. 8 at 15 (reviewing status of GOSIP and its inclusion of ODA/ODIF but not SGML).

[252]McDonald 1988 Archivist Paper at 5.

teletext, videotext, and facsimile recommendations, including T.400, "document transfer and manipulation," concerned with document architecture in the interchange of documents.[253]

The following standards are of particular interest: ODA/ODIF, MACDIF (Map and chart digital interchange format), ISO8211 (data descriptive file for information exchange), FTAM (file transfer and access methods) and TOP (technical and office protocols).[254]

## 2. OSI

The International Standards Organization is developing an Open Systems Interconnection (ISO/OSI) standard divided into seven layers.[255]   Layer 1 defines the physical and electrical connection between the computer and the network and defines the network's topology.   Layers 2 and 3 define the data link, in terms of the structure of packets used to address other stations and to transmit and receive information, and how packets are routed and relayed. Layer 4 defines how physical locations on the network can be addressed. Layers 5, 6 and 7 define how applications interface with the network.[256]   OSI is not completely defined and therefore is not widely available as a useful standard in actual products.[257]   OSI is, however, a widely accepted conceptual framework for thinking about specific standards and conventions for communicating and transferring information among different computer systems.

## 3. ASN.1/OSI

Abstract Syntax Notation 1 (ASN.1)[258] defines documents in the Office Document Architecture ("ODA") and Office Document Interchange Format ("ODIF") and permits data interchange in the Information Resource Dictionary System ("IRDS")[259] and Remote Database Access ("RDA") standards.   NIST used ASN.1 for the document transfer prototype software developed in its

---

[253]McDonald 1988 Archivist Paper at 5.

[254]McDonald 1988 Archivist Paper at 6.

[255]*See* ISO 7498-1984, Information processing systems, Open Systems Interconnection Basic Reference Model, available from Sales Department, American National Standards Institute, 1430 Broadway, New York, NY 10018. (212) 642-4900.

[256]*See* Byte, July 1987, at 152.

[257]See generally H. Perritt, How to Practice Law With Computers chap. 3 (1988 & Supp. 1990) (explanation of standards).

[258]ASN.1 is a component of OSI application layer protocol level.

[259]IRDS is defined in FIPS 156.

report because ASN.1 is sufficiently robust to be useful for defining arbitrary and complex data types and values.[260]

## 4. IRDS

An Information Resource Directory System ("IRDS") is an enhanced data dictionary. A data dictionary is a system database that contains definitions and descriptions of data stored in user databases. A data dictionary might include the following kinds of information about a particular document, file, or data value:

-- Category of the data item
-- Relationship of the data item to other data items
-- When and by whom the data were created
-- When and by whom the data were modified
-- Total number of modifications since the data were created
-- Description of the data, such as format and range of values
-- Databases or files in which the data items appear
-- Location of the data item in those databases or files.[261]

Data Dictionaries sometimes are integral with database management systems, and sometimes they are developed and implemented separately.[262] An IRDS typically has data organized in four levels. The lowest application level has specific values corresponding to entities in attributes.

The second IRD level has descriptions of the types of data stored at the first level. This description is known as a "schema." The third level has data constituting the schema for the IRD level, and the fourth level has data defining the schema for the third level.[263]

The IRDS concept would permit a uniform approach to a logically standardized, centralized, shareable database for all catalogued information.[264] Each agency might maintain a description of information resources in a directory using a standardized IRDS. Then, the archives organization would access the descriptions through IRDS service interface. Nonelectronic as well as electronic records can be included.[265] The same approach could be used to

---

[260]NIST Report §2.3.3 at 8. NIST also reports that ASN.1 is expected to be incorporated into SQL within a few years.

[261]McDonald 1989 Archivist Paper at 3-4 (close paraphrase of description).

[262]McDonald 1989 Archivist Paper at 4.

[263]McDonald 1989 Archivist Paper at 4.

[264]See McDonald 1989 Archivist Paper at 5; see generally 55 Fed.Reg. 11424 (Mar. 28, 1990) (NIST release on POSIX, discussing data base management standards SQL (FIPS 127) and IRDS (FIPS 156)).

[265]McDonald 1989 Archives Paper at 5.

manage transfer of electronic information from the agency to the archives.[266] A combination of IRDS and SQL or a specification of SQL features would permit easier archives access to agency records, and public access to both.[267]

The basic scenario could involve all departments and agencies using a standardized IRDS along with SQL. The original database environment need not be compatible with the archives system; the only requirements are that both support the IRDS and SQL standards.[268]

The IRDS approach would facilitate determination of data holdings by providing a single standardized dictionary of data holdings, eliminating the need either for hard copy indexes or for separate formats for dictionaries of each agency database.[269] IRDS also would facilitate the appraisal process because it includes structures specifically oriented toward appraisal, such as structures containing information about provenance and original order or context of the records[270] Acquisitions will be facilitated by IRDS because of the potential for standardizing export/import files.[271]

Most significantly, IRDS provides a feasible way for preserving the long term accessibility of agency created databases without the need to maintain an exploding universe of database management systems used to create the databases.[272]

The status of the IRDS standard is as follows. The ISO version has diverged to some extent from the ANSI version, but the ANSI version is more mature, likely to result in commercial products early in the 1990's.[273] A Canadian archives commission study of IRDS resulted in six recommendations, including development of a pilot project based on the ISO, IRDS, and acquisitions of a commercial ANSI, IRDS product.[274]

In the U.S. Accredited Standards Committee X3H4 voted in 1983 to adopt the National Bureau of Standards draft of a Federal Information Processing Standard for data dictionary systems as the basis for further work on IRDS. Since then, NBS and X3H4 have worked together to develop a Draft Proposal American National Standard ("DPANS") for information resource dictionary

---

[266]McDonald 1989 Archivist Paper at 6.

[267]McDonald 1989 Archivist Paper at 7.

[268]IRDS status report at 5.

[269]Protocols Standard and Communications, Inc., Situation Report on the Information Resource Dictionary System ("IRDS") (March, 1989) (PAC-ARC002-1 prepared for the National Archives of Canada) [hereinafter "IRDS Situation Report"].

[270]IRDS Situation Report at 30.

[271]IRDS Situation Report at 30.

[272]See IRDS Situation Report at 31.

[273]Id. at 34.

[274]Id. 37-38 (recommendations d, f).

systems. The DPANS, IRDS was approved as ANS X3.138 in October 1988. NIST officially adopted X3.138-1988 as FIPS156.[275]

## 5. SQL

Structured Query Language ("SQL") is a database language standard embraced by the American National Standards Institute ("ANSI.") SQL has three components: a data definition language, a data manipulation language, and a data control language. Each of the components can be executed directly or indirectly, through a database application development language.[276] SQL is a programming language, designed to implement the relational database model faithfully. It permits a database structure to be specified unambiguously, and for queries to be expressed generically. Microcomputer and mainframe computer software developers are rushing to include SQL compatibility in their products.

## 6. File Transfer, Access and Management ("FTAM")

ISO Standard 8571, File Transfer, Access and Management ("FTAM"), provides the capability to exchange data files and to manipulate data files remotely. It specifies an OSI application layer protocol, capable conceptually of accommodating relational and network databases, as well as ODA/ODIF documents.[277]

## 7. EDI

Electronic Data Interchange (EDI), is a set of standards generally grouped under ANSI X12. Under the standards developed by EDI and ANSI X12 groups, a growing proportion of businesses contract with each other electronically by exchanging prescribed data sets[278] between their computers. A purchase order might be issued electronically, or a request for bids and responding bids exchanged electronically.

EDI originated as a family of standards for exchanging numerical information as a part of electronic business transactions. EDI's heritage is in database schemas, where the meaning of a data element is determined in part[279] from its position in a stream of data or in a file. An EDI standard

---

[275] See 54 Fed. Reg. 13729 (Apr. 5, 1989).

[276] See H. Perritt, How to Practice Law with Computers 389 (1988 & Supp. 1990).

[277] See Canadian ODA/ODIF Report at 8.

[278] These prescribed data sets or data structures are called transaction sets.

[279] The meaning also is determined by the value of the element.

might say[280] that the first seven bytes are the message number, the next ten bytes the purchase order number, the next twenty five bytes a textual description of the product ordered, the next six bytes the total price, and the final ten bytes the signature.

NIST gave short shrift to EDI as a possible standard for electronic archives, but it could become more interesting as its features converge with those of SGML. EDI standards development is moving toward accommodating document types of interest to federal agencies. Transaction Set 864 for text messages encompasses contracts.[281] Transaction Sets 848 (Materials Safety Data Sheet), and 841 (Product Data) move EDI away from strictly commercial purchase transactions to encompass a wider universe of structured documents. EDI may be well suited for administrative records.[282] It also may represent a framework for developing a better database standard, related to SQL.

EDI is essentially one dimensional, while SGML is more two dimensional.

## 8. SGML

Standard Generalized Markup Language ("SGML") is defined in FIPS Pub. 152. A generalized markup language defines styles such as "headline1," "headline2" and "bodytext" in a header, in terms of their typeface, point size and style. Then portions of text to be set as headline1 would be marked "headline1," body text would be marked "bodytext," and the typesetter or printer driver would perform the procedural steps to set the marked text according to the defined styles. Procedural knowledge is separated from the text, but no content is indicated directly by the markup codes. SGML is a standard way to specify the tags and procedural definitions. Both SGML and page description languages like Adobe's Postscript originated in connection with typesetting and desktop publishing.

SGML, like the other text standards considered by NIST, accommodate compound documents.

---

[280] The example is greatly simplified from a real EDI standard.

[281] Transaction Set 864, however, does not permit contracts to be processed by a computer. Early efforts are underway to launch an electronic contract transaction set drafting effort within X12.

[282] *See* §IV(B)(1) for a the three suggested categories of agency information: decisional, administrative and technical.

ISO Standard 8879[283] notes the interaction of SGML with word processing features like WYSIWYG.[284] It implicitly distinguishes between SGML and an EDI transaction set by noting that the ISO standard does not identify or specify "standard" document types, document architectures, or text structures.[285] It does however, note that the efficiency of document formatting can be improved by the use of context from document descriptions: "element declarations."[286] At several points, the standard implicitly or explicitly suggests X12-like efforts to develop standard document type definitions, rather like EDI transaction sets.[287] The supporting documentation explicitly encourages X12-like efforts[288] to standardize document type definitions and other markup constructs.[289] It gives the "Blue Book" citation convention as a legal example of such a standard.[290]

The Association of American Publishers has adapted SGML for author manuscripts, electronic composition and for electronic publishing formats like CDROM. There has been some movement toward use of the AAP SGML standard, but it is hardly in universal use.

SGML is useful but it is not the only standards approach suited for the long run. SGML is not a part of OSI or GOSIP, although the Defense Department CALS[291] program mandates SGML for certain purposes. In order to be useful, SGML requires a parser to reconstruct the document, and it does not easily permit compound documents, although it does accommodate pointers. On the other hand SGML can represent anything, so it is inherently a flexible foundation for a collection of standards, including supplementary definitions of document types and presentational options.

SGML in the electronic archives context requires the following supplementary standards:

---

[283]International Standard 8879 - Information processing - Text and office systems - Standard Generalized Markup Language (SGML) (1st ed. - 1986-10-15) (Rev. no. ISO 8879-1986(E)) [hereinafter "ISO Standard"].

[284]ISO Standard at 59 (word processing); id. at 100 (WYSIWYG).

[285]ISO Standard para 1(a) note (a).

[286]ISO Standard at 21 (para. 6); id. at 63.

[287]ISO Standard at 66 (para B.1.2); id. at 71 (B.4.2- allowable structures); id. at 74 (attributes and chart resembling transaction set illustrations).

[288]Document structures and entities are user-definable in SGML. ISO Standard at 11 (para 4.120).

[289]ISO Standard at 110.

[290]ISO Standard D.1.1(a).

[291]Computer Assisted Acquisition Logistics System.

1. A Document Type Definition ("DTD") for each document or record type.[292]
2. Document Style Semantics and Specification Language ("DSSSL.")[293] DSSSL supplies formatting or presentational information for an SGML document that already exists. It is equivalent to stylesheet definitions in sophisticated word processing programs like Microsoft Word. DSSSL is the link from SGML to SPDL.
3. SGML Document Interchange Format ("SDIF"). SDIF is used to interchange SGML text files. It specifies a data structure which allows an SGML marked-up document developed in separate parts to be packed into a data stream for interchange in a manner that permits the recipient to reconstitute the separate entities.[294]
4. Office Document Language ("ODL")[295] ODL specifies how SGML documents can be coded so they are equivalent to ODA documents. ODIF and ODL representations are technically equivalent, permitting a document to be transformed from one to the other without loss of information about the document constituents and attributes.[296]

SGML's application in the SEC EDGAR system illustrates the potential convergence between SGML, EDI and page description languages. SGML is used by EDGAR to indicate content, rather than style alone. Ultimately, it is reasonable to expect a convergence between SGML and potential EDI standards for text. These standards approaches may merge conceptually with a merger of formatting technologies for printed impressions and for database management and retrieval and may eventually include a merger between page description and hypertext.

## 9. ODA/ODIF

ODA/ODIF[297] is regarded by many knowledgeable students of standards as the most promising standards effort for federal electronic records.[298] It accommodates compound document characteristics. To date, most of the

---

[292] Appendix C at 12.

[293] ISO-TECJTC1/SC18-WG8N606.

[294] ISO 9069-1987; Appendix C at 12.

[295] ISO 8613-5 (1988) (an SGML application; see attachment C at 13).

[296] *See* Appendix C at 20 (quoting ISO 8613-5).

[297] ISO 8613-1988. ODIF is a data stream used to transmit ODA-structured documents. Canadian ODA/ODIF Report at 26.

[298] *See* Canadian ODA/ODIF Report at Annex A (comparing SGML and ODA/ODIF).

ODA/ODIF development effort has been European and Canadian rather than American. ODA may be more machine dependent than SGML because it has format information in it. Parts 1-6 of the ODA/ODIF standard were approved in January 1988.[299] The current standard defines architectures for characters, bitmapped graphics images and object-oriented graphics images.[300] Future extensions are anticipated to encompass digitized sound, color and spreadsheets.[301]

ODA/ODIF permits exchange of documents in formatted form, permitting display of an image, but not further processing, or in processible form.[302] ODA provides two complementary descriptions of a document: a logical structure, composed of hierarchically arranged content elements like sections and paragraphs; and a layout structure, composed of presentation elements like pages and columns.[303] A document profile accompanying an ODA document specifies character sets, styles, and presentation device requirements.

The ODA/ODIF standard assumes an abstract document processing model that encompasses the major stages of adding value introduced in §IV(B)(2) of this report.[304]

The National Archives of Canada is sponsoring pilot projects with ODA/ODIF in an agency environment.[305] A report commissioned by the Canadian archives concluded that most major software vendors have expressed their intention to support ODA/ODIF, although no commercial products exist yet.[306]

The NIST OSI implementors' workshop has established an ODA special interest group to develop pilot ODA profiles.[307]

## 10. SPDL

Standard Page Description Language ("SPDL")[308] is a page description language, functionally like Adobe Postscript. SPDL is not revisable, which has some advantages for archival documents. The NIST report suggests that

---

[299]Canadian ODA/ODIF Report at 17.

[300]Canadian ODA/ODIF Report at 24.

[301]Canadian ODA/ODIF Report at 24.

[302]Canadian ODA/ODIF Report at 17.

[303]Canadian ODA/ODIF Report at 20.

[304]*See* Canadian ODA/ODIF Report at 29 (discussing and distinguishing editing process, document layout process, document imaging process).

[305]*See generally* Canadian ODA/ODIF Report.

[306]Canadian ODA/ODIF Report at 5-6.

[307]Canadian ODA/ODIF Report at Appendix C, page 10 (discussing conformance testing).

[308]ISOJTC1/SC18/WG8N561 (third working draft February 19, 1988).

NARA maintain documents in SPDL format while the agencies may wish to keep them in ODA/ODIF format so that they can be revised and portions reused.

## 11. Applications Portability

Most of the standards introduced in this section relate to portability of files among hardware platforms and operating systems. It also is important that applications software be portable. A report prepared for the National Archives of Canada[309] identified the following properties for application portability across different hardware and operating systems:

-- Consistent user interface, permitting a user trained on any one platform to use any other platform and use the application without additional training.
-- Same set of functions and facilities on all systems.
-- Exchange of data generated on one system to or from another system, including alphanumeric data or more complex structures such as documents, spreadsheets, and graphical images.

The report identified POSIX,[310] the NIST Application Portability Profile,[311] and the X/open activity, the Open Software Foundation, IBM's SAA and X/Windows as approaches that could result in application portability.

The APP architecture divides the operating environment in which an application resides into six functional components: operating system, database management, data interchange, network services, user interface, and programming services. The APP defines a set of standard elements for each component using nonproprietary standards.[312]

## 12. Uses of standards by agencies and the National Archives

The NIST report suggests that documents be marked with SGML tags, their format described using DSSL, and then translated to SPDL for storage and retrieval. ODA/ODIF documents also could be translated to SPDL for storage and retrieval. ODA structured documents would be represented for interchange by ODIF, while ODA structured SGML documents could be represented for interchange by ODL, permitting documents to be converted

---

[309]Protocol Standards and Communication, Inc., 1757 Bank Street, Ottawa, Ontario K1V7Z4, Application Portability (FE6 1989 PSC-ARZ003-1).

[310]*See also* FIPS Pub. 151-1, 55 Fed.Reg. 11424 (March 28, 1990) (adopting POSIX standard).

[311]*See also* Proposed Federal Information Processing Standard (FIPS) for the User Interface Component of the Applications Portability Profile, 54 Fed.Reg. 24372 (June 7, 1989).

[312]Canadian 1989 Status Report at 2.

from one to the other without loss of information about the document constituents and attributes.[313]

An important limitation of some of the standards candidates is that documents following them are not revisable. Nonrevisability may be a virtue for the basic archives document, but assuming that archives users get their own copies of archives documents, they very well may want to be able to revise them, if only to incorporate portions of them into other documents.

Standards by themselves do not resolve important electronic records management issues. For example, ODA/ODIF cannot be applied as an archival transfer standard unless retention and disposition of electronic documents is managed effectively among various users of a LAN. A data interchange format such as ISO 8211 cannot be effective if data are poorly documented or cared for.[314]

An important use issue is whether an archives organization should provide conversion services or otherwise assist in converting data to formats maintained by an archives organization to formats desired by users. This is of course the same issue as FOIA electronic formats questions.

While the National Archives is correct not to rush to a standard that has not been tested and for which software is not available, it may not be appropriate for the National Archives to wait until comprehensive records management procedures are put in place. As other parts of this report note, appropriate records management procedures require reforming organization structures and cultures. This is a process that is difficult in the best of circumstances and probably never can be achieved perfectly. While shortcomings will continue to exist in records management procedures, agencies continue to turn out large amounts of information, more and more of it in electronic form. The National Archives must develop an interim strategy to preserve and to accessions this mass of electronic data while it works to improve the organizational environment within which the information is generated and recorded in the first place.

---

[313] *See* NIST Report Appendix C at 20 (quoting ISO 8613-5). ODIF is defined in 8613-5 as is ODL.

[314] *See* McDonald 1987 Archivist Paper at 7.

# V.  Institutional Arrangements

## A.  Individual Agency Responsibility and Archivist Authority

The most difficult problem in implementing any electronic records management policy is that records management responsibility is dispersed throughout government agencies.  The increasing adoption of microcomputer systems increases this dispersal.  There are no technical impediments to the design of standalone PC-based systems to meet records management needs, and to enhance records management on such systems by making it transparent to users.  But every standalone PC user is the ultimate records manager for the files on her PC, and education is even more necessary than when fewer decisionmakers control time sharing host computers or network file servers.

Corporate culture affects the feasibility of implementing records management decisions.  For example, an organization like NIH has a corporate culture involving very defuse control by central authority.

It is important for everyone to realize that the main problem in effective records management and archives enhancement is not too little information but too much information.  The more an agency or an archives entity keeps, the less likely are users to be able to find what they need. The major challenge is in defining classes of records that should be destroyed in sufficient quantity, not increasing the already excessive legal tendency to encourage people to keep everything forever.

NARA has an inspection program, which is reviewing agency practices at the rate of two to three agencies a year, potentially providing encouragement to senior agency officials possibly reinforced by the Congress or OMB to meet records management responsibilities.

NARA's reluctance to play a stronger enforcement role is motivated by an unwillingness to jump into the middle of politically sensitive issues that high level government officials care a lot about, and the practical unenforceability of over-ambitious and over-mandatory requirements that are not "owned" by responsible agencies in which opposed burdens without meeting immediate agency needs for information.

Most agencies do not defy NARA explicitly, but sometimes they make promises they do not keep.  Some agencies are, however, forceful in refusing to transfer records to the national archives, frequently when they have a strong sense of their own institutional history and constituencies that want direct access to archival documents.

In many respects, the states are ahead of the federal government in taking a leadership role on solving the problems and recognizing the potential of electronic technologies for records management. Leadership at the national level would be enhanced if NARA continues its recent inclination to develop concepts and propose solutions.

## B. Judicial Role

The *Armstrong* case raises institutional responsibility issues that must be confronted in any legal assessment of electronic records retention. First, what is the nature of the injury under the records statutes, the Freedom of Information Act, and the Administrative Procedure Act when electronic records are destroyed? This injury assessment determines whether standing exists. Second, what is the appropriate standard for judicial review, assuming standing exists? The answer to this question determines the scope of discovery and of any trial of a dispute over electronic records management. Does a court examine the merits of retaining documents, or only review the procedure used by the agency in classifying electronic documents? Procedural review may involve review of computer programs that perform the classification functions.

## C. Long Term Federal Government Arrangements

The existence of agency archival records in electronic form, the availability of easily copiable electronic storage media, and potential accessibility via telecommunications links suggest eventual national archives systems that blur traditional institutional boundary lines.

For example, duplication of an optical disk containing archival information at the request of an archives user implicates printing functions historically within the jurisdiction of the Government Printing Office.[315] Maintenance of agency records already overlaps jurisdiction of NARA and GSA. Collecting information for a research use historically was within the jurisdiction of the Library of Congress as well as the NARA. The Library of Congress has an

---

[315]Duplication of records, and single copies on demand, is not, of course, the same thing as mass production printing functions traditionally associated with the Government Printing Office. NARA has made paper and microfilm copies of records for many years. 44 U.S.C. §2307 gives the National Archives Trust Fund Board special authority to duplicate records and print publications.

extensive pilot program underway to develop experience with machine readable formats.[316] To the extent that telecommunications links and new storage media facilitate distributed maintenance of archival information, and remote retrieval of such information, the depository library system or something performing the same function has an obvious role to play. Some of these agencies have not been very effective in playing a leadership role with respect to electronic information. NARA, GPO, and the Library Community are examples. The Library Community has been particularly parochial and unhelpful in formulating solutions as opposed to protecting its traditional turf and existing institutional arrangements.

The are some conflicts among new roles. NTIS wants to charge for its electronic information. If it gives the information to NARA, NARA effectively gives it away, undermining the NTIS pricing structure. The same negative inducement for transferring information to NARA exists with government contractors.

Depositing electronic information with NTIS is not the same as depositing it with the National Archives. The National Archives serves audiences who may not be influential in resource allocation decisions by agencies, by NTIS, or by GPO or the Library of Congress.

The National Archives only recently was removed from the General Services Administration, and it is not constructive to suggest revisiting old organizational arrangements. Nevertheless, it is not inappropriate to think about long term functional realignments. Eventually, it may be appropriate to define a single agency to manage physical storage, including particularly the computer systems on which a growing proportion of agency archival records would be maintained. This agency also would manage telecommunication links integrated appropriately with new high speed research communication systems and the new federal government telephone system. New organizational arrangements may be appropriate for duplicating archival media, and for facilitating user access to archival information. In concept, this would be more efficient than having several different agencies performing all of these functions. The disadvantage would be less flexibility to respond to differing needs of agencies and the communities that use information generated by them.

One must be careful, however, to preserve a locus of archivist expertise and sensitivity to historical uses of information The orientation of the library and archivist professions may be different in important respects. For example, members of the archivist profession say that the library community historically has had little interest in the provenance of information,[317] which archivists

---

[316] *See* generally Permanent Papers at 102.

[317] Provenance refers to the origin or source of information.

understand and support. Some archivists believe that provenance related information issues will become increasingly important for electronic records. For example, knowing the origin or source of particular information may be particularly important in authenticating information for legal purposes. It also may be important to preserve intellectual property rights.  As noted previously, agency information cannot be practically accessible without some significant effort to link the organization of the information to the organization and mission of the agency when it was created.   This linkage and this providing of context is an archivist specialty.

On some occasions at least, NARA has not evidenced in aggressive attitude in getting ready for an electronic records environment.  For example, in answers to questions submitted by Representative Weiss, NARA appeared to take the position that paper formats will be adequate for 5 to 10 years and that something surely would come along to deal with the proprietary formats problem, even while acknowledging  that Lotus 123 files cannot be interpreted completely without reference to the software used to create the files.[318]

A passive NARA is not desirable.  At the very least, NARA should seek out opportunities to do pilot and demonstration projects with standards possibilities for text and database information.  For example, DOD's CALS organization almost certainly would be willing to participate in an SGML pilot.   The SEC's EDGAR project probably also would be willing to participate in an SGML project. The Federal Maritime Commission probably would be willing to participate in proving database standards, possibly including SQL.

Standards is not a technical question; it is a policy question.  Deciding how standards will be set and what they should be involves allocating responsibility between public and private sectors and among public sector agencies.

## D.  Public Private Sector Responsibilities

No strong private sector interest in records management is apparent yet, except by private sector companies hoping for contract opportunities to develop and manage records management systems. Two different kinds of markets are conceivable for electronic records activities.  The first is a market represented by the federal government as a purchaser of hardware, software, and services.  This market likely will evolve around new standards, and

---

[318]*See* Permanent Papers Hearings at 44.

copying and refreshing activities. A second market is a potential market for archived information in electronic form.[319]

NARA identified genealogy as a potential market, and already has sought to serve it directly, with most demand experienced for paper records. Another possibility is climate meteorological information represented by NOAA's 13,000 tapes of satellite information. Nevertheless, no private sector interest has been expressed in market opportunities associated with these data. The ODISS Project involving the state of Tennessee conversion of Civil War records to optical disk formats, while resulting in useful information applicable to other similar projects, confirmed expectations that the records involved were not subject to large public demand.

Eventually, as direct electronic access to government information of all kinds becomes more widely available, private sector opportunities may be more obvious in connection with electronic access to archive information. Indeed FOIA and archives type adhoc access will become less distinguishable from electronic publishing with its inherently greater market.

# VI. Recommendations for Consideration by ACUS

Recommendation 1(a) parallels Part A of Recommendation 88-10. It starts with the premise that the basic policy balances have already been struck and does not seek to reopen them. Existing policy reflected in the records statutes[320] and in National Archives and Records Administration (NARA) regulations and guidelines should be applied to the new electronic formats, with the objective that changing from paper to electronic media should not diminish the historical record of the government or its accessibility. There are some instances in which a rule designed for paper information, when applied to electronic information, may produce significant differences in result.

In other instances, electronic formats present entirely new issues for records management, as with relational databases,[321] whose content is constantly changing, and whose use is different in character from traditional documents. In these instances, NARA and other agencies should identify explicitly the records management and records preservation issues presented

---

[319]NARA notes that it is in the nature of archives that they are used only infrequently, and therefore present only a very limited market, usually not justifying private investment.

[320]*See*, for example, The Federal Records Act and related statutes in 44 U.S.C. Chaps. 21, 22, 29, 31, 33 (1988).

[321]A "relational database" is composed of separate tables from which data are extracted and presented to a user as though they came from one database. A relational database is sometimes also a "distributed" database, meaning that it is made up of tables physically located at different places on a network.

and seek to resolve them in accordance with the basic purposes of a government-wide records management and archives system. The recommendation is not intended to discourage agencies from taking advantage of an enhanced ability to preserve additional records that may result from technological change.

OMB, NARA, and GSA should approve a records management appendix to OMB circular A-130, such as that drafted by NARA. NARA should develop guidelines for federal agency use in implementing the Federal Information Resources Management Regulation (FIRMR) requirement that records management concerns be addressed in the system design of new agency information systems. NARA recognizes the need for this.[322]

NARA should continue and intensify its education efforts to train agency policy level personnel and working level clerical and word processing personnel on how to design records managements systems and how to meet responsibilities under the records statutes.[323]

Recommendation 1(a) also states that the substantive legal rules and the boundaries of judicial review should not be changed by shifts toward electronic formats. This means that the evolution of rules concerning standing to enforce the requirements, and of the relationship between the Freedom of Information Act and the records statutes should continue, and that agency treatment of records in electronic formats should be subject to the same scrutiny as is applied to records in paper formats having the same content. In sum, the guiding principle should be that the content of the record, and not the format of its storage, should control the rules governing its retention and accessibility.

Recommendations 1(b) and 1(c) extend the basic principle of Recommendation 1(a) to public access. Electronic information formats have the potential to permit enhanced public access even as the volume of information grows, because of the potential for better indexes that are computer searchable and the possibility of free-text search. However, a great threat to longterm public access to electronic information formats is technological obsolescence, the possibility that, by the time someone wants to read information from a magnetic or optical disk, computer hardware and software capable of reading the information will not be available. This threat must be avoided--not by refusing to accept electronic information formats, but by working to develop and adopt standards for information exchange. Such standards must also accommodate newer more sophisticated document and database structures such as hypertext--or other compound documents composed of graphical, audio, and video, as well as textual components--and relational

---

[322]Technical Information Paper No. 8 at 20.

[323]Existing and not-yet-released NARA educational materials are a very good start in this direction. The GSA pamphlet also is a useful model.

distributed databases.[324]   Otherwise, solutions to technological obsolescence will themselves become obsolete as agencies adopt future technologies.

Recommendation 1(d) urges that records managers and archivists avoid archival practices that impair the use of electronic information technology in carrying out the agencies' programmatic activities.  For example, it might not necessarily serve the public interest to prohibit standalone microcomputers on the grounds that records management functions can be accomplished with greater effectiveness on time sharing or other network systems. NARA should begin to accept optical media, even though nonproprietary format standards are not adequately defined. NARA should develop interim procedures for recopying and thus preserving the information contained on optical media as format standards evolve. The high density of optical media and the utility of such media for small computers would considerably reduce the costs of agency records management and improve the implementation of records management guidelines.

Recommendation 1(e) encourages agencies to coordinate their use and development of electronic record-keeping technology and standards with the private sector to the fullest extent possible.

Recommendation 2 recognizes the need to give agency managers and counsel guidance as to the admissibility of electronic records as evidence in various proceedings. Without such guidance, agencies may be reluctant to eliminate  redundant paper versions of electronic records.

Recommendation 3 addresses problems relating to preservation of the records of agencies and commissions that are established on a temporary basis.

Recommendation 4 urges that NARA take a more active role in showing agencies how to harmonize records preservation objectives with agency modernization, and in exploring standards that can mitigate potential problems of incompatibility and technological obsolescence.  While NARA's reluctance to adopt document transfer or database transfer standards that do not have an established commercial base is appropriate, NARA should also take the initiative in promoting the development of appropriate standards through private standard-setting organizations,[325] and should encourage agencies to make use of available commercial products embracing the most promising standards.

---

[324]Arelational database is composed of separate tables from which data are extracted and presented to a user as though they came from one database.  A distributed database is made up of tables physically located at different places on a network.

[325]*See* ACUS Recommendation 78-4, Federal Agency Interaction with Private Standard-Setting Organizations in Health and Safety Regulation, 1 CFR §305.78-4.

While NARA's reluctance to adopt document transfer or database transfer standards that do not have an established commercial base is appropriate,[326] NARA should also take the initiative in promoting the development of appropriate standards through private standard setting organizations,[327] and should encourage agencies to make use of commercial products embracing the most promising standards.[328] NARA should evaluate an SGML-based standard for textual information, working with NIST and commercial software vendors.[329] NARA should develop standard document type definitions and other markup constructs. NARA should act on its expressed intent[330] to participate in the development of CALS to address the use of the SGML standard and its implications for broader electronic records management and standards for text and database management. The Defense Department should welcome this participation. In conjunction with its participation in the CALS initiative, NARA should define preliminarily the circumstances under which agency electronic documents should be transferred in a page image that cannot be altered easily, as under the SPDL standard, and the circumstances under which it would be acceptable or preferable to transfer the content of an agency electronic document so that archives users could edit the content further upon obtaining an electronic copy.[331]

NARA should develop new paradigms for databases, determining the relevant information for historical purposes, considering as possibilities the state of the database against which queries were executed, the queries themselves, or periodic "snapshots" of major databases. NARA should develop new paradigms for compound documents, determining how such documents can be archived, beginning with concepts developed for databases. NARA should develop concepts for archiving textual documents published on demand, including guidelines as to which version should be archived. NARA should work with agencies to develop free text algorithms to select documents as candidates for retention, to eliminate the need for human intervention to code

---

[326]*See* National Archives, A National Archive Strategy for the Development and Implementation of Standards for the Creation, Transfer, Access, and Long-term Storage of Electronic Records of the Federal Government (National Archives Technical Information Paper No. 8 June 1990) [hereinafter "Technical Information Paper No. 8"].

[327]*See* Technical Information Paper No. 8 at 17-19.

[328]National Technical Information Paper No. 8 at 8-9 (describing NARA posture on the relatively immature standards for document and database transfer).

[329]*See generally* 55 Fed.Reg. 11424, ___ (NIST release on POSIX, discussing FIPS Pub. 152 standard for SGML).

[330]*See* Technical Information Paper No. 8 at 16 (describing intention to assist in the development of CALS).

[331]*See* Technical Information Paper No. 8 at 814.

documents.     The ultimate goal should be making records management transparent to agency employees.

NARA should encourage agencies to obtain, and ensure that its center for electronic records obtains, experience with IRDS products as they become available.[332]   In the meantime, NARA should announce its willingness to accept SQL formats for databases to be transferred.[333] Ultimately NARA should select a standard database to use for retrieval, and should update media and technologies used to maintain the electronic database instructions more frequently than the content of the database transferred.   NARA should maintain in electronic form the necessary descriptive and indexing information. NARA should undertake a data administration function using data dictionary techniques.[334] Until a NARA database standard is developed and implemented, databases should be transferred in flat file form, with accompanying electronic documentation of the database schema.   NARA should act on its expressed intent[335] to participate actively in a dialogue with records management professionals, standard setting organizations, and software and hardware developers to intensify development of standards for document and database transfer and automated records management.

## RECOMMENDATIONS

1.      Federal agencies, including those responsible for archival and records policy, should ensure that:

(a) changes in the technology of recordkeeping, including the transfer of records from paper to electronic formats, do not (i) affect the content of information that is retained for eventual transfer to the National Archives, (ii) alter the availability of, or criteria for, judicial review of agency compliance with records law, or (iii) otherwise alter the substance of records law;

(b) changes in the format of agency information from paper to existing and future electronic media do not reduce the accessibility of information to the public;

---

[332]Accord Technical Information Paper No. 8 at 12-13.

[333]See Dept. of Commerce, National Institute of Standards and Technology, Approval of Federal Information Processing Standards Publication 127-1, Database Language SQL, 55 Fed.Reg. 3627 (Feb. 2, 1990).

[334]See NIST Report §3.2.2.1 at 20.

[335]See Technical Information Paper No. 8 at 22.

(c) accessibility is not degraded by technological obsolescence of electronic formats;

(d) policies and procedures aimed at enhancing records management should complement and, in any event, not impair the utility of information systems for the performance of agency missions; and

(e) maximum use is made of generally available technology and, whenever feasible, that agencies conform to standards that are widely agreed to and in use in the private sector.

2.     The Office of Management and Budget and the Department of Justice should promulgate guidance to agencies and, as necessary, proposals for statutory changes on questions relating to admissibility of electronic records as evidence and other reliability issues.[336]

3.     Temporary agencies and commissions should, in consultation with NARA, manage their recordkeeping (consistent with the agency's mission) in such a way as to ease the transfer and preservation of their records upon the agency's dissolution.

4.     The National Archives and Records Administration should seek out opportunities for pilot and demonstration projects, covering candidates for standards for text and database information that can ensure the transferability of such information from agencies to NARA and ensure longterm accessibility to the public. NARA and the White House Office of Administration should develop concepts for a turnkey Presidential records system that could go to a presidential library along with electronic presidential records, providing immediate public access to records to which access is permissible.

---

[336]Legal issues relating to reliability include signature requirements and contract documentation. *See*, for example, the federal statutory counterpart to the Statue of Frauds, 31 U.S.C. §1501.

# VII. APPENDIX: OTHER STUDIES AND REPORTS

## A. Interagency Committee on IRAM

In 1986 and 1987 a special task group on electronic recordkeeping of the Interagency Committee on Information Resources Management conducted five case studies of agency electronic recordkeeping. The case studies focused on systems serving program functions that also contain rules or procedures for indexing, storing, retrieving, and disposing of information. Simple electronic mail systems were excluded. The task group examined:
--    the Forest Service information processing system
--    the Federal Trade Commission's legal research system
--    the EPA's Superfund system
--    the Social Security Administration claims modernization project
--    the Air Force Communications Command.
The 1987 Forest Service system was a Data General distributed network involving Data General Eclipse series MB400 and MB8000 minicomputers located in each of about 900 Forest Services offices around the country, supporting about 10,000 workstations serving about 30,000 agency employees. The system included formal and informal correspondence, office automation activities such as electronic record filing, electronic mail, and calendars and most administrative recordkeeping. Formal communication subsystems forced personnel to follow the agency chain of command. The electronic system documented clearances and concurrences, but paper was used as the medium for the official archival record, generated at the time a document was officially signed. No procedures existed for retaining draft documents, and since official records were maintained on paper, retention/disposition scheduling, reports and forms management was not integrated into the automated system.

The Federal Trade Commission Legal Research System ("LRS") indexed and summarized legal, economic, and procedural documents used by the Commission, including correspondence, memoranda, reports and speeches, some of which were prepared outside the agency. The FTC did not consider the contents of the LRS to be Federal Records Act records.

EPA used a manual recordkeeping system, converting any electronic records into paper form.

The Social Security System was aimed at automating the processing of initial retirement and survivors insurance claims, beginning with the 20 most

used forms. Initial claims applications and supporting documentation were created on line via buffer mainframes at six program service centers strategically located around the country. Electronic records were maintained on line for 120 days and then copied onto magnetic tape for indefinite in house retention. Tape records were retrievable within 24 hours. Although there were no "drafts," changes to initial data entry were noted by date, change and individual making the modification. This practice created an audit trail. Records retention policies were somewhat vague.

The Air Force system was a pilot program using off-the-shelf McDonnell-Douglas developed "Augment" software providing electronic mail, document production and control, and personnel information management. Comprehensive document identification was required for each new document, including office symbols, subject, author, suspense date, comments, key words, disposition code, and special instructions. Records were created on hard disk and transferred to magnetic tape if the records were not accessed after 120 days. Documents accessed during the 120 days on line were kept on line for another 120 days. There was no default destruction feature. A NARA-approved interim disposition schedule consisted of five items encompassing the entire universe of documents.[337]

The study group concluded that:

-- "Record keeping rules and procedures should be built into major electronic information systems from the outset.

-- "This means that agencies must involve records managers in the design as well as implementation of electronic information systems."

-- "Records management considerations have direct consequences for system hardware and software; among other things they affect system memory requirements, online response capabilities, magnetic tape costs and storage, space management, and the entire system environment."

-- "Records management design and implementation should entail procedures for auditing information processed by the system, or

[337]The study group thought the Air Force disposition schedule warranted consideration by other agencies. The schedule provides for:

temporary,a retention period of two years

middle,retention period of two years

middle,retention period of eight years

long,a retention period of twenty years,

indefinite,disposition based on appending event or action, and

permanent,records offered to NARA.

designating responsibilities for ensuring that records management procedures are followed, and periodic review for adopted records management practices."

--    "IAC/IRM[338] should request GSA,[339] in connection with the Department of Justice, to supply guidance and develop standards regarding the legal admissibility of electronic records."

--    "Agencies have insufficient awareness and knowledge of NARA's role in records management and overall information management benefits to be gained from good records management practices. This is a problem for OMB and GSA as well as NARA."

--    "The study group suggested that central management agencies might profitably develop a records management appendix to OMB circular number A-130,[340] to accompany relevant portions of the FIRMR[341] and NARA regulations."


## B. National Academy of Public Administration

A 1989 report by the National Academy of Public Administration ("NAPA Report")[342] is a comprehensive and thoughtful discussion of the interaction between electronic technologies and the historical record of the government. The report notes that new computer technologies have potentially great impact on the organization of governmental work[343] but that agencies use the technologies with varying intensity.[344] While increasing use of the technologies creates increased risks of data loss,[345] it also presents the following possibilities:

--    more sophisticated electronic document forms,[346]

---

[338]Inter Agency Committee on Information Resources Management.

[339]General Services Administration.

[340]Circular A-130 deals with management of federal information resources.

[341]Federal Information Resources Management Regulation. *See e.g.*, 41 CFR §201.30.007 (1988) (life cycle design).

[342]National Academy of Public Administration, The Effects of Electronic Recordkeeping on the Historical Record of the United States Government, A Report for the National Archives and Records Administration (Jan. 1989) (hereinafter "NAPA Report.")

[343]NAPA Report at 28.

[344]NAPA Report at 25-26.

[345]NAPA Report at 38.

[346]NAPA Report at 32.

-- "live," instead of "dead" archives,[347]
-- new tools for indexing, storage and retrieval of records,[348] especially if those features and other record management goals are designed for at the beginning of electronic information system design,[349] and
-- a way of dealing with the mountains of paper that represent the real crisis in records management and preservation.[350]

The electronic technologies do not change certain problems of records management:

-- the failure of information system designers to talk to records managers or to take records management into account in designing practical work systems;[351]
-- the difficulty of defining the concepts of "record" and "lasting value" in the abstract so that variations in application can be minimized;[352] and
-- the fact that the real enemy of public access to government records is the quantity of unorganized information, as much as loss or destruction of the information.[353]

The report's recommendations include attention to the need for format standards to facilitate records transfer among organizations, for example between agencies and NARA, or between NARA and citizens, and among technologies,[354] and the need for law reform, including mainly the need for additional authority in the Archivist, and less discretion for agencies.[355]

## C.  1988 NARA Task Force

In 1988 a task force[356] commissioned by the Acting Archivist reported.[357] The task force concluded that the Archivist of the United States should be

---

[347]NAPA Report support for recommendation 4 (discussing "active collections.")

[348]NAPA Report at 45.

[349]NAPA Report at recommendation 6.

[350]NAPA Report at 43.

[351]NAPA Report at 36; recommendation 6.

[352]NAPA Report at 41.

[353]NAPA Report at 42.

[354]NAPA Report Recommendation 6.

[355]NAPA Report, Recommendation 10.

[356]Hereinafter "NARA Report." This effort also is known informally as the "CAPA Report."

given additional authority to issue binding regulations interpreting[358] the existing statutory definition[359] of records. It considered the possibility of amending the FOIA to incorporate the definition of records developed by the Archivist, in order to reduce the tendency of confusion over the FOIA definition of records to interfere with agency records creation and management practices.[360] It recommended legislation emphasizing the need for adequate documentation of the process of decision and policy making,[361] and requiring agencies to promulgate record keeping and disposition requirements in accordance with NARA standards.[362]

The NARA report generally suggests greater oversight by NARA[363] and by agency officials, possibly agency inspectors general. It urged, however, that primary responsibility for creating and maintaining records must remain with agency heads because records creation and maintenance is intrinsic to the administration and functioning of any institution or program.[364] It recommended avoidance of new authority that would compel the Archivist to decide on a case-by-case basis whether something qualifies as a record,[365] rejecting giving the Archivist standing to sue, but taking no position on whether the Archivist should become a special master to courts hearing records controversies.[366] It urged legislation to increase the Archivist's authority to inspect and eventually to "accession" records whose access may be restricted by statute, to limit the effect of *American Friends Service Committee v. Webster*.[367]

---

[357]NARA, NARA and Federal Records: Laws and Authorities and Their Implementation: A Report of the Task Force on NARA Responsibilities for Federal Records and Related Documentation (Feb. 1988) (hereinafter "NARA Report.")

[358]NARA Report at 2 (II(B)); id. at 15 (suggesting language).

[359]NARA Report I(B)).

[360]NARA Report at 17 (suggesting amendment language).

[361]NARA Report at 4 (IV(B)(5).

[362]NARA Report at 15 (suggesting language).

[363]NARA Report at 12, *citing* 44 U.S.C. §§2904(c)(7) and 2906(a)(1) (authorizing Archivist inspections and surveys)).

[364]NARA Report at 14.

[365]NARA Report at III(B)).

[366]NARA Report at 19.

[367]720 F.2d 29 (D.C.Cir. 1983) (affirming district court order compelling development of adequate records disposal plan and enjoining destruction of records until plan developed; finding standing and reviewability). See NARA Report at 20 (suggested language). The *American Friends* case is discussed in §II(D)(2) of this report.

The NARA Report's guidelines cover not only archival records, but all documentary materials. They propose that agencies systematically maintain documents containing data having informational value[368] or that:

    -- facilitate action by agency officials and their successors;[369]

    -- make possible scrutiny by Congress and other institutions "and other persons properly and directly concerned" about the manner in which public business has been discharged;[370]

    -- protect financial, legal and other rights of the Government and of persons affected by governmental actions;[371]

    -- contain essential information on formulation and execution of basic policies and decisions or on major actions;[372]

    -- document significant decisions reached orally face-to-face, by telephone, or in conference;[373]

    -- document important board, committee or staff meetings, or matters considered at or resulting from such meetings.[374]

The guidelines generally are supported by NARA's "documentation standards."[375]

Authority for the guidelines is provided by 44 U.S.C. §§2104(a), 2904(a) and 2904(c).[376]

## D. United Nations

In 1990, the United Nations Advisory Committee for the Co-ordination of Information Systems ("ACCIS") released a report entitled, "Management of Electronic Records: Issues and Guidelines."[377] The report resulted from the work of a technical panel established in 1987 and chaired by Richard Berry, Chief of the Information Services Division of the World Bank. The report

---

[368]NARA Report at A-5.

[369]NARA Report at A-5, citing 36 CFR §1222.20(a)).

[370]NARA Report at A-5, citing 36 CFR §1222.20(a)). The scrutiny guideline is worded so as to support an inference that scrutiny by the general public is not a consideration.

[371]NARA Report at A-5, citing 36 CFR §1222.20(a)).

[372]NARA Report at A-5, citing 36 CFR §1222.20(b) (emphasizing importance of working papers).

[373]NARA Report at A-5, citing 36 CFR §1222.20(b)).

[374]NARA Report at A-5, citing 36 CFR §1222.20(b)).

[375]36 CFR §1222 0.

[376]NARA Report at 12.

[377]Hereinafter "UN Report."

surveys electronic records management within the United Nations system and concludes that all United Nations organizations are "in the midst of significant growth in the use of both electronic records and electronic communication systems; but most are unprepared to cope with them."[378] It identified policy level issues and sets of practical tools for arriving at policy solutions, and reported on emerging technology standards for facilitating information exchange and retrieval.

It suggested the need for policy guidance on the following issues, among others:

-- how should the concepts of record and nonrecord information be defined so that people and systems can implement the concepts?
-- what criteria for retention will lead to acceptable results for electronic records, while being consistent with criteria for paper records?
-- how can timely disposition of records be assured?
-- how can legality of electronic records be assured?
-- what organizational units should be involved in managing electronic records, with what specific responsibilities?
-- who should have physical custody of archival electronic records?
-- how can the format of records be prevented from being a barrier to access?
-- how can intellectual control and documentation be assured?
-- how can accessibility and security needs be balanced?

The report concluded with recommendations for specific steps, and for adoption by the UN of the OSI standard, along with its components ASN.1, ODA/ODIF and FTAM.[379]


## E.  Information Locator Report

In mid-1990, the GSA Regulatory Information Service Center and OMB's Office of Information and Regulatory Affairs released a report prepared for both agencies on the possibilities for electronic federal information inventory and locator systems.[380] The report recommended that OMB formulate policy initiatives for review and discussion, conducting additional research and developing a long range plan for the development of a government wide

[378]UN Report at ix.

[379]These standards are discussed in §IV(L).

[380]See C.McClure, Ann Bishop, Philip Doty, and Pierrette Bergeron, Federal Information Inventory/Locator Systems: From Burden to Benefit (July 27, 1990).

inventory and locator system.  It recommended building consensus among the stakeholders interested in an inventory and locator system and avoiding the temptation to design a system outside the context of stakeholder views and without additional research.[381]

## F.  NAPA/NARA Database Project

The National Academy of Public Administration, under  contract with NARA, has undertaken to review electronic database management practices by federal agencies, with a view toward developing recommendations for appropriate records management and archives practices for such electronic information storage, retrieval and communication techniques. The project report should be available in 1991.  The author of this report is a member of the NAPA panel overseeing the project.

---

[381] *See* McClure et al. at 89.