FEDERAL ELECTRONIC RECORDS MANAGEMENT: WHAT IS THE PLAN? WHAT IS OUR PROGRESS?

HEARING

BEFORE THE

SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY, INTERGOVERNMENTAL RELATIONS AND THE CENSUS

OF THE

COMMITTEE ON GOVERNMENT REFORM

HOUSE OF REPRESENTATIVES

ONE HUNDRED EIGHTH CONGRESS

FIRST SESSION

JULY 8, 2003

Serial No. 108–132

Printed for the use of the Committee on Government Reform

Available via the World Wide Web: http://www.gpo.gov/congress/house
http://www.house.gov/reform

U.S. GOVERNMENT PRINTING OFFICE

93–006 PDF WASHINGTON : 2004

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2250 Mail: Stop SSOP, Washington, DC 20402–0001

COMMITTEE ON GOVERNMENT REFORM

TOM DAVIS, Virginia, *Chairman*

DAN BURTON, Indiana
CHRISTOPHER SHAYS, Connecticut
ILEANA ROS-LEHTINEN, Florida
JOHN M. MCHUGH, New York
JOHN L. MICA, Florida
MARK E. SOUDER, Indiana
STEVEN C. LATOURETTE, Ohio
DOUG OSE, California
RON LEWIS, Kentucky
JO ANN DAVIS, Virginia
TODD RUSSELL PLATTS, Pennsylvania
CHRIS CANNON, Utah
ADAM H. PUTNAM, Florida
EDWARD L. SCHROCK, Virginia
JOHN J. DUNCAN, JR., Tennessee
JOHN SULLIVAN, Oklahoma
NATHAN DEAL, Georgia
CANDICE S. MILLER, Michigan
TIM MURPHY, Pennsylvania
MICHAEL R. TURNER, Ohio
JOHN R. CARTER, Texas
WILLIAM J. JANKLOW, South Dakota
MARSHA BLACKBURN, Tennessee

HENRY A. WAXMAN, California
TOM LANTOS, California
MAJOR R. OWENS, New York
EDOLPHUS TOWNS, New York
PAUL E. KANJORSKI, Pennsylvania
CAROLYN B. MALONEY, New York
ELIJAH E. CUMMINGS, Maryland
DENNIS J. KUCINICH, Ohio
DANNY K. DAVIS, Illinois
JOHN F. TIERNEY, Massachusetts
WM. LACY CLAY, Missouri
DIANE E. WATSON, California
STEPHEN F. LYNCH, Massachusetts
CHRIS VAN HOLLEN, Maryland
LINDA T. SANCHEZ, California
C.A. "DUTCH" RUPPERSBERGER, Maryland
ELEANOR HOLMES NORTON, District of Columbia
JIM COOPER, Tennessee
CHRIS BELL, Texas

BERNARD SANDERS, Vermont (Independent)

PETER SIRH, *Staff Director*
MELISSA WOJCIAK, *Deputy Staff Director*
ROB BORDEN, *Parliamentarian*
TERESA AUSTIN, *Chief Clerk*
PHILIP M. SCHILIRO, *Minority Staff Director*

SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY, INTERGOVERNMENTAL RELATIONS AND THE CENSUS

ADAM H. PUTNAM, Florida, *Chairman*

CANDICE S. MILLER, Michigan
DOUG OSE, California
TIM MURPHY, Pennsylvania
MICHAEL R. TURNER, Ohio

WM. LACY CLAY, Missouri
DIANE E. WATSON, California
STEPHEN F. LYNCH, Massachusetts

EX OFFICIO

TOM DAVIS, Virginia

HENRY A. WAXMAN, California

BOB DIX, *Staff Director*
LORI MARTIN, *Professional Staff Member*
URSULA WOJCIECHOWSKI, *Clerk*
DAVID MCMILLEN, *Minority Professional Staff Member*

CONTENTS

	Page
Hearing held on July 8, 2003	1
Statement of:	
Carlin, John W., Archivist of the United States, National Archives and Records Administration; L. Reynolds Cahoon, Chief Information Officer, NARA; Harriet Riofrio, E-Records Management Policy and Program Lead, U.S. Department of Defense; and Linda Koontz, Director, Information Management Issues, U.S. General Accounting Office	11
Sprehe, J. Timothy, president, Sprehe Information Management Associates; Robert F. Nawrocki, CRM, director, Records Management and Imagining Services Division, Library of Virginia; Caryn Wojcik, State Government Records Management, Michigan; and Dr. Richard Lyusakowski, director, Collaborative Electronic Notebook Systems Association [CENSA]	86
Letters, statements, etc., submitted for the record by:	
Cahoon, Reynolds, Chief Information Officer, NARA, prepared statement of	26
Carlin, John W., Archivist of the United States, National Archives and Records Administration, prepared statement of	13
Koontz, Linda, Director, Information Management Issues, U.S. General Accounting Office, prepared statement of	51
Miller, Hon. Candice S., a Representative in Congress from the State of Michigan, prepared statement of	7
Nawrocki, Robert F., CRM, director, Records Management and Imagining Services Division, Library of Virginia, prepared statement of	95
Putnam, Hon. Adam H., a Representative in Congress from the State of Florida, prepared statement of	3
Riofrio, Harriet, E-Records Management Policy and Program Lead, U.S. Department of Defense, prepared statement of	40
Sprehe, J. Timothy, president, Sprehe Information Management Associates, prepared statement of	89
Wojcik, Caryn, State Government Records Management, Michigan, prepared statement of	101

FEDERAL ELECTRONIC RECORDS MANAGEMENT: WHAT IS THE PLAN? WHAT IS OUR PROGRESS?

TUESDAY, JULY 8, 2003

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY,
INTERGOVERNMENTAL RELATIONS AND THE CENSUS,
COMMITTEE ON GOVERNMENT REFORM,
Washington, DC.

The subcommittee met, pursuant to notice, at 10 a.m., in room 2154, Rayburn House Office Building, Hon. Adam Putnam (chairman of the subcommittee) presiding.

Present: Representatives Putnam, and Miller.

Staff present: Bob Dix, staff director; John Hambel, counsel; Lori Martin, professional staff member; Ursula Wojciechowski, clerk; David McMillen, minority professional staff member; and Jean Gosa, minority assistant clerk.

Mr. PUTNAM. Good morning. A quorum being present, this hearing of the Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census will come to order. I want to welcome everyone to today's hearing entitled, "Federal Electronic Records Management, What is the Plan and What is our Progress?"

In today's world of emerging technology, Federal agencies are conducting more and more of their business through electronic means. As an example, e-mail has become a primary form of communication, often used now in place of the telephone and as a vehicle for sending memos or other important documents. While this has increased the efficiency and productivity of the modern day office, it has also presented a new set of challenges to the issue of records management and preservation as the Federal Government moves from being paper based to more and more electronic and technology based.

The importance of records management and preservation has a long history. Not only is the legacy of our history at issue, but this issue has a direct impact on the continuing ability of the Federal agencies to function properly. Without a comprehensive and cohesive strategy, records can be misplaced or even lost. This not only has the potential to hinder day-to-day operations, but also has potentially significant ramifications on the national archival process.

The management of all Federal records and now particularly electronically generated records creates new and additional challenges. Examples include what types of electronic records need to be saved, what process or technology should be utilized for their

preservation, how will the retrieval of these records be guaranteed and through what process in the future. This whole issue raises another set of questions as well, what happens with all of the existing records in their various formats? How many, if any, of these existing records will be converted to an emerging process or technology? How will all of these records and exhibits continue to be preserved, and how are these records being cataloged? And what is the magnitude of the process required to retrieve an existing record and how are classified versus nonclassified records treated?

The Federal Records Act has provided the National Archives and Records Management Agency with the responsibility for oversight of records management within the Federal Government. Under the provisions of FRA, NARA is to provide guidance and oversight to Federal agencies as they execute the electronic records strategy. Certainly having a common set of goals and objectives in those strategies will ensure consistency and continuity in this process.

Presently, NARA is working to develop the necessary tools to support that guidance. A number of projects including the development of their electronic records archive program and NARA's leadership role as a managing partner of the Electronic Records Management E-government Initiative will put them in a position to help define the national standards for electronic records management.

At the same time, Federal agencies themselves must begin to work more diligently in making electronic records management a higher priority, many agencies have not fulfilled their obligations to confer with NARA about the progress of their records management plans. While NARA has been charged with oversight responsibility regarding these matters, they have been provided little, if any authority, to enforce compliance.

As it currently stands, it is the responsibility of each individual agency to develop and implement an electronic records management strategy. One of the areas we will explore at this hearing is whether the authority and accountability that presently exists related to this issue is sufficient to get the job done.

These are just some of the challenges facing the subject of Federal electronic records management. Through the research and preparation for this hearing, the subcommittee has come to recognize the enormity of the task. It is our objective to learn from a variety of stakeholders, institution, academic, government and private sector, and to work with NARA together with Federal agencies to determine the progress in developing and implementing an effective management strategy.

Today's hearing can be viewed live via Webcast by going to reform.house.gov and clicking on the link under live committee broadcast. It is an important hearing, and we have two distinguished panels of witnesses, and we are fortunate to have a vice chairman of this subcommittee who served as the Secretary of State in Michigan and has worked extensively on this committee in a number of issues, but particularly in the modernization and the utilization of technology to improve the efficiencies of government operations.

[The prepared statement of Hon. Adam H. Putnam follows:]

COMMITTEE ON GOVERNMENT REFORM
SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY, INTERGOVERNMENTAL RELATIONS AND THE CENSUS
CONGRESSMAN ADAM PUTNAM, CHAIRMAN

OVERSIGHT HEARING
STATEMENT BY ADAM PUTNAM, CHAIRMAN

Hearing topic: *"Federal Electronic Records Management: What is the Plan? What is our Progress?"*

Tuesday, July 8, 2003
10:00 a.m.
Room 2154 Rayburn House Office Building

OPENING STATEMENT

In today's world of emerging technology, Federal agencies are conducting more and more of their business through electronic means. As an example, e-mail has become a primary form of communication, often used now in place of the telephone, and as a vehicle for sending memos or other important documents. While this has increased the efficiency and productivity of the modern day office, it has also presented a new set of challenges to the issue of records management and preservation as the federal government moves from being primarily paper based to more and more use of electronic and technology-based tools.

The importance of records management and preservation has a long history. Not only is the legacy of our history at issue, but this issue has a direct impact on the continuing ability of the federal agencies to function properly. Without a comprehensive and cohesive strategy, records can be misplaced...or even lost. This not only has the potential to hinder day-to-day operations, but also has potentially significant ramifications on the national archival process. The management of all federal records...and now, particularly electronically generated records, creates new and additional challenges. Examples would include: what type of electronic records need to be saved?...What process or technology should be utilized for the preservation?...How will the retrieval of these records be guaranteed and through what process in the future?

This whole issues raises another set of questions as well...What happens with all of the existing records in their various formats?... How many, if any, of these existing records will be converted to an emerging process or technology?... How will all of these records and exhibits continue to be preserved?... How are these records being catalogued and what is the magnitude of the process required to retrieve an existing record?...How are classified versus non-classified records treated?

The Federal Records Act has provided the National Archives and Records Management Agency (NARA) with the responsibility for oversight of records management within the federal government. Under the provisions of FRA, NARA is to provide guidance and oversight to Federal agencies as they develop and execute and electronic records management strategy. Certainly, having a common set of goals and objectives in those strategies will insure consistency and continuity in this important process.

Presently, NARA is working to develop the necessary tools to support that guidance. A number of projects, including the development of their Electronic Records Archive (ERA) program and NARA's leadership role as a managing partner of the Electronic Records Management e-Government initiative, will put them in a position to help define the national standard for electronic records management.

Simultaneously, Federal agencies themselves must begin to work more diligently in making electronic records management a higher priority. Many agencies have not fulfilled their obligation to confer with NARA about the progress of their records management plans. While NARA has been charged with oversight responsibility regarding these important matters, they have been provided little, if any, authority to enforce compliance. As it currently stands, it is the responsibility of each individual agency to develop and implement an electronic records management strategy. One of the areas we will explore at this hearing today is whether the authority and accountability that presently exists related to this issue is sufficient to actually get the job done.

These are just some of the challenges facing the subject of Federal electronic records management. Through the research and preparation for this hearing, the Subcommittee has come to recognize the enormity of this task. It is our objective to learn from a variety of stakeholders...institutional, academic, government and private sector...and to work with NARA, together with Federal agencies, to determine the progress in developing and implementing an effective electronic records management strategy.

Mr. PUTNAM. So I now yield for her opening comments to the vice chairwoman of the subcommittee, the gentlelady from Michigan, Mrs. Miller.

Mrs. MILLER. Thank you, Mr. Chairman. I appreciate you holding this hearing today. I'm sorry, I was just a few minutes late. I know you like to start your hearings on time. We should have a hearing on the elevators in this building I think someday. But I certainly appreciate your commitment to exploring ways, examining ways to improve the business of the Federal Government. I think that the challenge facing the Federal Government in regards to electronic records management is certainly very complex. It is extremely daunting. The National Archives and Records Administration [NARA], as we commonly call it, possesses the responsibility to provide guidance and oversight to Federal agencies regarding the records management strategies. Under current law, Federal agencies are required to submit record schedules to NARA, which must be approved and then, in turn, to allow NARA to aid agencies in their records management strategy implementation.

However, as with so many other statutory requirements concerning the modernization and the improved efficiency of government, many Federal agencies are not cooperating. NARA has been working for years on developing an effective electronic records management policy, but the responsibility of implementing the strategy really falls directly on each individual agency, quite frankly.

Today we will be examining the strategy that NARA has developed and the progress that has been made. I feel it is very important to also examine the successes and the failures of Federal agencies in regards to this enormous task. Though each agency is responsible for its own success, the valuable lessons learned throughout government at all levels must be utilized as an important resource. For this reason, I'm very pleased that we're going to be hearing from witnesses at both the Federal and State levels of government, and then representatives from the private sector also.

As the Federal Government progresses into the 21st century and moves to modernize its operations, a problem has arisen in that the Federal agencies seem to place different priorities on utilizing technology. Some agencies have seen the benefits of implementing substantial technologies, while others seem to be stuck in a time warp, stuck in the past. Something needs to be done to force these particular agencies to realize the importance of modernizing their records management.

The current scenario is simply more than an adjustment to the current frame of thinking. It also includes agencies not complying by federally mandated statutes. Currently, there is no real enforcement mechanism that forces agencies to abide by statutes outlined by the Federal Records Act and to submit schedules required by NARA. Many Federal agencies have not placed any degree of priority on electronic records management, and this has resulted, of course, in a wide variety of problems.

I notice Linda Koontz, in her written testimony, states that NARA has developed a strategy for raising awareness amongst agency management of the need to place electronic records management at a very high priority, and I'm also—I found that very encouraging. I'm also looking forward to Mr. Carlin's testimony also,

hoping that he can elaborate on the measures taken by NARA within the last year to impress upon agency officials the need to implement an effective management strategy, and I certainly especially want to thank Caryn Wojcik, who will be testifying in the second panel taking some time to testify before our subcommittee today, as the chairman mentioned, during my tenure as Michigan Secretary of State, the archives were sort of under the umbrella of the Michigan Secretary of State's office, and we think because of the great people that led our agency there, our State's archiving and records management projects, Michigan has certainly become one of the national leaders by incorporating a lot of information technology into the everyday activities of government, and in addition, our State managers have been working very hard to ensure that the State is prepared as its technology evolves well into the 21st century.

So I hope that Ms. Wojcik will be able to inform the subcommittee of some of the different things that have happened. Since I've been gone, I know you're going like rapid fire there as well. But Federal agencies certainly must utilize the knowledge acquired by State governments so we're not reinventing the wheel. Although electronic records management at the Federal level is very daunting, as we say, there are resources available that can enrich the process, and again I reiterate my point that these Federal agencies must make a concerted effort to implement a successful management policy. As records sit untouched or the technology that has been used to create these records becomes obsolete, it is certainly imperative that action be taken and agencies prepare for the future, and certainly it is important that Congress play a very active role in ensuring that electronic records management be implemented effectively and in a timely manner as well, so I'm looking forward to all of the testimony from the witnesses today and I certainly want to thank you all for coming and thank the chairman for calling this hearing. Thank you.

[The prepared statement of Hon. Candice S. Miller follows:]

Congresswoman Candice S. Miller

Opening Statement
Committee on Government Reform
Subcommittee on Technology, Information Policy, Intergovernmental Relations, and the Census
July 8, 2003

OPENING STATEMENT

Mr. Chairman, thank you for holding this hearing today. I appreciate your desire to examine ways to improve the business of the Federal government. The work that you do as Chairman is highly laudable, and I am pleased to work with you as vice chair of this Subcommittee.

The task facing the Federal government in regards to electronic records management is complex and extremely daunting. The National Archives and Records Administration (NARA) possesses the responsibility to provide guidance and oversight to Federal agencies regarding their records management strategies. Under current law, Federal agencies are required to submit record schedules to NARA, which must be approved and, in turn, allow NARA to aid agencies in their records management strategy implementation. However as with so many other statutory requirements concerning the modernization and improved efficiency of government, many Federal agencies are not cooperating. NARA has been working for years on developing an effective electronic records management policy, but the responsibility of implementing this strategy falls directly on each individual agency.

Today, we will examine the strategy that NARA has developed and the progress that has been made. I feel it important to also examine the successes and failures of Federal agencies in regards to this enormous endeavor. Though each agency is responsible for its own success, the valuable lessons learned throughout government at all levels must be utilized as an important resource. For this reason, I am pleased that we will hear from witnesses from both the Federal and state levels of government and from the representatives of the private sector.

As the Federal government progresses into the 21st century and moves to modernize its operation, a problem has arisen in that Federal agencies seem to place differing priorities on the utilization of technology. Some agencies have seen the benefits of implementing substantial

technological, while others seem to be stuck in the past. Something needs to be done to force these particular agencies to realize the importance of modernization. The current scenario is simply more than an adjustment in the current frame of thinking, it also includes agencies not complying by federally mandated statutes. Currently, there is no substantive enforcement mechanism that forces agencies to abide by statutes outlined in the Federal Records Act and to submit schedules required by NARA. Many Federal agencies have not placed any degree of priority on electronic records management, and this has resulted in a wide variety of problems.

Linda Koontz, in her written testimony, states that NARA has developed a strategy for raising awareness among agency management of the need to place electronic records management at a high priority. This is very encouraging, and I look forward to Mr. Carlin's testimony. I hope that he can elaborate on the measures taken by NARA within the last year to engrain upon agency officials the need to implement an effective management strategy.

I especially want to thank Caryn Wojcik for taking the time to testify before the Subcommittee today. During my tenure as Michigan's Secretary of State, the State Archives division was under the direction of my office. Because of the great people that led our state's archiving and records management projects, Michigan has become a national leader by incorporating information technology into the everyday activities of the government. In addition, state managers are working hard to ensure that the state is prepared as technology evolves well into the 21st century.

I am hoping that Ms. Wojcik will inform the Subcommittee as to the details of the projects the state of Michigan has undertaken and the lessons they have learned. Federal agencies must utilize the knowledge acquired by state governments so that we are not re-inventing the wheel. Though electronic records management at the Federal level is a daunting task, there are resources available that will enrich the process. I reiterate my point that these Federal agencies must make a wholehearted effort to implement a successful management policy. As Federal records sit untouched and the technology used to create these records become obsolete, it is imperative that action be taken and agencies prepare for the future.

The number of issues facing government at all levels are immense and challenging, and it is important that Congress play an active role in ensuring that electronic records management be implemented effectively and in a timely manner. I look forward to the testimony of all the witnesses today. Thank you each for coming.

Thank you, Mr. Chairman.

Mr. PUTNAM. Thank you, Mrs. Miller.

We'll now begin with our first panel of witnesses. Each has kindly prepared written testimony which will be included in the record of this hearing, and I ask each of you to summarize your thoughts in a 5-minute presentation to give us ample time for questions and dialog. You'll notice that there's a timer with a light on the table in front of you. The green light means begin your remarks, the point that you see the yellow light, we ask you to sum up, and at the red light, please bring it to a close. In order to be sensitive to everyone's time schedule, we ask that witnesses cooperate with us in adhering to the time system.

As is the custom with the Government Reform Committee, we'll swear in our witnesses. I would ask the first panel of witnesses to stand and raise your right hands, and if you're accompanied by anyone who will be called upon to answer the questions or elaborate on a question, if they would stand as well, please.

[Witnesses sworn.]

Mr. PUTNAM. Note for the record that the witnesses responded in the affirmative.

Our first witness this morning is Mr. John Carlin. As archivist of the United States, John Carlin is the head of the National Archives and Records Administration, an independent agency of the Federal Government with more than 2,900 employees and 33 facilities throughout the Nation. NARA is the Nation's recordkeeper. Its mission is to ensure for the citizen and the public servant, for the President and the Congress and the courts, ready access to essential evidence.

Mr. Carlin was appointed archivist in 1995 by President Clinton. In 1998, he launched a major initiative to build the Electronic Records Archives to preserve and provide access to virtually any type of electronic record created anywhere in the Federal Government.

Prior to being named archivist, Mr. Carlin had a distinguished career in politics, business and education. A native Kansan, he first entered public service in 1971 by serving in the legislature of the State of Kansas. He became Speaker of that State's House of Representatives, and in 1978, won election to the Kansas Governorship serving two terms through January 1987. His fellow Governors elected him chairman of the National Governors Association in 1984. Following his political career, he joined the faculty of Wichita State University teaching graduate courses in public administration.

In 1987 he received an honorary doctorate of laws degree from his alma mater, Kansas State. As Governor of Kansas, he strongly supported the Kansas State Historical Society and the State archives within it, in which he deposited his own gubernatorial papers. Before heading NARA, he served on the National Archives Foundation Board.

Governor Carlin, we welcome you to the subcommittee. You're recognized.

STATEMENTS OF JOHN W. CARLIN, ARCHIVIST OF THE UNITED STATES, NATIONAL ARCHIVES AND RECORDS ADMINISTRATION; L. REYNOLDS CAHOON, CHIEF INFORMATION OFFICER, NARA; HARRIET RIOFRIO, E-RECORDS MANAGEMENT POLICY AND PROGRAM LEAD, U.S. DEPARTMENT OF DEFENSE; AND LINDA KOONTZ, DIRECTOR, INFORMATION MANAGEMENT ISSUES, U.S. GENERAL ACCOUNTING OFFICE

Mr. CARLIN. Thank you, Mr. Chairman. I wish to express to you our real appreciation for you holding this important hearing today on the subject of——

Mr. PUTNAM. If you could just turn the mic on, please.

Mr. CARLIN. That might help.

Mr. PUTNAM. So that we can have it for the archives.

Mr. CARLIN. That is a little embarrassing. We're going to talk about technology today, and I can't turn the mic on.

We're still appreciative of your holding this hearing today. Obviously, your knowledge of technology has been clearly demonstrated, Mr. Chairman, in your brief tenure as the Chair of this committee, and I know your Vice Chair, a former Secretary of State, likewise has great experience, and we appreciate the opportunity to work with both of you.

As you all know, the rapid evolution of information technology has produced huge volumes of diverse and complex digital records. These electronic records pose the biggest challenge ever to keeping records in the Federal Government. When you combine the rate of technological obsolescence with the explosive number of electronic records being created by the government every day, then you can begin to imagine the challenge that we face.

In the National Archives and Presidential libraries, NARA is responsible for preserving and providing sustained access to records of all three branches of the Federal Government, and our challenge is magnified by the need to preserve and deliver authentic records for generations of Americans who will not be born for 100 years or more. The National Archives and Records Administration's statutory responsibilities relating to electronic records management are rooted in the Federal Records Act, which is codified under Title 44 of the United States Code.

Under the statute, the archivist shall "provide guidance and assistance to Federal agencies with respect to ensuring adequate and proper documentation of the policies and transactions of the Federal Government and ensuring proper records disposition."

The archivist also has the responsibilities to approve the disposal of any temporary Federal record and to take into the National Archives of the U.S. Federal records that "have sufficient historical or other value to warrant their continued preservation by the U.S. Government." And these statutory responsibilities apply to electronic records as well as records in other formats.

All this we have summed up in a simple succinct statement of mission which you have already stated. The mission of the National Archives and Records Administration is to ensure for the citizen and the public servant, for the President and the Congress and the courts, ready access to essential evidence.

The scope of NARA's responsibilities compounds the challenge, for the historically valuable electronic records that come to the National Archives and the Presidential libraries come from applications deployed across the entire Federal Government. So they can be virtually in any digital format, from word processing and e-mail, to Web sites, data bases, geographic information systems, digital photography and motion video, computer-assisted engineering or manufacturing applications, laboratory simulations, satellite observations and many others.

Within the past decade, there has been significant progress in developing software products which enable agencies to apply records management discipline to electronic records that are typically produced on individual desktops. An example, the Department of Defense has developed a program for certifying these products as complying with Federal records management requirements. NARA has worked closely with the Defense Department on this program, and has endorsed its use by all Federal agencies.

However, implementing such a product is a time and resource-intensive effort. As a result, agencies are trying to manage many records and paper filing systems despite the fact that some of the new electronic formats cannot be rendered well, or, in some cases, at all in a paper environment.

Through the E-government Electronic Records Management Initiative, NARA works with its agency partners, in providing guidance on electronic records management that is applicable governmentwide, and will enable agencies to transfer their permanent records to NARA in a variety of data types and formats.

In a few moments, Mr. Reynolds Cahoon, the assistant archivist for Human Resources and Information Service, will give you details on our programs and initiatives designed to effectively manage electronic records throughout their life cycle.

In closing, I'd like to thank you again for your interest in electronic records and the challenges they pose for agencies the government as a whole and our Nation. The records of our country have played a vital role in our history, and it is imperative that we find solutions for electronic records. For I'm sure you will all agree with me when I say that records matter for our citizens, our government and the future of our democracy.

Thank you, Mr. Chairman.

Mr. PUTNAM. Thank you very much, Governor.

[The prepared statement of Mr. Carlin follows:]

FORMAL STATEMENT

JOHN W. CARLIN
Archivist of the United States
National Archives and Records Administration

before the

Subcommittee on Technology Policy, Information Policy, Intergovernmental Relations, and the Census
of the
Committee on Government Reform
U. S. House of Representatives

July 8, 2003

I am John W. Carlin, Archivist of the United States, and I wish to thank you Chairman Putnam for holding this important hearing today on the subject of Electronic Records Management. It is a pleasure to be with you, Congresswoman Miller, Congressman Clay and other distinguished members to discuss a subject of obviously great importance to the National Archives, and all of us as Americans. I feel at home here today. Your knowledge of technology has been clearly demonstrated even in your brief tenure as Chairman and I know that as a former Secretary of State, your vice chairman has very specific experience in this area. As for Congressman Clay, in the spirit of full disclosure, I go to work everyday on the campus of his alma mater, the University of Maryland, with which the National Archives enjoys a very good relationship and a very special partnership in the area of electronic records.

Defining the Challenges of Electronic Records

As you all know, the rapid evolution of information technology has produced huge volumes of diverse and complex digital records. These electronic records pose the biggest challenge ever to record keeping in the Federal Government. When you combine the rate of technological obsolescence with the explosive number of electronic records being created by the Government everyday, then you can begin to imagine the challenge that we face.

The very fact of this Committee hearing, and your opening statement, Mr. Chairman, demonstrate that this committee is well aware of both the character and the significance of the challenges posed by increasing reliance on digital technology. But let me highlight the uniqueness and the intensity of the challenge that the National Archives and Records Administration (NARA) faces in this arena.

In the National Archives and the Presidential libraries, NARA is responsible for preserving and providing sustained access to records of all three Branches of the Federal Government. We share the common challenge entailed by rapid changes in the technology. This challenge is two-edged: on the one side, rapid obsolescence makes it difficult to maintain old hardware, software and digital formats; on the other, progress in technology offers better solutions which offer the possibility of improving service to customers. In NARA's case, this challenge is magnified by the need to preserve and deliver authentic records for generations of Americans who will not be born for a hundred years or more.

The ordinary person faces the dilemma of deciding whether to buy a new device today, or wait some months to take advantage of reduced prices and improved performance. So, not only does NARA have to deliver solutions critically needed to accomplish its mission, but we must also ensure that those solutions do not create future difficulties, recognizing that 5, 10, and more years in the future, Americans will want to use the best available technologies to access the records of their government.

The scope of NARA's responsibilities compounds the challenge. Under the provisions of the Clinger-Cohen Act and other statutes, each agency is required to acquire information technology that supports its core mission functions, reduces costs, improves effectiveness, and makes maximum use of commercial, off-the-shelf technology. NARA's challenge is not just to use the best available information technology for accomplishing its core mission, but also to cope successfully with the results of all other agencies' selections of technologies that optimally support their missions.

The historically valuable electronic records that come to the National Archives and the Presidential libraries come from applications deployed across the entire Federal Government. They can be in virtually any digital format -- not only those that result from common end-user applications like word processing and email, but also from web sites, databases, geographic information systems, digital photography and motion video, computer assisted engineering and manufacturing applications, from laboratory simulations, satellite observations, deep space probes, medical information and many others. Combining the variety of types of information that are produced and collected in the conduct of the Government's business with the recognition the digital formats in which the information is recorded change as rapidly as the software used to create and store it, leads to the conclusion that NARA needs to cope with thousands of different digital formats.

Inital Progress to Meet the Challenge

Within the past decade there has been significant progress in developing software products which enable agencies to apply records management discipline to electronic records that are typically produced on individual desktops. The Department of Defense has developed a program for certifying these products as complying with Federal records management requirements. I'm pleased to note that NARA has worked closely with the Defense Department on their Standard 5015.2 and has endorsed its use by all Federal

agencies. However, implementing such a product is a time- and resource-intensive effort. As a result, agencies are trying to manage many records in paper filing systems, despite the fact that some of the new electronic formats cannot be rendered well, or in some cases at all, in a paper environment.

Through the e-Government Electronic Records Management (ERM) Initiative, NARA, working with its agency partners, is providing guidance on electronic records management applicable government-wide and enabling agencies to transfer their permanent electronic records to NARA in a variety of data types and formats.

Achieving Government-wide ERM requires a multi-year effort by all agencies. Differences in agency cultures and information technology infrastructures today make even an agency-wide implementation of ERM a challenging task. Moreover, as previously mentioned, the existing technology solutions for managing electronic records are not yet mature, and do not allow seamless integration into business processes. Clearly, for e-Government to be sustained over time, electronic records must be managed consistently within an agency as well as throughout the Government. We must not look only to solve the technological problems we face, but also to ensure that management, preservation, and reliable access to essential electronic records is built into the fabric of the next generation's national information infrastructure.

NARA's Statutory Responsibilies

The National Archives and Records Administration's statutory responsibilities relating to electronic records management are rooted in the Federal Records Act, which is codified under Title 44 of the United States Code. Under 44 U.S.C. 2904(a) the Archivist of the United States, as head of the National Archives and Records Administration, "shall provide guidance and assistance to Federal agencies with respect to ensuring adequate and proper documentation of the policies and transactions of the Federal Government and ensuring proper records disposition." Subsection 2904(c) specifies ways in which the Archivist carries out these guidance and assistance responsibilities.

The Archivist also has responsibilities under section 3303a to approve the disposal of any temporary Federal record and under section 2107 to take into the National Archives of the United States Federal records that "have sufficient historical or other value to warrant their continued preservation by the United States Government." And these statutory responsibilities apply to electronic records as well as other formats. Section 3301 includes in the definition of federal records "**machine readable materials, or other documentary materials, regardless of physical form or characteristics**."

Title 44 further specifies the kinds of records that Federal officials must create and preserve with NARA's guidance. Section 3101 stipulates that: the head of each Federal agency shall make and preserve records containing adequate and proper documentation of the organization, functions, policies, decisions, procedures, and essential transactions of the agency and designed to furnish the information necessary to protect the legal and financial rights of the Government and of persons directly affected by the agency's activities. Definitions in Section 2901 extend recordkeeping requirements to elements of

the legislative and judicial branches, as well as executive branch agencies, and Section 2203 requires similarly that the President of the United States "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 . . ."

Thus, NARA shares responsibility with Federal officials throughout the government for "adequacy of documentation"--for seeing that certain kinds of records are created, kept, and made accessible. In Title 44 NARA has an additional and unique role to file centrally and to publish Federal laws and administrative regulations, the President's official orders, and the structure, functions, and activities of Federal agencies through the daily *Federal Register*. And in Section 2504 NARA, through the National Historical Publications and Records Commission, has the responsibility for making grants to manage, preserve, edit, and publish the papers of "outstanding citizens of the United States," and "documentary sources significant to the history of the United States." The Information Security Oversight Office (ISOO), which administratively is part of NARA, is responsible to the President for policy oversight of the Government-wide information security classification system and the National Industrial Security Program.

NARA's Unique Mission: Ready Access to Essential Evidence

All this we have summed up in a simple, succinct statement of mission that both reflects our statutory mandates and expresses our sense of their significance:

> ***The mission of the National Archives and Records Administration is to ensure, for the Citizen and the Public Servant, for the President and the Congress and the Courts, ready access to essential evidence.***

This statement acknowledges our statutory responsibility for records in all three branches of the Federal Government. The statement acknowledges our statutory responsibility to help Federal officials manage records effectively for their own use as well as for the use of the public. And the statement acknowledges our commitment to making it as convenient as we can for officials and the public to get access to what sections 3101 and 3301 call "evidence" of "essential transactions" of the Federal Government.

The statutes quoted above speak of protecting rights, of maintaining adequate and proper documentation of what officials responsible to the public do, and of preserving records of historical and other value. Accordingly, we have defined "essential evidence" in our mission statement not just as documents needed for court cases, but as material generated by or received by the Federal Government, that documents: the rights of citizens; the actions of Federal officials; and the national experience.

Documentation of the rights of citizens means material that enables them to establish their identities, protect their rights, and claim their entitlements. Documentation of the actions of Federal officials means material that enables them to explain past decisions, form future policy, and be accountable for consequences. Documentation of the national experience means material of importance for understanding and evaluating the effects of Federal actions.

Essential evidence is not limited in form. It can include: written paper records; maps; drawings, and pictorial materials of documentary value; records generated in multiple formats by computers; artifacts as well as papers in Presidential library collections; and donated manuscripts, Federal Register publications, and other materials that help document rights and entitlements, Federal actions, and historical experience.

We use the term "essential evidence" to sum up, not to supplant, statutory definitions of records or traditional archival concepts. Recognizing that we cannot save everything, nor need to do so, our commitment to essential evidence in our mission statement underscores the particular importance we attach to safeguarding, within the body of Federal and Presidential records, those materials, informational as well as evidentiary, in technical archival terms, that document the identities, rights, and entitlements of citizens, the actions for which Federal officials are accountable, and the effects of those actions on shaping the national experience.

NARA alone is the archives of the Government of the United States, responsible for safeguarding records of all three branches of the Federal Government. The Smithsonian Institution maintains archives of its own as well as artifacts from a wide range of sources, on a wide range of subjects, including American history; and the Library of Congress preserves private manuscript and pictorial collections, as well as books, of an equally wide range. But neither the Library of Congress nor the Smithsonian has a mandate to protect Federal and Presidential records, or exercises responsibility for seeing that the activities of the three branches of the Federal Government are accessibly documented. Historians and other researchers make use of the holdings of NARA, the Smithsonian, and the Library of Congress, but there is virtually no overlap in what we statutorily have the responsibility to collect and preserve. NARA alone is mandated to provide ready access to essential records of what the Federal Government does--why, how, and with what consequences. Our mandate is unique.

Records Matter

Many of the records we hold have survived hundreds of years, but the same cannot be said of electronic records. As we know all too well, records created just a few years ago are already unreadable by today's technology. It is imperative that we find a way to authentically preserve and provide access to electronic records for as long as they are needed. Electronic records, like records in traditional forms, are critical for the effective functioning of a democracy. If they are not managed effectively throughout their lifecycle, vital records of the U.S. Government will be lost.

Thanks to the leadership of Chairman Davis, Congressman Turner and others, the role of the Archives in e-Government was furthered strengthened by our very visible role in the e-Government Act, now Public Law 107-347. We were pleased to be included in the Interagency Committee on Government Information and look forward to working with OMB and the CIO Council on the important issues of cataloging of information and access to electronic information. As the Archivist of the United States, I am required to issue policies based on recommendations from that Committee and the area of standards

that we will talk a great deal about as our hearing goes on, is central to that policy requirement.

Currently we are working to address the challenges of electronic records in a comprehensive and strategic fashion, aiming to both preserving electronic records for the future, while at the same time striving to make effective electronic records management part of the infrastructure of e-government. The capability to create and maintain trustworthy records is an essential component of all government activities. Federal agencies require trustworthy records to meet their legal and internal business needs. Their business partners – whether an individual member of the public, a business, or another governmental entity – rely on the Government to have trustworthy records. Records management is the tool and process for providing such trustworthy records.

In a few moments, Mr. L. Reynolds Cahoon, the Assistant Archivist for Human Resources and Information will give you details on our programs and initiatives designed to effectively manage electronic records throughout their lifecycle.

In closing, I would like thank you all again for your interest in electronic records and the challenges they pose for agencies, the government as a whole, and our nation. The records of our country have played a vital role in our history, and it is imperative that we find solutions for electronic records. For I'm sure you will all agree with me when I say that records matter… for our citizens, our government and the future of our democracy.

Thank You.

Questions for the Record

before the

Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census
Committee on Government Reform
U.S. House of Representatives

July 8, 2003

QUESTION: Dr. Sprehe mentioned in his testimony loss of records at BIA and FBI. Did NARA know about this?

RESPONSE:

BIA

We are assuming that Mr. Sprehe's testimony to the Subcommittee concerning the loss of records at the Bureau of Indian Affairs (BIA) was focused on issues and specific court findings related to the Cobell v. Norton litigation concerning the Federal government's management of Individual Indian Money (IIM) Trust Fund accounts. NARA is certainly aware of the many records management issues involved in the ongoing litigation.

NARA evaluations of the BIA records management program in 1969, 1980, and 1990 indicated serious deficiencies in the management of BIA records. Since 1994, NARA has actively worked with both BIA and the Department of Interior to address critical BIA records management weaknesses, many of which date back 35 years or more. Over the last 10 years, staff in NARA headquarters and regional offices has actively assisted BIA in a wide variety of records management initiatives to help address critical BIA issues. Numerous records management training classes and briefings to BIA managers have been provided at all agency levels, and NARA has provided assistance in inventorying and scheduling BIA records at headquarters, area, and agency offices. NARA has established a national records management team to assist BIA, the Office of Special Trustee, and the Department of Interior in developing a new comprehensive textual records schedule for trust fund records in FY 2004, and to assist with scheduling BIA's electronic systems, especially those related to Indian trust fund management. Several

other BIA record scheduling projects for non-trust related BIA programs and functional areas are planned for FY 2005.

FBI

We are also assuming that Mr. Sprehe's testimony to the Subcommittee concerning the loss of records at the Federal Bureau of Investigation (FBI) is a reference to the Oklahoma City bombing investigation. Contrary to its agreement with Timothy McVeigh's defense attorneys, the FBI failed to turn over all of its Oklahoma City bombing records until shortly before his scheduled execution. As a result, his execution was delayed for one month.

It is important to note that no inquiries or investigations into this matter have determined that records were destroyed without authorization. Rather, because of internal management and communications problems and an atypically broad agreement on the part of the government to give Mr. McVeigh's attorneys everything relating to the case in FBI records, several thousand pages of FBI documents were found not to have been turned over well after the end of the trial and shortly before Mr. McVeigh's scheduled execution. The Justice Department delayed the execution for one month to give Mr. McVeigh's attorneys time to review the records for exculpatory material. No exculpatory material was found in the records. The FBI has undertaken a major reorganization of its records management program to prevent such an occurrence in the future, and is working closely with NARA staff to assure that all FBI records in textual and electronic format are properly scheduled for disposition.

In testimony before the Congress following the revelation of this issue, then Director Louis Freeh committed to a complete renovation of the FBI's records management program. Since then, the FBI has re-established its Records Management Division, which is currently the largest, non-investigative division in the FBI. NARA has worked very closely with the FBI Records Management Division in assessing the management of its investigative records, assisting with records management training, scheduling records and contributing to planning for their conversion to electronic recordkeeping.

QUESTION: How does information from a website get handled?

RESPONSE:

Most Federal agencies use the Internet to assist in carrying out their mission. They may simply disseminate information also available in other forms or conduct business (e.g., e-government initiatives). Agencies must document all of their agency programs (44 U.S.C. 3101) including web sites which are part of its overall public message. Agency web pages meet the definition of a Federal record and therefore must be managed as such. Records relating to managing web sites, tracking use (hits), documenting what was posted and when, and documenting decisions about web content and design, are also Federal records. However, most agencies have not yet scheduled their web records, which include both the content displayed on the web site and the operational and management records associated with maintaining a web site.

NARA has been developing guidance on managing web records and is on record with the General Accounting Office (GAO) for a 9/30/03 public release. The guidance is currently at OMB to begin the clearance process with federal agencies. We expect the interagency comment period will be 4 weeks for federal agencies to comment, since the guidance needs to go to records officers, the CIO Council, and web program staff and we want a unified agency response. Depending on the number and complexity of responses, the September release will proceed as planned.

We plan to issue further guidance on records management needs associated with electronic records, including web records, as part of implementing section 207 of the E-Government Act of 2002.

QUESTION: What is NARA's overlap, if any, with Library of Congress and their digital library?

RESPONSE:

NARA and the Library of Congress have overlapping interests in the development of technically sound and economically viable software products, methods, and standards for digital preservation; however, there is no overlap in operational deployment of the resulting technologies and full

coordination and cooperation has been the rule from the beginning of both agencies' efforts.

The Congress began appropriating research and development dollars to the National Archives and Records Administration for digital preservation in FY 2000, at the inception of an Electronic Records Archives project. From the beginning, findings and successes of that research were shared with the Library and others in the digital library community through a variety of forums including the Federal Library and Information Center Committee, the Federal Depository Library Program, the Defense Technical Information Center User Conferences, the CENDI interagency group, and the Digital Library Federation Forums.

In December 2000, Congress authorized the Library of Congress to develop and execute a congressionally approved plan for a National Digital Information Infrastructure and Preservation Program (NDIIPP). A congressional appropriation was made to establish the program. According to Conference Report (H. Rept. 106-1033), "The overall plan should set forth a strategy for the Library of Congress, in collaboration with other federal and nonfederal entities, to identify a national network of libraries and other organizations with responsibilities for collecting digital materials that will provide access to and maintain those materials. ... In addition to developing this strategy, the plan shall set forth, in concert with the Copyright Office, the policies, protocols and strategies for the long-term preservation of such materials, including the technological infrastructure required at the Library of Congress."

NARA and the Library collaborate on a number of fronts dealing with digital preservation. These include the Library's National Digital Strategy Advisory Board, as well as several working groups of the Library's NDIIPP, the LoC National Audio-Visual Conservation Center, the National Science Foundation's Advanced Computational Infrastructure and Research, Digital Library, and Digital Government research initiatives, the National Institute of Standards and Technology's Information Technology Laboratory, and projects undertaken by the Digital Library Federation.

Operationally there is no disagreement that the mission of the Archives regarding lifecycle management, preservation and access to the Federal record, and the Library's role as both a collector of non-Federal information and its leadership as the National Library has no fundamental overlap. Each

will benefit from the other's research and full cooperation has been the required approach for each institution.

Mr. PUTNAM. Our next witness is Reynolds Cahoon. Mr. Cahoon was appointed Assistant Archivist for Human Resources and Information Services and Chief Information Officer at the National Archives and Records Administration in February 1996. His responsibilities include the Nationwide Information Technology Program, oversight of NARA's Electronic Records Archives Program, human resources staff and organizational development services, NARA's Record Management Program.

Mr. Cahoon currently serves as cochair of the component subcommittee of the Federal CIO Council's Architecture and Integration Committee, and as a commissioner for the International Commission for the Accreditation of Professional Genealogists. Prior to his appointment, Mr. Cahoon served as managing director of the family history department of the Church of Jesus Christ of Latter Day Saints.

In that position he had global operating responsibility for one of the world's largest genealogical research support institutions. He previously served as the director of administration and controller and is the director of projects and planning for the Family History Department. He also served on the board of the Civil War Trust and was a member of the International Council and Archives Committee on archival automation from 1986 to 1997. Welcome to the subcommittee. You're recognized.

Mr. CAHOON. Thank you. I join the archivist in thanking you, Chairman Putnam, for recognizing the importance of electronic records management and for holding this hearing. NARA's plans for electronic records management aim at making it an integral part of electronic government and at delivering electronic records to future generations. Both records management support of current government business and preservation of electronic records for the long term face the technical problem of making digital information assets independent of the specific technologies used to process, store, and communicate them.

Key objectives of the E-government Act of 2002 depend on how well the government creates, retains, and manages records that document its decisions and its performance.

Ultimately, for the trustworthiness of the government, these assets must remain authentic and reliable in the short term as they move between different systems and in the long term when they are retrieved by our great grandchildren and their descendents. To respond to the challenge in a sustainable way, we have launched three major interrelated initiatives:

Our first initiative, redesign of Federal Records Management, builds a foundation that aims to make records management a normal and integral part of agency's asset and risk management processes and add real value to the conduct of government business. We have developed and are working on 14 interconnective strategies to improve records management. Several of the strategies including flexible scheduling and resource allocation are being prototyped by partner agencies. Implementation plans for five more are currently being developed.

In our second initiative, the electronic records management project in the e-government portfolio, we are working with other agencies to provide guidance on electronic records, tools for agen-

cies to manage them, and more format options for agencies to transfer electronic records to NARA. We have produced capital planning guidance on electronic records management application acquisition. We have promulgated and endorsed version 2 of the Department of Defense 5015.2 standard. We have authorized two new transfer technologies, digital linear tape and file transfer protocol and three new data formats for transfer of electronic records to NARA and, finally, initiated automating the process of transferring records to NARA using extensible markup language [XML].

In addition, NARA will lead acquisition of records management components to be included in the Federal enterprise architecture service component infrastructure. Service component software can be reused and leveraged in many different systems enabling agencies to integrate management and use of electronic records within the systems they actually use to transact business.

The electronic records archives must be scalable—excuse me. In our third initiative, the Electronic Records Archive Program will authentically preserve and provide access to any kind of electronic record free from dependency on any specific hardware or software enabling NARA to carry out its mission into the future. The Electronic Records Archive must be scalable and evolvable to accommodate both growth in volume and new types of electronic records and take advantage of improvements in technology. We will implement a flexible approach starting with physical preservation of electronic records in their original formats and, where appropriate, converting them to more durable or more accessible formats.

In the long-term, electronic government and electronic business will drive the emergence of standards, products, and services that make information assets independent of specific technology. We expect agencies will increasingly adopt open-standard infrastructure independent formats, such as XML, which are also suitable for long-term preservation and access.

In the ERA program, we have spent the last 2 years developing the management infrastructure to ensure proper stewardship of this critical program with an eye to its contribution to both electronic government and to posterity. We are completing requirements development and refining our acquisition strategy with the target of issuing a request for proposals by the end of this calendar year. In this process we have engaged both our customers and the IT industry in an extensive dialog.

Facing the challenge of electronic records is a difficult, serious endeavor, but we have no alternative. To this committee, I respectfully testify that we, as the National Archives, face the challenge squarely and with full purpose and resolve to leave for our descendents a trustworthy record of our turn on this planet. Thank you, Mr. Chairman. I'd be happy to answer any questions you might have or your subcommittee might have.

Mr. PUTNAM. Thank you very much.

[The prepared statement of Mr. Cahoon follows:]

FORMAL STATEMENT

L. REYNOLDS CAHOON
Assistant Archivist for Human Resources and Information Services,
and Chief Information Officer
National Archives and Records Administration

before the

Subcommittee on Technology Policy, Information Policy,
Intergovernmental Relations, and the Census
of the
Committee on Government Reform
U. S. House of Representatives

July 8, 2003

I am L. Reynolds Cahoon, Assistant Archivist for Human Resources and Information Services, and Chief Information Officer of the National Archives and Records Administration (NARA). I wish to join the Archivist in thanking you, Chairman Putnam, for recognizing the importance of Electronic Records Management and holding this hearing.

Two Perspectives on Records Management

As the Chief Information Officer of NARA, I have to maintain a dual perspective on the needs and responsibilities of this agency. In one direction, I have to see to it that NARA has the information technology that enables it to accomplish its mission and serve the citizens. In the other direction, I must take into account NARA's government-wide responsibility for leadership in the lifecycle management of records, which are critical information assets needed to make the U. S. Government work, to ensure that it is accountable to the citizens, and to document our national experience.

These two perspectives must coalesce in a comprehensive vision encompassing both how NARA accomplishes its own business and the direction NARA gives to other agencies. Only concerted, complementary efforts in both directions will enable the Federal Government to address the challenges posed by electronic records and make the transition to e-Government in a way that takes full advantage of the opportunities offered by networking and information technology while preserving what is best in our democratic traditions.

While my responsibilities as CIO focus on issues and topics related to information technology, there is not a shadow of a doubt that the goal of my efforts, and those of my staff, is not to acquire or deploy technology, but to help realize the contribution that effective records management can make in the transition to e-Government, in interagency

and intergovernmental collaborations, and in the value that American citizens can realize in the information assets of their Government.

Legal Mandates for Records Management

The Archivist has described NARA's need to preserve and provide sustained access to key records of the U.S. Government in the National Archives and the Presidential libraries. Equally daunting is NARA's responsibility for guiding other agencies' management of the electronic records they create. This guidance must provide a sound and comprehensive basis for achieving the objectives of records management as set out in the Federal Records Act. Those objectives do not make records management a goal in its own right, and they do not divorce the preservation of historically valuable records from the on-going performance of the Government's business. The objectives of this law include:

- Accurate and complete documentation of the policies and transactions of the Federal Government;
- Effective and economical government operations;
- Simplification of activities, systems, and processes; and
- Prevention of unnecessary Federal paperwork; as well as
- judicious preservation and disposal of records. (44 U.S.C. 2902)

While these objectives were set out in the Federal Records Management Amendments of 1976 (H.R. 13828), they resonate well with key purposes of the E-Government Act of 2002. Among other worthy goals, the E-Government Act aims

- To improve the ability of the Government to achieve agency missions and program performance goals.
- To reduce costs and burdens for businesses and other Government entities.
- To promote better informed decisionmaking by policy makers.
- To make the Federal Government more transparent and accountable.
- To transform agency operations by utilizing, where appropriate, best practices from public and private sector organizations; and
- To provide enhanced access to Government information and services in a manner consistent with laws regarding protection of personal privacy, national security, records retention, access for persons with disabilities, and other relevant laws.

Government performance, decision-making, and accountability, as well as access to Government information all depend on how well the Government creates, retains, and manages the records that document its decisions and performance; in other words, how well agencies achieve the objectives of Federal records management.

NARA's responsibility for government-wide direction in records management makes it imperative for us to ensure that our guidance on managing electronic records is appropriate and effective. This is a strategic requirement which we address with as much seriousness and commitment as we apply to the task of preserving key electronic records of the U.S. Government for the citizens. In addressing this strategic requirement, we

must remember that what makes something a federal record has nothing to do with its form – whether it's a report or a web page, a letter or an email – and nothing to do with the information technology used to create, communicate or store it. What makes it a government record is either its connection to the conduct of government business or the value of the information it contains.

The often intricate connection between the technology and electronic records and the potential for improved service to the citizen and better return on investment offered by appropriate selection and use of information technology, compels us to adapt and enrich our records management guidance to make it highly useful to individual agencies as they implement IT solutions, and to align it with the government-wide objectives of the E-Government Act, the Government Performance and Results Act, and other key statutes.

Responding to the challenges of electronic records

Our response is comprehensive, multi-faceted, and future oriented. 'Like the dual vision I mentioned earlier, this orientation aims both at delivering electronic records to future generations and at positioning electronic records management as an integral element in effective e-Government. It is essential to tackle both of these objectives in parallel. The solutions NARA adopts for preservation and sustained access must complement the solutions other agencies implement to improve the conduct of the Government's current business.

NARA's actions in the short-term to improve the lifecycle management of electronic records, from the moment of creation, are critical for its long-term goal of ensuring that the history of our time is accessible to future generations through authentic records. Without neglecting the importance of transmitting those electronic records that already exist to future generations, continuing expansion of the application of IT in government, combined with continuing expansion in the kinds and the complexity of information recorded in digital form mean, that we have to anticipate that the challenges entailed by electronic records will explode to a level that makes everything we have faced up to now look small by comparison.

NARA looks for a coherent solution to the challenge of electronic records because both electronic records management to support current business and archival preservation of electronic records for the long-term face the same, fundamental technical problem: the need to make digital information assets independent of the specific hardware and software used to process, store, or communicate them at any time. In the archival context, this problem manifests itself over time in the twin problems of obsolescence and progress. In the context of e-Government, this problem manifests itself in the need for systems to interoperate across agency lines, and more broadly for agencies to interact electronically with citizens, with businesses, and with state, local, and tribal governments, and with other nations.

For e-Government to achieve the goals which the Congress and the President have established in the areas of efficiency and timeliness, it is necessary that the systems involved in a transaction be able to interoperate and automatically exchange required

information. Both far-reaching and fundamental differences in requirements and simple differences in the frequency and timing of when institutions and individuals acquire and replace technology make it extremely unlikely that all the stakeholders with whom the U.S. Government needs to interact will ever have fully compatible systems. Moreover, developments in information technology make that unnecessary. For example, an online transaction in e-government might involve, on one end, an agency application running on a sophisticated database management system and, on the other either a company using a system that is fine-tuned to its business, or a citizen using a personal computer.

Rather than requiring that all parties use the same, or compatible, systems, the technology is providing generic products that translate among disparate systems. Notably, many commercial products implement the eXtensible Markup Language (XML). The XML standard provides a simple, open way to identify what kinds of data are being exchanged online. For different systems to participate in electronic transactions, each one only needs to recognize the XML markup of the data it receives, and translates the data into its own internal format. Conversely, it needs to be able to export its data in the shared format. As predicted by industry analysts 5 years ago, this mediated approach is proving to be one of the major enablers of e-business and e-government.

The Need for Authenticity

But for quality of service, and ultimately for the trustworthiness of the Government, information assets must not suffer any diminution in authenticity or reliability as they go through repeated transformation from one internal format, to an intermediate common format, and to another internal form in the course of a transaction. Effective records management is in fact the discipline which can ensure that authenticity and reliability remain intact in transmission of information across time, space, and institutional and technological boundaries, and it can also contribute to timeliness and efficiency.

NARA is uniquely positioned and legally required to ensure that federal records management does make this contribution to e-government. Of itself, records management cannot satisfy the requirements for authenticity and reliability, but it provides a well-established framework for defining the requirements, for evaluating the solutions that technology offers, and for ensuring that the real-world results achieve their objectives.

In broad terms, we regard the mediated approach that technologists have developed to solve the problems of sharing information in real time as the most viable approach for sustained access to authentic electronic records over the long-term. No one can pretend to know what information technology will be like 25, 50, or more years in the future. But it is a reasonable assumption that Americans will want to use the best technologies available to them to access the electronic records of their Government, regardless of when those records were created. In effect, the solution NARA implements to provide sustained access to these records for the long-term must mediate between the myriad digital formats of federal electronic records that exist now, and that will be created in years to come, and the technologies that best serve the needs of citizens at any time in the future.

The most basic requirement for the architecture and the technology that NARA adopts to accomplish its core mission is that the system be able to evolve over time, keeping pace with progress in IT, and responding to citizens' expectations for the best available service. Building this solution will not be easy. Based on market research conducted over the past two years, an extensive dialogue with the IT industry, and continuing collaborations with leading researchers in computer science and engineering, we have learned that key technologies enabling such a solution are available today, and that it will be possible to develop a full solution gradually over time.

NARA's Initiatives on Electronic Records Management

NARA's comprehensive approach involves three major initiatives: the Records Management Initiatives, the Electronic Records Management project in the Administration's e-Government portfolio, and the Electronic Records Archives Program, as well as a number of other activities. It starts close to home with an open, everything-on-the-table review of our records management guidance.

Under the leadership of the Deputy Archivist, Dr. Lewis Bellardo, our Records Management Initiatives aim to create mutually supporting relationships with agencies, whereby

- records management is viewed as a normal part of asset and risk management;
- records are managed effectively for as long as they are needed;
- NARA's records management program adds value to the agency business processes; and
- records of archival value are preserved and made available for future generations.

We will adapt our guidance to the digital environment; make it adaptable by agencies to suit the specific needs entailed by their missions and the different constituencies they serve; and we are supplementing our government-wide guidance with targeted assistance to individual agencies.

Government-wide Records Collaborations

We are coordinating our efforts with other government-wide initiatives in furtherance of the President's Management Agenda, especially his e-Government goals, notably with the Office of Management and Budget (OMB) and the Federal CIO Council. Our work with the Federal Enterprise Architecture (FEA) Program Management Office in OMB is reflected in the position of records management in the June 2003 release of the Business Reference Model (BRM) of the FEA.

Lifecycle management of records is positioned in all four of the main areas of business defined in the BRM. In the Services for Citizens area, it appears under the heading of "Cultural and Historic Preservation." In the Mode of Delivery area, it is part of "Information Infrastructure Management." The Support Delivery of Services area includes "Central Records and Statistics Management" under the General Government

line of business. And the Management of Government Resources area includes Records Retention as a sub-function of Information and Technology Management. Thus the Business Reference Model simultaneously promotes records management and positions records management to contribute to e-Government.

Our coordination is not only at the intellectual level of elaborating the framework of the Federal Enterprise Architecture. It also has practical elements. Working with other agencies, NARA has undertaken four Electronic Records Management projects in the Administration's e-Government portfolio, providing specific examples of the contributions of good records management to e-Government.

The Electronic Records Management Initiative will provide the tools that agencies will need to manage their records in electronic form, addressing specific areas of electronic records management where agencies are having major difficulties. This project will provide guidance on electronic records management applicable government-wide and will enable agencies to transfer electronic records to NARA in a variety of data types and formats so that they may be preserved for future use by the government and citizens. Ultimately, to be successful, the management of electronic records must be an integral part not only of the FEA, but also of the information architecture and infrastructure of each agency.

The FEA includes a Technical Reference Model (TRM) which provides a foundation to support the construction, delivery, and exchange of business and application components, called 'Service Components', that may be used and leveraged in many different applications. Components can be large or small, written in different development environments, and may be platform independent. Components can be executed on stand-alone computers, or a LAN, Intranet, or the Internet. The FEA Service Component Framework provides for the definition, development and deployment of common service components which will be available for use and reuse of applications running in the FEA.

I have the privilege of serving as the co-chair of the Components Subcommittee of the Architecture and Infrastructure Committee, chartered by the CIO Council. The subcommittee's objective is "to foster the identification, maturation, use and reuse of Component-Based Architectures and Architectural Components in the federal government." In this context, NARA will lead the acquisition of records management components which will be included in the FEA's Service Component infrastructure and available to any application. This will be a significant improvement over the current situation where records management is implemented as a stand-alone application which is not integrated with the applications which are used to conduct business.

Electronic records are, in fact, created in such business applications. Agencies will be better equipped to optimize the use of these information assets in e-Government if they can be retained and managed within the systems they actually use to transact business. Records management component services will make this possible, by identifying electronic records wherever they exist, by enabling both government officials and agency

customers to navigate among records, and by ensuring that the integrity of the records is protected and that they are properly retained and destroyed.

Electronic Records Standards
We are also strengthening records management through support for the development of relevant standards and alignment with the emerging national information infrastructure.

First, standards. We have worked with NASA and representatives of other American and international organizations since 1995 in the development of the ISO standard for Open Archival Information Systems. NARA has also contributed to the development of the ISO standard on records management.

We are also working with the Association for Image and Information Management, International, to establish integrated EDMS/ERMS functional requirements for electronic document management system (EDMS) and electronic records management systems (ERMS), and on a proposed standard for the use of the Adobe Portable Document Format (PDF), for the long term storage of multi-page documents that may contain a mixture of text, raster images and vector graphics.

NARA has made substantial contributions to the development and the success of the DoD standard for Records Management Applications (DoD 5015.2-STD), the de facto standard for records management software, adopted by private companies, as well as by other governments, such as the State of Michigan. Currently there are over 40 commercial off-the-shelf software products certified as compliant with the Department of Defense standard. And NARA is supporting DoD in its efforts to update and enhance this standard.

NARA is also one of the original members of the Federal Geographic Data Committee (FGDC). We have contributed to the development of the FGDC's policies on data management and on historical data, as well as its data transfer and content standards. We are proud to be part of the nationwide effort to realize the vision of the National Spatial Data Infrastructure, an effort actively embraced by all levels of government in this country.

National Information Infrastructure
Second, development of the next generation national information infrastructure. The challenges NARA faces in the realm of electronic records are immense and complex, and rendered more so by the prospect of continuing, open-ended change in information technology. But developments in the technology also offer improving prospects for viable solutions. Rather than seeking to surmount these challenges alone, or expecting the market to develop *specific* solutions for all the problems we face, we seek solutions in mainstream developments that promise to become major elements in the emerging national information infrastructure needed to support not only e-Government and e-business, but also education and research in a networked world.

Rather than looking for *specific* archival and records management solutions, we have sought to identify mainstream developments that, with modest redirection, might be adapted to *specific* archival and records management requirements. We have pursued this approach since 1998. In doing so, we have benefited from the prior work of agencies such as NASA, the Defense Advanced Research Projects Agency, the National Science Foundations, the National Institute of Standards and Technology, and the Army Research Laboratory.

One of the major assets we have is the very nature of the challenge we face: its great complexity and the immense size. NARA's electronic records challenge has attracted the attention of world-class researchers in institutions around the country, such as the San Diego Supercomputer Center, the University of California /Berkeley, the Georgia Tech Research Institute, the University of Illinois at Urbana-Champagne, Ohio State University, and the University of Maryland's Institute for Advanced Computer Systems.

The importance of the dynamic challenge NARA faces has been recognized at the highest levels of government. Every year the White House issues what is known as the 'Blue Book.' This is an official supplement to the President's Budget which explains how the Executive Branch coordinates the Government's investment in networking and information technology research and development. In the Blue Books for the last two Fiscal Years, 2002 and 2003, the White House has identified digital preservation in general, and the appraisal and management of electronic records in particular, as significant areas of research. The Administration's recognition of the need for research in this area is critical because, as long as the technology continues to develop, and as agencies find new ways to apply it in conducting the public's business, NARA will have to continue to work with leaders in research and development to find new solutions to new problems.

But we cannot delay action. NARA needs to move ahead aggressively to acquire and implement solutions that will enable us to ensure that important records of our era are not lost. Records such as:

- the State Department's worldwide diplomatic correspondence, which was converted to digital form starting in 1972;
- the digital returns from the 2000 Census of Population;
- the digital map of the United States;
- operational and intelligence records of the war in Iraq; and,
- the automation of Congress's legislative process and the planned coverage of its proceedings in High Definition Television.

The Electronic Records Archives Program

The insights we have gained and the lessons we have learned to date from our research collaborations have emboldened us to launch a frontal assault on the technological challenges entailed in preserving and providing sustained access to valuable electronic records in the National Archives and in the Presidential libraries. This strategic initiative is our Electronic Records Archives (ERA) Program. This program includes three

elements. It is responsible for the research partnerships I have just mentioned. It is managing the effort to acquire a system which will enable us to manage the lifecycle of electronic records and to preserve them for American citizens, and it supports business process redesign and change management activities to ensure that NARA and its staff are able to use the ERA system successfully.

NARA's vision for the ERA system is that it "will authentically preserve and provide access to any kind of electronic record, free from dependency on any specific hardware or software, enabling NARA to carry out its mission into the future." The key features of the Electronic Records Archives will be (1) an architecture which is scalable to expected growth in the volume of electronic records and evolvable both to accommodate new types of electronic records and to take advantage of improvements in technology, and (2) a system with the ability to accept, preserve, and provide sustained access to any kind of electronic record for as long as needed.

The information and expert advice we have received on the state of the art supports the conclusion that it is possible to develop such an architecture now. With respect to preserving and providing sustained access to electronic records, the proven methods that are currently available are limited to relatively few formats. But we are optimistic that we will be able, over the course of time, to expand the varieties of electronic records that we can preserve indefinitely. There are two elements to this optimism. One is that the requirements of e-Government and of e-business are driving the emergence of standards, products, and services that make information assets independent of any specific hardware or software. The other is that these same requirements will lead agencies increasingly to adopt open-standard, infrastructure-independent formats, such as XML.

In the interim, we will implement a flexible approach to preservation and access. We will start with guaranteed physical preservation of all electronic records in their native formats. While it is desirable to convert the records to formats which would be both infrastructure independent and immune to obsolescence over substantial lengths of time, until the market for products supporting such formats matures, we will adopt methods for replicating these records in formats that are either less sensitive to obsolescence or better for meeting short term demands for access. A number of open formats that are available now include the Portable Document Format for textual information, the Tagged Image File Format for scanned paper, and the relational or other standard database formats not only for data files, but also for semi-structured records, such as e-mail.

Our judgement that it is possible to start building the Electronic Records Archives is supported by market research and by our dialogue with the IT industry. Over the past three months, we have had representatives of 72 companies come to NARA, to tell us about their products and services, which they believe could contribute to the solution. We have had extensive discussions with a dozen major systems integrators who are candidates to lead the development of the Electronic Records Archives system.

What we heard from these companies is, in sum, that it is possible to build a digital archives, which not only can be sustained over the long-term, but which can grow and

evolve in response to the dynamic challenge NARA faces. The need for the Electronic Records Archives system, the possibility of developing it, and the feasibility of an evolutionary development solution has also been affirmed recently in a study we have commissioned from the Computer Science and Technology Board of the National Academy of Science.

The Electronic Records Archives Program focuses on the archival end of the records lifecycle, on delivering historically valuable electronic records to Americans in the future. But it is not ignoring the front of the lifecycle in doing so. Just as, in generic terms, we are seeking archival solutions in mainstream developments in the National Information Infrastructure, we are specifically seeking to build an Electronic Records Archives that is optimally integrated with approaches and technologies used to manage records in active support of the Government's current business.

Partnering to Find Solutions

I would like to close with a final point. Throughout my remarks, I have mentioned many of NARA's partners across the Government and the Nation. NARA's Strategic Plan repeatedly recognizes the need for us to partner with others in addressing the challenge of electronic records. It not only recognizes the need, but also commits us to working in collaboration. We have wholeheartedly embraced this collaborative strategy in the Electronic Records Archives Program. The partners I have named do not by any means exhaust the list, and merely naming them does not do justice to the benefits NARA has derived from working with them.

Our research partners have opened our eyes to possibilities we had not conceived; they have expanded our horizons and strengthened our eagerness to move ahead. The managing partners in our research activities – DARPA, the National Science Foundation, and the Army Research Lab – have opened the doors to a wealth of knowledge, expertise and creativity while enabling us to avoid the need to develop within NARA the full administrative overhead needed to manage high technology research.

We have also been enriched by peer collaborations with other agencies and institutions that also need to preserve digital information for long periods of time. Our earliest partners in this area were NASA, the Patent and Trademark Office, and the InterPARES project. The InterPARES project focuses squarely on the preservation of authentic electronic records. Headquartered at the University of British Columbia, the project brings together archivists, records managers, information scientists, computer scientists and engineers, chemists, conservators, museum directors, and even artists from around the world.

We have also developed profitable partnerships with the library community. We are pleased to be collaborating with the Library of Congress in its initiative to preserve America's cultural and intellectual heritage in digital formats. We also profit from interactions with librarians and technologists from leading universities as a member of the Digital Library Federation.

We are also grateful to the many companies that have not only informed us about their products and capabilities, but have also shared with us the benefits of their experience and insights into the risks, the complexities, and the best practices for developing large, complex IT systems. We look forward to working even more closely with industry in the development of the ERA system in the near future.

Finally, I need to mention our relationship with the General Accounting Office. Two year ago, the Congress asked GAO to review the ERA Program, and they have been with us ever since. We are benefiting from GAO's understanding of systems development, gained in decades of experience across the Government. It is an enviable situation to have GAO going along with us in each step of the program, sharing its insight and recommendations so that we can avoid problems before they happen.

The Archives of the Future

One sign of the vitality and success of our commitment to approaching the challenge of electronic records through extensive partnerships is in the fact that, in 2001, for the first time in history, the Archivist of the United States was invited to give the keynote address at a Symposium of the Institute of Electrical and Electronics Engineers. I would like to close my testimony with a quote from Mr. Carlin's speech to the IEEE:

> "The 'Archives of the Future' will not consist of many buildings scattered across the country. Instead, the 'Archives of the Future' will be available on the desktop of any American who chooses to explore the records of his or her country.... Building this new, digital archives is not and will not be easy. But we have no alternative....
>
> "A society whose records are closed cannot be open. A people who cannot document their rights, cannot exercise them. A nation without access to its history cannot analyze itself. And, a government whose records are lost cannot accountably govern."

Building the Electronic Records Archives is a difficult, serious endeavor, one which NARA could not achieve without the support contributions of our many partners, not the least of which is the Committee on Technology, Information Policy, Intergovernmental Relations, and the Census. I thank you for the opportunity you have given us to explain our strategy for ensuring that electronic records management becomes a critical and successful element in the infrastructure of e-Government, and that Americans will continue to have access to valuable records.

Mr. Chairman, I would be happy to answer any questions that you or your subcommittee might have.

Mr. PUTNAM. Our next witness is Ms. Harriet Riofrio. Ms. Riofrio serves as electronic records management policy lead and senior staff officer for knowledge management for the Office of Assistant Secretary of Defense for Networks and Information Integration. You must have the longest business card of anybody in government. You have to flip it over to read what your address is.

In this capacity her primary role is to assist the deputy CIO and director IM to enable netcentric process change throughout the Department of Defense. Duties include strategic planning for and application of the principles and techniques of ERM and KM. She is responsible for the management and evolution of the Department of Defense 5015.2 STD—only the Department of Defense can come up with this stuff—and the Records Management Application Certification Program at the Joint Interoperability Test Center. She represents DOD and the DOD-NARA partnership that maintains as its objective the development of DOD standard to improve records management processes and electronic transfer of records between the two agencies.

She also originated and leads the DOD-KM community of practice. We did our best to include every conceivable acronym in your biography. Previously, Ms. Riofrio headed the Executive and Information System and Microtechnology Group at HQ Defense Logistics Agency, where she was responsible for designing and implementing innovative technology solutions to include major agency executive information and decision support systems.

She has numerous speaking engagements, awards and publications to her credit. She was born here in Washington, DC, and received a Masters Degree in 1980. Ms. Riofrio is known as an innovator and thought leader in business process design, new technology insertion and management disciplines. She joined the OSD staff in 1997. Welcome to the subcommittee. Hopefully your testimony will be more clear than your biography.

Ms. RIOFRIO. It took coming here to realize how many acronyms we use. I apologize for that.

Mr. Chairman and members of the subcommittee, I am pleased to appear before the subcommittee this afternoon to describe the evolution of the Department of Defense Electronic Records Management Program and standard, culminating in its current relevance to and integration with our plans for advancing the Department to a netcentric government environment. Netcentricity will provide an unprecedented accessibility and usability of data to include departmental records in every electronic media. Through the development of our Global Information Grid Enterprise Services Initiative, we will be converting our information infrastructure from a platform based to a netcentric environment.

We will be applying a metatagging standard to DOD data in general and facilitating its just-in-time discovery. Official records are a critical part of our enterprise knowledge base. The intent is to make them more visible and usable while utilizing sound content management principles to ensure their proper storage, preservation and protection.

I will describe how our approach to electronic records management has been developed and applied to date and how it represents

one of several foundational disciplines converging to achieve our vision of enterprise wide netcentricity.

Looking back, it seems we have been working for the last 10 years to ready our records management processes for netcentricity. In 1993, records management was analyzed as part of a functional process improvement initiative. A baseline analysis of records management in the Department was developed. This work culminated in the first version of the DOD 5015.2 standard that was signed in 1997. The focus of the standard has been to prescribe essential records automation functions consistent with the law and regulations. The intent has been to help DOD components bid, test or buy compliant records management application software. Its objective has been to be unambiguous and not to dictate design.

In 1998, we began the process of developing the second revision to the standard. We added a section on national security classification markings, and the declassification schedule as requested by the Intel community. We also included recommendations from NARA. We added a requirement for information related to section 508 of the Federal Rehabilitation Act. Version 2 of DOD 5015.2 standard was signed in June 2002.

Currently, the Joint Interoperability Test Command [JITC], of the Defense Information Systems Agency, manages the compliance testing process for the DOD standard. This testing is mandatory for DOD and endorsed by NARA. At this time, JITC has certified approximately 43 records management applications as compliant to our standard.

An example of how this standards process has begun to influence our enterprise-wide environment can be seen in its application to the Navy Marine Corps Intranet. The Navy and Marine Corps are looking to merge records and document management for over 350,000 users which they claim will make the Department of the Navy the world's largest electronic records management customer.

In 2003, DOD formally partnered with NARA on electronic records management as part of the President's E-government initiatives. We are working to certify direct export from RMA's from record management applications to NARA to include coordination of transfer and specification for transfer file format. This effort, especially upon its adoption by the Federal community, has the potential for substantial improvements in the timeliness, cost and quality of permanent records transfers throughout the government.

As we begin to operationalize netcentricity through our Global Information Grid Enterprise Service Initiative, we are cognizant of the need to assure electronic records management becomes a part of the Department's enterprise services.

In summary, we are finding that the DOD standard could not be more opportune. It is being sought after and used by some of our States, by the Federal Government and even the international community. It is open enough to encompass different functions and infrastructures but serves to assure consistently useful electronic records products. We look forward to planning for integration of ERM into our new netcentric environment and working closely with the Federal community to develop solutions for our common data and records problems.

We welcome your support as we continue to shape the future of this critical area together. Thank you.

Mr. PUTNAM. Thank you very much.

[The prepared statement of Ms. Riofrio follows:]

STATEMENT FOR THE RECORD

HARRIET J. RIOFRIO

DEPARTMENT OF DEFENSE ELECTRONIC RECORDS

MANAGEMENT POLICY

BEFORE

THE HOUSE GOVERNMENT REFORM COMMITTEE

INFORMATION AND TECHNOLOGY SUBCOMMITTEE

Mr. Chairman and Members of the Subcommittee:

I am pleased to appear before the subcommittee this afternoon to describe the evolution of the Department's electronic records management (eRM) program and standard, culminating in its current relevance to and integration with our plans for advancing the Department to a netcentric environment. Netcentricity will provide an unprecedented accessibility and usability of data to include Departmental records in every electronic media. Through the development of our Global Information Grid Enterprise Services (GES), we will be converting our information infrastructure from a platform-based to a netcentric environment. We will be applying a metatagging standard to DoD data and facilitating its just-in-time discovery. Official records are a critical part of our enterprise knowledgebase. The intent is to make them more visible and usable while utilizing sound content management principles to ensure their proper storage, preservation and protection. I will describe how our approach to electronic

records management has been developed and applied to date and how it represents one of several foundational disciplines converging to achieve our vision of enterprise-wide netcentricity.

Looking back, it seems we have been working for the last ten years to ready our records management processes for netcentricity. In 1993, Records Management was analyzed as part of a functional process improvement initiative.
A baseline analysis of records management in the Defense Department was developed. In 1995, an automated document conversion master plan was developed. We accomplished the groundwork for establishing a certification process and published "Baseline Requirements And Data Elements For DOD Records Management Application Software". A Records Management Software Test and Evaluation Program Plan was developed. This work culminated in the first version of the DoD 5015.2-STD that was signed in 1997.

The focus of the standard has been to prescribe essential records automation functions consistent with the law and regulations. The intent has been to help DoD components build, test or buy compliant records management application software. Its objective has been to be unambiguous and not to dictate design.

The eRM standard's requirements for the Records Management Applications (RMAs) have been to assure basic capabilities to:

- Define a file plan - record categories/series and their associated disposition schedules
- Identify/declare records, provide context
- Store, preserve, protect electronic records
- Search for and retrieve electronic records
- Track records' disposition schedule status
- Execute disposition instructions--cutoff, transfer, destroy

In 1998 we began the process of developing the second revision to the standard. We added a section on national security classification markings and the declassification schedule as requested by the Intelligence Community. We also included recommendations from the National Archives and Records Administration (NARA). We added a requirement for information related to Section 508 of the Federal Rehabilitation Act. We added a section on email that allows editing of sender and addressee metadata fields to comply with NARA's regulation requiring that email records include the real names of senders and addressees. Version 2 of the standard explicitly requires the RMAs to have the ability to export email from the records repository to an email application to allow forwarding, and replying.. We enhanced the auditing section and added requirements on vital records. Additionally, we added requirements to enter the security classification level and allow the implementation of classification guides. Version 2 of the DoD 5015.2-STD was signed June 2002.

Currently, the Joint Interoperabilty Test Command (JITC) of the Defense Information Systems Agency (DISA) manages the compliance testing process for the DOD eRM standard. The Compliance Program verifies that the RMA software satisfies mandatory requirements. This testing is mandatory for DoD and endorsed by NARA. At this time, JITC has certified approximately 43 records management applications as compliant to DoD 5015.2-STD.

There are numerous compliant electronic records implementations in DoD. Here are some examples:

- The Army has implemented the Army Records Information Management System (ARIMS). This "Recordkeeping Redesign" applies to every function and program in the Army. It places emphasis on important records, provides secure management of electronic records and simplifies the process for soldiers and action officers to identify and preserve important records in any medium. ARIMS is a web-based set of applications and tools to help in the management of the Army's hardcopy and electronic records. Key ARIMS components include:
    - An Army Electronic Archives to preserve electronic records in any format
    - A central index to all records, paper and electronic
    - Web based tools and features to help users find, collect and preserve record information

  ARIMS is available for use via the World Wide Web without cost to the

user and provides a searchable web-based database of all retention and disposition information on Army records.

- The Defense Logistics Agency has implemented a compliant RMA for their Financial Services Office. They have digitized five years of records and thus provided a speed and flexibility of function unavailable previously when the staff was required to manually and laboriously search through large volumes of paper records.

- Through their Enterprise-wide Navy Marine Corps Internet (NMCI), the Navy and the Marine Corps are looking to merge records and document management for over 350,000 users.

"Navy is equipped to provide a robust foundation to support the enterprise-wide requirement for ERM and EDM solutions. The Department of the Navy is excited to become the world's largest electronic records management customer... Although electronic records management (ERM) tools are not widely used within the Navy and Marine Corps at this time, the implementation of the Navy Marine Corps Intranet (NMCI) is poised to bring about a fundamental transformation...The Department of the Navy (DON) required the winner of the NMCI contract to provide for the 'retention of electronic information files consistent with applicable DoD (DoD Standard 5015.2-STD) and DON policy (SECNAVINST 5212.5D),'

as part of the basic service for all NMCI seats'...."

http://www.chips.navy.mil/archives/02_Summer/authors/index2_files/DONERM.htm

In 2003, DoD partnered with NARA on Electronic Records Management as part of the President's E-Government Initiatives. Joint research has begun regarding transfer of permanent electronic records to NARA with ultimate inclusion in the next version of the DoD standard. This next version is anticipated to include additional metadata for special permanent record classes to include digital photographic images, scanned images of hard copy comments; web pages; email attachments and "portable document format" or "PDF". NARA is partnering with DoD to certify direct export from RMA to NARA to include coordination of transfer and specification for transfer file format. This effort, especially upon its adoption by the Federal community, has the potential for substantial improvements in the timeliness, cost and quality of permanent records transfers throughout the Government.

Additionally, and in preparation for netcentricity collaboration services, we are looking at the implications for eRM. It has been suggested that different records classes could have different metadata sets. Structure of metadata fields and export file formats will have to be specified in the next version of the standard.

Adoption of DOD 5015.2 STD version 2.0 for civilian agency use was specified in NARA Bulletin issued 1/15/03. It endorses the DoD standard for all agencies and recommends using the DoD standard and the DoD –certified products as a

baseline when selecting an RMA to manage the agency's electronic records:

> *"We recommend that agencies use the DoD standard and the DoD-certified products as a baseline when selecting an RMA to manage the Agency's electronic records. NARA has evaluated the DoD standard for record keeping functionality and has determined that DoD-certified products comply with the relevant provisions of the Federal Records Act and NARA regulations with respect to the creation, maintenance and use, and disposition of Federal records. Continued partnership between NARA and DoD to develop future enhancements to the DoD standard will enable DoD 5015.2-STD certified RMAs to export electronic records and directly transfer them to NARA. "(NARA Bulletin 2003-03,* January 15, 2003)

As we begin to operationalize netcentricity through our GES initiative, we are cognizant of the need to assure electronics record management becomes a part of the Department's enterprise services. Netcentric principles such as the following will be particularly transformational for records management:

- "Only handle information once." (OHIO) Collecting information or replicating data entry is costly and adversely affects efficiency in both combat and business operations. "Only handling information once" requires that processes be re-engineered, and that technology and processes are integrated to minimize time and effort dedicated to data collection and entry.

- "Post before processing" means that access to data for disparate needs is not delayed by unnecessary processing. Everyone is a provider and consumer of information. As a provider, they have the responsibility to post data before they use it; as a consumer they will have the technical capability to securely access all data when they want it and in the format they need.

- Users will "pull" data as needed instead of having massive amounts of information "pushed" to them regularly, regardless of whether it is needed.

- Collaboration technologies will be employed to assist users in making sense of the data that is pulled. For example, subject matter experts from diverse units or organizations are frequently called upon to come together to make sense out of special situations. The ability to pull expertise from within a unit as well as from across the Department is a value-added feature of a net-centric environment.

These netcentric principles and their accompanying infrastructure and capabilities have the potential to solve some of the Department's most stubborn records management problems.

The DoD 5015.2-STD could not be more opportune. It is being sought after and

used by the Federal and even the international community. It is open enough to encompass different functions and infrastructures but serves to assure consistently useful electronic records products. We look forward to working closely with the Federal community to develop solutions for our common data and records problems. We welcome the support of this committee and look forward to a continuing dialogue about this critical area. Thank you.

Mr. PUTNAM. Our final panelist for panel one is Ms. Linda Koontz. Ms. Koontz is Director of Information Management Issues at the U.S. General Accounting Office. She is responsible for issues concerning the collection, use and dissemination of government information in an era of rapidly changing technology. Recently she has been heavily involved in directing studies concerning e-government, privacy, electronic records management and governmentwide information dissemination issues. She is a frequent panelist before the subcommittee, and we always appreciate your insight and support. You're recognized for your testimony.

Ms. KOONTZ. Thank you, Mr. Chairman. I appreciate the opportunity to participate in the subcommittee's hearing on the challenges of records management in an electronic era. As you know, agencies are increasingly moving to an operational environment in which electronic rather than paper records are used to document their activities and business processes. This migration is likely to accelerate in light of the E-government Act of 2002 which encourages the expansion of electronic government. This transformation is leading to improvements in the way Federal agencies work and interact with each other and with the public. However, the rapid evolution of information technology is creating challenges in managing and preserving electronic records. Complex electronic records are increasingly being created in a decentralized environment and in volumes that make it difficult to organize them and make them accessible.

Further storage media themselves are affected by the dual problems of obsolescence and deterioration. For example, few computers today have disk drives that can read information on 8 or 5¼ diskettes, even if the diskettes themselves remain readable. These problems are compounded as computer hardware and application software become obsolete. They may leave behind electronic records that can no longer be read. Unless these challenges are addressed, the government will be unable to effectively leverage the information it has and valuable government information may be lost forever.

We report last year that while NARA has responded to these challenges, most electronic records, including data bases of major Federal information systems, remain unscheduled, and records of historical value were not being identified and provided to NARA, and as a result, they were at risk of loss.

Three factors contributed to this condition. First, NARA has acknowledged that its policies and processes on electronic records have not yet evolved to reflect the modern recordkeeping environment; second, records management programs were generally afforded low priority by Federal agencies. A related issue was that agency management had not given priority to acquiring the technology tools required to manage electronic records.

Third, NARA was not performing systematic inspections of agency record programs. Such inspections are important as a means to evaluate individual agency records management programs, assess governmentwide progress in improving records, and manage and identify areas where guidance needs to be strengthened. We recommended that NARA develop strategies for raising agency man-

agement awareness of the importance of records management and for performing systematic inspections.

In the last year, NARA has taken steps to improve its guidance and address the lack of technology tools. In response to our recommendations, it has devised a reasonable strategy for raising awareness among senior agency management. In addition, it has developed a comprehensive approach to assessing and improving agency records management programs that includes identification of risks and priorities and inspections, but it has not yet described how this will be made an ongoing program and an implementation plan for the strategy has not yet been established. Until NARA fully addressed these issues, the risk is increased that records management programs will continue to show the weaknesses that led to the scheduling and disposition problems that we and NARA described in our earlier work.

NARA also faces significant challenges in acquiring an electronic record archive, an advanced system that is intended to address the problems associated with preserving electronic records and making them accessible. Specifically the plans, policies and practices that NARA is using to acquire the system do not, in many cases, conform to standards or to applicable Federal acquisition guidance.

In addition, NARA is unable to track the cost and schedule of the product adequately. Unless NARA addresses these issues, the risk is increased that the system will fail to meet user's needs, will cost more than currently estimated and will be delivered later than planned. In view of these risks, we have recommended that NARA address weaknesses in its acquisition plans and project schedule.

In conclusion, NARA and the Federal Government face significant challenges in managing electronic records. While NARA is responding to these challenges and is making progress, much work remains to be done to gain control over the massive numbers of electronic records that continue to grow and prevent the loss of valuable permanent records.

That concludes my statement. I would be happy to answer any questions.

Mr. PUTNAM. Thank you very much.

[The prepared statement of Ms. Koontz follows:]

United States General Accounting Office

GAO

Testimony

Before the Subcommittee on Technology, Information Policy, Intergovernmental Relations, and the Census, Committee on Government Reform, House of Representatives

For Release on Delivery
Expected at 10 a.m. EDT
Tuesday, July 8, 2003

ELECTRONIC RECORDS

Management and Preservation Pose Challenges

Statement of Linda D. Koontz
Director, Information Management Issues

GAO-03-936T

GAO
Accountability•Integrity•Reliability

Highlights

Highlights of GAO-03-936T, a testimony before the Subcommittee on Technology, Information Policy, Intergovernmental Relations, and the Census, Committee on Government Reform, House of Representatives

Why GAO Did This Study

The difficulties of managing, preserving, and providing access to the vast and rapidly growing volumes of electronic records produced by federal agencies present challenges to the National Archives and Records Administration (NARA), the nation's recordkeeper and archivist. Complex electronic records are being created in volumes that make them difficult to organize and keep accessible. These problems are compounded as computer hardware, application software, and even storage media become obsolete, as they may leave behind electronic records that can no longer be read. As a result, valuable government information may be lost. GAO was requested to testify, among other things, on NARA's recent actions to address the challenges of electronic records management, including its effort to address the problem of preserving electronic records by acquiring an advanced Electronic Records Archive (ERA).

www.gao.gov/cgi-bin/getrpt?GAO-03-936T.

To view the full product, click on the link above.
For more information, contact Linda Koontz at (202) 512-6240 or koontzl@gao.gov.

July 8, 2003

ELECTRONIC RECORDS

Management and Preservation Pose Challenges

What GAO Found

As reported in GAO's past work, most electronic records—including databases of major federal information systems—remained unscheduled: that is, their value had not been assessed, and their disposition—to destruction or archives—had not been determined. In addition, records of historical value were not being identified and provided to NARA; as a result, they were at risk of loss. NARA has begun to address these problems by taking steps to improve federal records management programs; among other things, it has (1) updated guidance to reflect new types of electronic records, (2) devised a strategy for raising awareness among senior agency management of the importance of good federal records management, and (3) devised a comprehensive approach to improving agency records management that includes inspections and identification of risks and priorities. Through these and other actions, NARA is making progress, but its approach to improving records management does not include provisions for using inspections to evaluate the efficacy of its governmentwide guidance, and an implementation plan for the approach has yet to be established. Without these elements, the risk is increased that federal records management problems will persist.

In addition to its efforts to improve records management, NARA is also acquiring ERA as a means to archive all types of electronic records and make them accessible. GAO found, however, that NARA faces significant challenges in acquiring ERA, a major information system. While NARA has made progress in building its organizational capabilities for acquiring major information systems, it has not developed adequate policies, plans and practices to guide the ERA acquisition or established the means to track the cost and schedule of the project. Unless NARA addresses these and other issues, the ERA system may not meet user expectations, and NARA may not have the information required to control the cost of the system or the time will take to complete it.

Master Copies of Electronic Records in NARA's Archives

Source: NARA

Mr. Chairman and Members of the Subcommittee:

I appreciate the opportunity to participate in the Subcommittee's hearing on the challenges of records management in an electronic era. As you know, agencies are increasingly moving to an operational environment in which electronic—rather than paper—records are used to document their activities and business processes. This migration to an increasingly electronic environment is likely to accelerate in light of the E-Government Act of 2002, which encourages the expansion of electronic government.[1]

This transformation is leading to improvements in the way federal agencies work and interact with each other and with the public, but it has also led to the generation of vast and rapidly growing volumes of electronic records. These records need to be properly managed, as they provide documentation and information that are necessary for essential government functions and for protecting government and citizen interests. Also, a small percentage of documents and other records are sufficiently significant that they must be preserved in archives for the historical record.[2]

Overall responsibility for the government's electronic records lies with the National Archives and Records Administration (NARA),[3] whose mission is twofold: (1) oversight of records management, which governs the life cycle of records (creation, maintenance and use, and disposition), and (2) archiving, which is the permanent preservation of documents and other records of historical interest. In carrying out these missions, NARA and agencies use a process known as scheduling to assess the value of records and determine their disposition.

As requested, in my remarks today, I will begin by providing some perspective on records management and its importance to the government in today's environment. In addition, I will

- describe the challenges that the government faces in managing and preserving electronic records,
- summarize the results of our June 2002 report on electronic record management issues,

[1] Electronic government refers to the use of technology, particularly Web-based Internet applications to enhance the access to and delivery of government information and services to citizens, business partners, employees, other agencies, and other entities.

[2] Of the total number of federal records, less than 3 percent are designated permanent.

[3] NARA's regulations implementing the Federal Records Act are found at 36 CFR 1200-1280.

- provide information on NARA's actions since we issued our report and assess its progress in responding to the recommendations we made to improve governmentwide records management, and
- assess NARA's efforts to acquire an advanced electronic record archive (ERA), intended to preserve and provide access to electronic records.

Our evaluation of records management challenges and issues is based on our June 2002 report, entitled *Information Management: Challenges in Managing and Preserving Electronic Records.* To update NARA actions since our report, we analyzed key documents and plans, and we interviewed appropriate agency officials. Our assessment of NARA's efforts to acquire ERA reflects the results of ongoing work, the results of which will be reported in August 2003.

We performed our update work from June to July 2003 in accordance with generally accepted auditing standards.

Results in Brief

The rapid evolution of information technology is creating challenges in managing and preserving electronic records. Complex electronic records are increasingly being created in a decentralized environment and in volumes that make it difficult to organize them and make them accessible. Further, storage media themselves are affected by the dual problems of obsolescence and deterioration: for example, few computers today have disk drives that can read information stored on 8 or 5¼-inch diskettes, even if the diskettes themselves remain readable. These problems are compounded as computer hardware and application software become obsolete, as they may leave behind electronic records that can no longer be read. Unless these challenges are addressed, valuable government information may be lost forever.

Our past work has shown that while NARA has responded to the challenges associated with managing and preserving electronic records, most electronic records—including databases of major federal information systems—remained unscheduled, and records of historical value were not being identified and provided to NARA; as a result, they were at risk of loss. A number of factors contributed to this condition:

- NARA acknowledged that its policies and processes on electronic records had not yet evolved to reflect the modern recordkeeping

environment: records created electronically in decentralized processes.

- Records management programs were generally afforded low priority by federal agencies. A related issue was that agency management had not given priority to acquiring the more sophisticated and expensive information technology required to manage records in an electronic environment.
- NARA was also not performing systematic inspections of agency records programs. Such inspections are important as a means to evaluate individual agency records management programs, assess governmentwide progress in improving records management, and identify agency implementation issues and areas where guidance needs to be strengthened.

We recommended that NARA develop strategies for raising agency management awareness of the importance of records management and for performing systematic inspections.

In the year since our report was issued, NARA has taken steps to improve its guidance and address the lack of technology tools. In response to our recommendations, it has devised a reasonable strategy for raising awareness among senior agency management. In addition, it has devised a comprehensive approach to improving agency records management that includes inspections and identification of risks and priorities, but its approach does not include provisions for using inspections to evaluate the efficacy of its governmentwide guidance, and an implementation plan for the approach has yet to be established. Until NARA fully addresses the need to assess and improve agency records management programs and develop an implementation plan, the risk is increased that records management programs will continue to show the weaknesses that led to the scheduling and disposition problems tha we and NARA described in earlier work.

To address the problems associated with preserving electronic records in a rapidly changing technological environment, NARA is acquiring ERA, an advanced archiving system whose acquisition presents significant challenges. First, while NARA has made progress in strengthening its organizational capabilities for acquiri major information systems, these capabilities are not fully established. For example, NARA has completed some elements of an information security program; however, several key areas—suc as developing individual system security plans—have not yet been addressed. Second, the policies, plans, and practices that NARA is using to acquire ERA do not in many cases conform to standards o

to applicable federal acquisition guidance. Finally, NARA is unable to objectively track the cost and schedule of the ERA project. Unless NARA improves its acquisition planning, the risk is increased that the ERA system will fail to meet user expectations, and that NARA will not have the information required to control the cost of the system or the time it will take to complete it. In view of these risks, we have recommended that NARA address weaknesses in its acquisition policies, plans, and practices and its ERA schedule.

Background

NARA's mission is to ensure "ready access to essential evidence" for the public, the president, the Congress, and the Courts. NARA's responsibilities stem from the Federal Records Act,[4] which requires each federal agency to make and preserve records that

- document the organization, functions, policies, decisions, procedures, and essential transactions of the agency and
- provide the information necessary to protect the legal and financial rights of the government and of persons directly affected by the agency's activities.

Federal records must be managed to ensure that the information that they contain is available when needed. According to NARA, without effective records management, the records needed to document citizens' rights, actions for which federal officials are responsible, and the historical experience of the nation will be at risk of loss, deterioration, or destruction.

Records management is defined as the policies, procedures, guidance, tools and techniques, resources, and training needed to design and maintain reliable and trustworthy records systems. Records must be managed throughout their life cycle: from creation, through maintenance and use, to final disposition. Temporary records—those used in everyday operations but lacking historic value—are ultimately destroyed. Permanent records—those judged to be of historic value—are preserved through archiving.

With NARA's oversight and assistance, each agency is responsible for managing its own records at all phases of the life cycle, with the exception of the archiving of permanent records (which is NARA's responsibility).

[4] 44 U.S.C. chapters 21, 29, 31, and 33.

NARA Is Responsible for Oversight of Records Management and for Archiving

NARA is responsible for

- issuing records management guidance;
- working with agencies to implement effective controls over the creation, maintenance, and use of records in the conduct of agency business;
- providing oversight of agencies' records management programs; and
- providing storage facilities for certain temporary agency records.

The Federal Records Act also authorizes NARA to conduct inspections of agency records and records management programs.

Records Are Managed through Scheduling

NARA works with agencies to identify and inventory records; to appraise their value; and to determine whether they are temporary or permanent, how long the temporary records should be kept, and under what conditions both the temporary and permanent records should be kept. This process is called *scheduling*. No record may be destroyed unless it has been scheduled. Thus, for temporary records the schedule is of critical importance, because it provides the authority to dispose of the record after a specified time.

Records are governed by schedules that are either (1) specific to an agency or (2) general—that is, common to several agencies or across the government. According to NARA, records covered by general records schedules make up about a third of all federal records. For the other two thirds, NARA and the agencies must agree upon specific records schedules. Once a schedule has been approved, the agency must issue it as a management directive, train employees in its use, apply its provisions to temporary and permanent records, and evaluate the results.

Guidance Addresses Electronic Records

While the Federal Records Act covers documentary material regardless of physical form or media, records management and archiving were until recently largely focused on handling paper documents. With the advent of computers, both records management and archiving have had to take into account the creation of records in varieties of electronic formats.

NARA's basic guidance for the management of electronic records is in the form of a regulation at 36 CFR Part 1234. This guidance is

supplemented by the issuance of periodic NARA bulletins and a records management handbook, *Disposition of Federal Records*.

For electronic records, NARA's guidance sets forth two basic requirements.

- First, agencies are required to maintain an inventory of all agency information systems. The inventory should identify (1) the system's name, (2) its purpose, (3) the agency programs supported by the system, (4) data inputs, sources, and outputs, (5) the information content of databases, and (6) the system's hardware and software environment.
- Second, NARA requires agencies to schedule the electronic records maintained in their systems. Agencies must schedule those records either under specific schedules (completed through submission and approval of Standard Form 115, *Request for Records Disposition Authority*) or pursuant to a general records schedule. NARA relies on this combination of inventory and scheduling requirements to ensure that management of agency electronic records is consistent with the Federal Records Act.

NARA has also established a general records schedule for electron. records. General Records Schedule 20 (GRS 20) authorizes the disposal of certain categories of temporary electronic records. It has been revised several times over the years in response to developments in information technology, as well as legal challenges.

GRS 20 applies to electronic records created both in computer centers engaged in large-scale data processing and in the office automation environment. GRS 20 authorizes the disposal of certain types of electronic records associated with large data base systems, (such as inputs, outputs, and processing files), as well as the deletion of the electronic version of records on word processing and electronic mail systems once a recordkeeping copy has been made. Since most agency recordkeeping systems are paper files, GRS 20 essentially authorizes agencies to destroy E-mail and word-processing files once they are printed. (Recall that records not covered by a general records schedule may not be destroyed unless authorized by a records schedule that has been approved by NARA.)

GRS 20 does not address many common products of electronic information processing, particularly those that result from the now prevalent distributed, end-user computing environment. For example, although the guidance addresses the disposition of certain types of electronic records associated with large databases, it does not specifically address the disposition of electronic databases created by microcomputer users. In addition, GRS 20 does not

address more recent forms of electronic records such as Web pages and portable document format (PDF) files.[6]

NARA Archives Permanent Records of Historical Interest

As the nation's archivist, NARA accepts for deposit to its archives those records of federal agencies, the Congress, the Architect of the Capitol, and the Supreme Court that are determined to have sufficient historical or other value to warrant their continued preservation by the U.S. government. NARA also accepts papers and other historical materials of the Presidents of the United States, documents from private sources that are appropriate for preservation (including electronic records, motion picture films, still pictures, and sound recordings), and records from agencies whose existence has been terminated.

To ensure that permanent electronic records are preserved, each agency must transfer electronic records to NARA in accordance with the agency's records disposition schedule. NARA accepts for archiving electronic records that are in text-based formats, such as databases and certain text-based geographic information system (GIS)[6] files. In addition, NARA accepts E-mail records and attachments, several forms of scanned images of text files, and PDF files. It does not accept Web pages, word processor files, or relational databases. (Although NARA does not as yet accept such files for archiving, they must still be scheduled.)[7]

Agencies Are Beginning to Automate Management of Electronic Records

In response to the difficulty of manually managing electronic records, agencies are turning to automated records management applications to help automate electronic records management lifecycle processes. The primary functions of these applications include categorizing and locating records and identifying records that are due for disposition, as well as storing, retrieving, and disposing of electronic records that are maintained in repositories.

[5] PDF is a proprietary format of Adobe Systems, Inc., that preserves the fonts, formatting, graphics, and color of any source document, regardless of the application and platform used to create it.

[6] A geographic information system is a computer system for capturing, storing, checking, integrating, manipulating, analyzing, and displaying data related to positions on the Earth's surface. Typically, a GIS is used for handling maps of one kind or another. These might be represented as several different layers where each layer holds data about a particular kind of feature (e.g., roads). Each feature is linked to a position on the graphical image of a map.

[7] NARA's requirements for the transfer of agency electronic records are provided through 36 CFR Part 1228. NARA's Web site provides further guidance on this topic under the heading Permanent E-Records Transfers to NARA Deliverables. (http://www.archives.gov/records_management/initiatives/transfer_to_nara.html)

Also, some applications are beginning to be designed to automatically classify electronic records and assign them to an appropriate records retention and disposition category.

The Department of Defense (DOD), which is pioneering the assessment and use of records management applications, has published application standards and established a certification program.[8] DOD standard 5015.2, endorsed by NARA, includes the requirement that records management applications acquired by DOD components after 1999 be certified to meet this standard.[9]

NARA Is Currently Pursuing Initiatives Related to Electronic Records Management

NARA is pursuing other interrelated efforts that address records management (including electronic records). Three major initiatives are

- NARA's effort on Redesign of Federal Records Management;
- the Electronic Records Management initiative, one of 25 e-government initiatives sponsored by the Office of Management and Budget (OMB), and
- the acquisition of an advanced Electronic Records Archives (ERA).

In 2000, NARA began a three-stage effort to redesign federal records management. First, in 2001, NARA produced a report based on information on federal records management that it collected and analyzed.[10] Second, it used this report as a starting point to revise the regulations, policies, and processes for managing federal records and to develop a set of strategies to support federal records management. As a result of this analysis, in July 2002 NARA issued a draft proposal for the redesign of federal records management. Third, based on comments received on the proposal, it is developing a redesigned records scheduling, appraisal, and accessioning process, as well as prototype and functional requirements for automated tools for the redesigned process. The redesign is planned as a multiyear process (2003 to 2006), during which NARA intends

[8] Department of Defense, *Design Criteria Standard for Electronic Records Management Software Applications*, DOD 5015.2-STD (November 1997) (http://www.dtic.mil/whs/directives/corres/html/50152std.htm).

[9] DOD 5015.2-STD requires that records management applications be able to manage records regardless of their media.

[10] SRA International, Inc., *Report on Current Recordkeeping Practices within the Federal Governmen.* (Dec. 10, 2001) (http://www.nara.gov/records/rkreport.html). This document reports results from two sources: a recordkeeping study performed by NARA's contractor, SRA International, and a series of records system analyses performed by NARA staff.

to address the scheduling and appraisal of federal records in all formats.

The overall purpose of the Electronic Records Management (ERM) initiative is to help agencies better manage their electronic records, so that records information can be effectively used to support timely and effective decision making, enhance service delivery, and ensure accountability. The initiative is intended to provide a variety of tools to address immediate and longer term agency needs. NARA is the managing partner agency for the overall ERM initiative.

The goals for the advanced ERA system are that it will be able to preserve and provide access to any kind of electronic record, free from dependency on any specific hardware or software, so that the agency can carry out its mission into the future. NARA plans for ERA to be a distributed system, allowing storage and management of massive record collections at a variety of installations, with accessibility provided via the Internet. NARA is planning to build the system in five increments, with the last increment scheduled to be complete in 2010.

Inherent Features of Electronic Records Make Management and Preservation Challenging

The rapid evolution of information technology makes the task of managing and preserving electronic records complex and costly. Part of the challenge of managing electronic records is that they are produced by a mix of information systems, which vary not only by type but by generation of technology: the mainframe, the personal computer, and the Internet. Each generation of technology brought in new systems and capabilities without displacing the older systems.[11] Thus, organizations have to manage and preserve electronic records associated with a wide range of systems, technologies, and formats. These records are stored in specific formats and cannot be read without software and hardware—sometimes the specific types of hardware and software on which they were created.

Several factors contribute to the challenge of managing and preserving electronic records:

[11] International Council on Archives, *Guide for Managing Electronic Records from an Archival Perspective* (Paris: February 1997).

- *Massive volumes of electronic data require automated solutions.*

  Electronic records are increasingly being created in volumes that pose a significant technical challenge to our ability to organize them and make them accessible. For example, among the candidates for archiving are military intelligence records comprising more than 1 billion electronic messages, reports, cables, and memorandums, as well as over 50 million electronic court case files. Managing such large volumes is clearly not possible without automation.

- *Control of electronic records is difficult in a decentralized computing environment.*

  The challenge of managing electronic records significantly increases with the decentralization of the computing environment. In the centralized environment of a mainframe computer, it is easier to identify, assess, and manage electronic records than it is in the decentralized environment of agencies' office automation systems, where every user is creating electronic files that may constitute a formal record and thus should be preserved.

- *The complexity of electronic records precludes simple transfer to paper.*

  Electronic records have evolved from simple text-based files to complex digital objects that may contain embedded images (still and moving), drawings, sounds, hyperlinks, or spreadsheets with computational formulas. Some portions of electronic records, such as the content of dynamic Web pages, are created on the fly from databases and exist only during the viewing session. Others, such as E-mail, may contain multiple attachments, and they may be threaded (that is, related E-mail messages are linked into send–reply chains). These records cannot be converted to paper or text formats without the loss of context, functionality, and information.

- *Obsolescent and aging storage media put electronic records at risk.*

  Storage media are affected by the dual problems of obsolescence and decay. They are fragile, have limited shelf life, and become obsolete in a few years. For example, few computers today have disk drives that can read information stored on 8- or 5¼-inch diskettes, even if the diskettes themselves remain readable.

- *Electronic records are dependent on evolving software and hardware.*

  Electronic records are created on computers with software ranging from word-processors to E-mail programs. As computer hardware and application software become obsolete, they may leave behind

electronic records that cannot be read without the original hardware and software.

Past GAO Work Highlighted Electronic Records Challenges

In June 2002,[12] we reported that NARA had responded to the challenges associated with managing and preserving electronic records. However, most electronic records—including databases of major federal information systems—remained unscheduled, and records of historical value were not being identified and provided to NARA; as a result, they were at risk of loss. A number of factors contributed to this condition:

- NARA acknowledged that its policies and processes on electronic records had not yet evolved to reflect the modern recordkeeping environment: records created electronically in decentralized processes.
- Records management programs were generally afforded low priority by federal agencies. A related issue was that agency management had not given priority to acquiring the more sophisticated and expensive information technology required to manage records in an electronic environment.
- NARA was also not performing systematic inspections of agency records programs. Such inspections are important as a means to evaluate individual agency records management programs, assess governmentwide progress in improving records management, and identify agency implementation issues and areas where guidance needs to be strengthened.

We also provided some confirmation of NARA's findings regarding records scheduling and disposition: our review at four agencies (Commerce, Housing and Urban Development, Veterans Affairs, and State) elicited a collective estimate that less than 10 percent of mission-critical systems were inventoried. As a result, for these four agencies alone, over 800 systems had not been inventoried, and the electronic records maintained in them had not been scheduled.[13] Scheduling the electronic records in a large number of major

[12] U.S. General Accounting Office, *Information Management: Challenges in Managing and Preserving Electronic Records*, GAO-02-586 (Washington, D.C.: June 17, 2002).

[13] According to information collected by the Office of Management and Budget in November 1999 as part of the federal government's effort to assess the Year 2000 computing challenge, 6,435 mission-critical systems were reported at the 24 major agencies. For the 4 agencies we reviewed, the number of mission-critical systems was reported to be 907. Subcommittee on Government Management, Information, and Technology, House Committee on Government Reform, *Federal Government Earns B+ on a Final Y2K Report Card*, news release (Washington, D.C.: Nov. 22, 1999).

information systems presents an enormous challenge, particularly since it generally takes NARA, in conjunction with agencies, well over 6 months to approve a new schedule.

Failure to inventory systems and schedule records places these records at risk. The absence of inventories and schedules means that NARA and agencies have not examined the contents of these information systems to identify official government records, appraised the value of these records, determined appropriate disposition, and directed and trained employees in how to maintain and when and how to dispose of these records. As a result, temporary records may remain on hard drives and other media long after they are needed or could be moved to less costly forms of storage. In addition, there is increased risk that these records may be deleted prematurely while still needed for fiscal, legal, and administrative purposes. Further, the lack of scheduling presents risks to the preservation of permanent records of historic significance.

Guidance on Management of Electronic Records Was Not Up to Date

NARA acknowledged in 2001 that its policies and processes on electronic records had not yet evolved to reflect the modern recordkeeping environment: records created electronically in decentralized processes.[14] Despite repeated attempts to clarify its electronic records guidance through a succession of bulletins, the guidance was incomplete and confusing. It did not provide comprehensive disposition instructions for electronic records maintained in many of the common types of formats produced by federal agencies, including Web pages and spreadsheets. To support their missions, many agencies had to maintain such records—often in large volumes—with little guidance from NARA.

Agency Records Management Programs Had Low Priority and Did Not Have Technology Tools

NARA's study concluded that records management was not even "on the radar scope" of agency leaders. Further, records officers had little clout and did not appear to have much involvement in or influence on programmatic business processes or the development of information systems designed to support them. New government

[14] National Archives and Records Administration, *An Overview of Three Projects Relating to the Changing Federal Recordkeeping Environment* (January 2001) *(http://www.nara.gov/records/rmioverview.html)*.

employees seldom received any formal, initial records management training. One agency told NARA that records management was "number 26 on our list of top 25 priorities."

Further, records management is generally considered a "support" activity. Since support functions are typically seen as the most dispensable in agencies, resources for and focus on these functions are often limited. Also, as NARA's study noted, federal downsizing may have negatively affected records management and staffing resources in agencies.

In our June 2002 report, we recommended that the Archivist of the United States address the priority problem by developing a documented strategy for raising agency senior management awareness of and commitment to records management principles, functions, and programs.

Related to the priority issue is the need for appropriate information technology tools to respond to the technical challenge of electronic records management: for electronic records to be managed effectively, agencies require a level of technology that was not necessary for paper-based records management programs. Unless management is focused on records management, priority is not given to acquiring or upgrading the technology required to manage records in an electronic environment. Agencies that do invest in electronic records management systems tend to do so because they value good records management and have a critical need to retrieve information efficiently. In other agencies, despite the growth of electronic media, agency records systems are predominantly in paper format rather than electronic. According to NARA's study, many agencies were either planning or piloting information technology initiatives to support electronic records management, but their movement to electronic systems is constrained by the level of financial support provided for records management.

Inspections of Agency Records Management Programs Were Limited

NARA is responsible, under the Federal Records Act, for conducting inspections or surveys of agency records and records management programs and practices. Its implementing regulations require NARA to select agencies to be inspected (1) on the basis of perceived need by NARA, (2) by specific request by the agency, or (3) on the basis of a compliance monitoring cycle developed by NARA.[15] In all instances, NARA is to determine the scope of the inspection. Such

[15] CFR 1220.54 (a).

inspections provide not only the means to assess and improve individual agency records management programs but also the opportunity for NARA to determine overall progress in improving agency records management and identify problem areas that need to be addressed in its guidance.

In 2000, NARA changed its method of performing inspections: rather than performing a small number of comprehensive agency reviews, it instituted an approach that it refers to as "targeted assistance." NARA decided that its previous approach to inspections was basically flawed, because it could reach only about three agencies per year,[16] and because the inspections were often perceived negatively by agencies, resulting in a list of records management problems that agencies then had to resolve on their own. Under the targeted assistance approach, NARA works with agencies, providing them with guidance, assistance, or training in any area of records management.

However, we pointed out in our June 2002 report that this approach, although it may improve records management in the targeted agencies, is not a substitute for systematic inspections and evaluations of federal records programs. Targeted assistance has significant limitations because it is voluntary and, according to NARA, initiated by agency request. Thus, only agencies requesting assistance are evaluated, and the scope and the focus of the assistance are not determined by NARA but by the requesting agency.

In light of these limitations, we recommended in June 2002 that the Archivist develop a documented strategy for conducting systematic inspections of agency records management programs to (1) periodically assess agency progress in improving records management programs and (2) evaluate the efficacy of NARA's governmentwide guidance.

NARA Is Continuing to Respond to Records Management Challenges, but Its Progress on Inspections is Limited

Since June 2002, NARA has taken steps to strengthen its guidance, to address the low priority accorded to records management programs and the associated lack of technology tools, and to revise its approach to inspections as part of a comprehensive strategy for

[16] Between 1996 and 2000, NARA performed 16 inspections of agency records management programs; only 2 of the 24 major executive departments or agencies were reviewed, with most evaluations focused on component organizations or independent agencies.

assessing agencies' management of records. However, NARA's plans to implement its comprehensive new strategy are not yet complete. Although the strategy describes a reasonably systematic approach that allows NARA to focus its resources appropriately and to use inspections and other interventions to assess and improve federal records management, it does not yet include a description of how NARA will establish an ongoing program.

NARA Is Improving Records Management Guidance

Since our 2002 report, NARA has taken steps to update its guidance on electronic records management in various areas. For example, although 36 CFR Part 1234, the basic guidance on electronic records, has not been updated to reflect new types of electronic records, NARA has produced a variety of guidance on electronic records. A new General Records Schedule, GRS 24, "Information Technology Operations and Management Records," was issued on April 28, 2003. In addition, "Records Management Guidance for PKI-Unique Administrative Records," which was jointly developed by NARA and the Federal Public Key Infrastructure Steering Committee's Legal/Policy Working Group, was issued on March 14, 2003. As part of its e-government initiative, NARA has just released guidance on evaluating funding proposals for electronic records management systems through capital planning processes.

NARA has also supplemented its disposition guidance[17] as a result of the project on transfer of permanent electronic records under its e-government initiative: this guidance covers transferring permanent E-mail records and attachments, several forms of scanned images of text files, and PDF, and it expanded the methods by which agencies could transfer electronic records to NARA for archiving. NARA is also planning to expand the capability of its current systems for archiving electronic records by accommodating additional electronic record formats and volumes. However, according to NARA, agencies have not yet transferred electronic records in these formats to NARA; these records may not be scheduled or may not yet be eligible for transfer.

In addition, as part of the policy analysis in its effort to redesign federal records management, NARA has stated that it plans to identify policies, procedures, regulations, and guidance that would need to be modified in light of the proposed redesign.

[17] These requirements are found in 36 CFR Part 1228.

Efforts Continue to Raise the Priority of Agency Records Management Programs and Address Technology Issues

In response to our recommendation that it develop a documented strategy for raising agency senior management awareness of records management, NARA devised a strategy intended to raise awareness of the importance of agency records management. The strategy includes two goals:

- increased senior-level awareness of the importance of records management, particularly electronic records management, across the federal government and in specific agencies, and
- increased senior-level understanding of how effective records management programs support the business needs of specific agencies and the federal government as a whole.

As part of its strategy, NARA identified a number of activities that its senior leaders will conduct, including briefing agency program leaders on the importance of records and information management in general and on specific issues (such as electronic record keeping requirements, litigation exposure, and vital records), participating in establishing or closing out certain targeted assistance agreements, and pursuing promotional activities such as making speeches and holding conferences.

NARA has also developed an implementation plan, which establishes goals, timeframes, and required resources for fiscal year 2003. For example, the plan contains a goal of conducting six agency briefings by the end of September; three have been completed to date, and a fourth has been scheduled for mid-July. A similar implementation plan for fiscal year 2004 is to be developed by September 1.

NARA's strategy for raising senior agency management awareness appears reasonable, and if carried out effectively could help to mitigate the problem of the low priority given to records management.

Since our June 2002 report, some steps have also been taken to address the lack of technology tools to manage electronic records. In January 2003, NARA recommended that agencies use version 2 of DOD standard 5015.2, which sets forth a set of requirements for records management applications, including that they be able to manage records regardless of their media. The effort to promulgate this standard was part of the electronic information management standards project under the ERM initiative. Under the standard, DOD is to certify records management applications as meeting the

standard; as of the end of June 2003, DOD had certified 43 applications. The availability of applications that conform to the standard may be helpful in encouraging agencies to adopt records management systems that address electronic records.

NARA Plans Comprehensive Approach to Improving Agency Records Management

In response to its own mission needs and our recommendations of June 2002 regarding its inspection program, NARA has documented a new strategy for assessing agencies' management of records. This strategy is described in draft documents that describe NARA's plans for setting priorities and for conducting inspections and studies. The new approach is now being piloted with the Department of Homeland Security; the results of the pilot—expected by September 30, 2003—will determine whether it is extended governmentwide.

The main features of the draft strategy are as follows:

- NARA will evaluate agencies and work processes in terms of risk to records, implications for legal rights and accountability, and the quantity and value of the permanent records; it will focus its resources on high-priority areas. This process of assessing risks and priorities will involve NARA staff with subject-matter and agency expertise, and it will address records management governmentwide.
- NARA plans to use a variety of means to address areas identified for attention through its risk and priority assessment. Among the means being considered are targeted assistance, records management studies, and inspections.[18] The strategy indicates that NARA has changed its approach to targeted assistance: Rather than using it only when an agency requests assistance, NARA intends to recommend that an agency accept targeted assistance when NARA has identified records management issues at that agency that require attention. In addition, NARA plans to perform studies on records management best practices as a means not only to encourage good records management practices throughout government, but also to recognize agencies whose records management programs have exemplary features. According to the strategy, inspections will be conducted only under exceptional circumstances, when the risk to records is deemed high, and after other means have failed to mitigate risks (e.g., targeted assistance, training, and so on).
- NARA intends to focus on the core functions of the federal government, rather than on individual agencies. It will use as its

[18] Other means include standard and tailored training programs, certification programs for records managers, and various forms of advocacy.

starting point the business areas defined in the Business Reference Model[19] of the Federal Enterprise Architecture.[20] By focusing on the Business Reference Model's broad activities and work processes, which cut across agency lines, NARA may inspect a single agency or a group of agencies in one line of business.

Although NARA's strategy appears to be a reasonably systematic approach that allows it to focus its resources appropriately and to use inspections and other interventions to assess and improve federal records management, it is not yet complete. Specifically, the draft strategy does not yet include a description of how NARA will establish an ongoing program. For example, the priority assessment plan does not indicate whether NARA will revise its risk identification process as circumstances warrant, or if this a single-time occurrence. NARA officials have said that the agency will update its priority and risk assessments periodically, but this is not yet reflected in the plan. Further, the strategy states that the results of studies may be used to improve guidance, but it does not create a similar feedback loop for inspection results. While records management guidance may benefit from the "best practices" identified in studies, inspection results could also identify areas where guidance needs to be clarified, augmented, and strengthened. Finally, no implementation plan or schedule for this new strategy has yet been devised.

Without a strategy that provides for establishing an ongoing program that includes a feedback cycle, as well as complete implementation plans that fully reflect that strategy, NARA's efforts to assess records management programs may not provide it with the information that it needs to improve its guidance and to support its redesign of federal records management.

NARA's Acquisition of ERA Continues to Face Risks

In addition to its efforts to improve records management across the government, NARA is also acquiring ERA as a means to archive all types of electronic records and make them accessible, regardless of changes to hardware and software over time. However, NARA faces significant challenges in acquiring ERA. ERA will be a major

[19] The Business Reference Model is a function driven framework for describing the business operations of the federal government independent of the agencies that perform them.

[20] The Federal Enterprise Architecture is a business-based foundation that provides a common framework for governmentwide improvements in key areas such as performance measurement and e-government. The Office of Management and Budget began development of the Federal Enterprise Architecture in February 2002.

information system; NARA has no previous experience in acquiring major information systems. Further, no comparable electronic archive system is now in existence, in terms of either complexity or scale. Finally, technology necessary to address some key requirements of ERA is not commercially available and will have to be developed. In light of these challenges, NARA will face significant difficulties in its ERA acquisition unless it addresses

- its information technology (IT) organizational capabilities;
- ERA system acquisition policies, plans, and practices; and
- its ability to control ERA's cost and schedule.

NARA has indicated that it needs to strengthen its IT organizational capabilities and has been taking steps to do so in three key areas:

- *IT investment management* provides a systematic method for agencies to minimize risks while maximizing the return on IT investments.
- *An enterprise architecture* provides a description—in useful models, diagrams, and narrative—of the mode of operation for an agency. It provides a perspective on agency operations both for the current environment and for the target environment, as well as a transition plan for sequencing from the current to the target environment. Managed properly, an enterprise architecture can clarify and help optimize the dependencies and relationships among an agency's business operations and the underlying IT infrastructure and applications that support these operations.
- *Information security* is an important consideration for any organization that depends on information systems to carry out its mission. Our study of security management best practices found that leading organizations manage their information security risk through an ongoing cycle of risk management.

NARA has made progress in strengthening these capabilities, but these efforts are incomplete. For example, NARA has improved its IT investment management. However, although it is continuing to develop an enterprise architecture, NARA does not plan to complete its target architecture in time to influence the ERA system definition and requirements. In addition, it has completed some elements of an information security program, but several key areas have not yet been addressed (such as individual system security plans), and NARA has not assessed the security risks to its major information systems.

In addition, NARA has developed policies, plans, and practices to guide the ERA acquisition, but these do not consistently conform to industry standards and federal acquisition guidance. NARA has chosen to follow Institute of Electrical and Electronics Engineers (IEEE) standards[21] in developing its policies, plans, and practices. Examples of these include (1) a concept of operations that describes the characteristics of a proposed system from the users' viewpoint and provides the framework for all subsequent activities leading to system deployment, (2) an acquisition strategy that establishes how detailed acquisition planning and program execution will be accomplished, and (3) a risk management plan to identify potential problems and adjusting the acquisition to mitigate them. However, key policy and planning documents are missing elements that are required by the standards and federal acquisition guidance: for example, the ERA acquisition strategy did not satisfy 15 of 32 content elements required by the relevant IEEE standard.

Further, NARA is unable to track the cost and schedule of the ERA project. The ERA schedule does not include all program tasks and lacks a work breakdown structure, which would include detail on the amount of work and resources required to complete each task.

Unless NARA can address these issues, the risk is increased that the ERA system will fail to meet user expectations, and that NARA may not have the information required to control the cost of the system or the time it will take to complete it.

In light of these risks, our briefing included recommendations to NARA to address the weaknesses in its acquisition policies, plans and procedures and to improve its ability to adequately track the project's cost and schedule.

In summary, NARA and the federal government face significant challenges in managing electronic records, which are largely due to the rapidly changing technological environment. While NARA is responding to these challenges with its various initiatives, much work remains to be done. Specifically, NARA has implemented our recommendation to address the low priority given to records management programs. The agency's advocacy strategy, if implemented effectively, could help raise awareness of the importance of records management governmentwide. However,

[21] The Institute of Electrical and Electronics Engineers, *12207.0 Standard for Information Technology—Software Life Cycle Processes*; 12207.1 *Standard for Information Technology—Software Life Cycle Processes—Life Cycle Data*; and 12207.2 *Standard for Information Technology—Software Life Cycle Processes—Implementation Considerations*.

NARA has not fully responded to our recommendation for a revitalized inspection program. Further, while it is making progress in building its capacity to acquire a major system, it has not developed adequate policies, plans, and practices to guide the ERA acquisition or established an effective means to track the project's cost and schedule. Until NARA fully addresses these challenges, the success of the ERA project remains at risk, the government may not be able to gain control over the massive numbers of electronic records that continue to grow every day, and permanent records of historical value may be subject to loss.

Mr. Chairman, this concludes my statement. I would be pleased to answer any questions that you or other members of the subcommittee may have at this time.

Contact and Acknowledgements

If you should have any questions about this testimony, please contact me at (202) 512-6222 or via E-mail at *koontzl@gao.gov*. Other major contributors to this testimony included Tim Case, Barbara Collier, Mirko Dolak, and Elena Epps.

(310375)

Mr. PUTNAM. I want to thank all of you for your both written and verbal testimony, and it has certainly set the table for an interesting hearing.

Before we start peeling this onion down to several layers, I want to start in a very basic way. If you would, answer for me how you determine what records are worthy of archiving, what records are worthy of being kept for posterity and how much of that is defined in the law. What form it should be saved in and what the life cycle is of those forms? In other words, an old TSR 80 RadioShack computer disk that looks like a cassette tape, what is the life span of that versus a 4½ inch disk, 3½ inch disk or microfiche, microfilm, CD ROM? How long do those things live in a manner that we can retrieve that information?

And what process do we have for converting that? And is the conversion of the format in and of itself changing history? In other words, we wouldn't be content to have run a Xerox copy of the Declaration of Independence and feel like we'd saved the Declaration of Independence, and that is an extreme example, but certainly you can address that spectrum of issue.

So that is a lot to throw out there. But Governor, if you could begin to attempt to address these issues for us to get us going.

Mr. CARLIN. By law, it is our responsibility. In fact, in the end, my sign-off on the scheduling of records, in practice we historically have worked with agencies to evaluate records to determine whether they should be ultimately an accession to the archives which works out to be 2 or 3 percent of what is originally created, which one should be disposed of earlier and how long they should be kept for appropriate business purposes as well as protecting rights and entitlements for accountability reasons and so forth. That has been a back-and-forth with agency process, and, on any disposal, communication with stakeholders through a publication in the Federal Register.

As indicated in testimony that you have here today, we are working to overhaul that process, because we know, with the massive size of the Federal Government, with all of the records now that have been created electronically, with more and more of them wanting to be the record copy, that we must have a more efficient, effective way of scheduling records. Things change too fast for it to be an old system when a schedule stays in place for a long period of time. We have made a lot of progress. We are literally testing some of those changes with agencies as we speak, and I'm confident in the end we will have a system that will allow us to stay current with the technology here in which we live and schedule records appropriately, but it is our responsibility. We accept that responsibility, and work with agencies to carry it out as well as with stakeholders.

On the technology issue, the second part of the question, if I might, I would yield to my colleague here who probably can better explain to you that aspect of the question.

Mr. CAHOON. Thank you, Mr. Chairman. The technology chosen for any particular type of record often depends on the technology that was used to create it, and from there the technology required to ensure that those records are available from generation to generation. Microfiche, for example, can last, if preserved properly, for

hundreds of years. When that microfiche begins to appear to deteriorate, then that microfiche can be recopied using the same technology because it is human readable.

Electronic records, however, present a more difficult challenge. Electronic records will come in many different formats when they are created but ultimately, for us to preserve those records for the long term, we need to find formats that will enable us to remove the dependence that we often find that those records have on the technology that created them. We need to find technology independent formats for them.

The media, which is one of the issues that you raised—the media presents an interesting problem. We find that the best way to deal with the media at this point is to migrate the media as technology changes. For example, we might record information on cartridge tapes, 3480 cartridge tapes. When those become obsolete, we might convert that information to 3590 cartridge tapes and continue to progress as the technology evolves. So we're continually making copies—faithful copies of those original records.

So to sum up, the technology used to migrate records or to—excuse me, to preserve records depends on the original format that they came in and on the strategies we need to engage to migrate them from one technology to another.

Mr. PUTNAM. That makes sense for ordinary memos and personnel records and payroll records, but when you get into the GIS or maps or digital photography, so much of what we generate in great volumes in this day and age. Do we save the original photograph or the original map, or are we content to scan it and leave it on a cartridge or a disk or microfilm? How was something like that handled that is, at best, different than just ordinary memorandums or other documents?

Mr. CAHOON. The technology that creates these electronic records, a geographic information system record, for example, is highly dependent on the software that created it and the format that record is actually stored in. And so, the intent would be to find a format, ultimately, that would free that record from its dependence on the format that it was created in.

So in your example, where we're concerned about scanning a map, that's different, Mr. Chairman, from the technology used to capture the record originally and electronically and evolve it over time. So for scanning a map, for example, we would find a technology independent format to store that map in and be able to migrate then that map from generation through generation of media technology.

For a geographic information system where the information for the data associated with it is proprietary, we would try to find a technology independent format for that, and then migrate it as faithfully as we could through generations of media technology.

Does that respond to your question?

Mr. PUTNAM. I'm not sure. Our capabilities today are so dramatically different than when Lewis and Clark were drawing their own maps or when Cartier was charting the St. Lawrence. I mean, it seems obvious to us today that is historic and worthy of preservation as it is, but if we're shooting tens of thousands of pictures over Baghdad in a 90-day period of time, do we save all of those as they

are because they are historic—as they are? Or do we change how they are saved and therefore lose the actual document that was used by decisionmakers to make particular policy decisions that are historic by definition?

But when you calculate that there's a constant stream of images being generated of Mars or the moon or a tropical storm in the Gulf of Mexico or where a gas line is in the city of Alexandria, how do we decide which of those things are worthy of keeping as they are because the very nature of that form is in itself historic just as if it had been drawn on the back of a deer skin?

Mr. CARLIN. I would just say this, Mr. Chairman, that when you move from historically the important documents whether they are written on parchment or paper, whether they are still photos or videotapes, as you get more and more into the technology realm, it changes in terms of preserving that original. In other words, ultimately 100 years from now to look at something that was created digitally today, it won't be essential that you have the equipment of the 2003 to use it to see it.

What you want to see in modern equipment of 100 years from now is what was there. So we won't be preserving that record that is on that first. The information will be preserved, and it will be preserved authentically. That is one of the real challenges, to make sure that you have authentic records over time, but in contrast to that Lewis and Clark map, which we will keep in its original form because it has intrinsic value as well as informational value, although even today we will scan, digitize and put up on a Web site that map to expand access to it, we will keep that original, yes; but when you move, it seems to me, into technology, we're not—the original is important only in the sense that we keep the authentic document, the information, the numbers, the diagrams, the descriptions and not keep a historical presentation itself or the equipment it was presented on.

Now, that is my cut at answering——

Mr. PUTNAM. If I heard you correctly, you said that it is less important to preserve the actual document and more important to preserve the information on that document using the best available technology. I believe that is what I heard you say.

Mr. CARLIN. As a layperson, that is what I communicate, yes.

Mr. PUTNAM. But my point is that the format that it is saved in today, which is the best possible technology today, that actual physical document, the map or the photograph or the thermal imagery of the fires below the rubble in Ground Zero of New York, in and of itself, will be historical because 100 years from now, people will think, well, isn't that quaint and charming that they could rely on something so rude, so crude or so rudimentary compared to what they have 100 years from now? And so the difference between the pictures of Ground Zero or the Pentagon from September 11th and the pictures of Pearl Harbor are the magnitude in terms of volume.

I mean, there's thousands of pictures of New York. There's probably dozens of pictures of Pearl Harbor. So how do you decide which of those things are worthy of keeping as they are because of what they are, in the format that they are in, because they are intrinsically historic?

Mr. CARLIN. Well, Mr. Chairman, to use the Pearl Harbor example, of photographs taken at that time, at some point, that record, that picture will diminish, and we will copy it. So 200 years from now they will not be looking at that original picture. They will be looking at a copy of that record, of that picture. That's been true in terms of still photos, color photographs, videos. The preservation over time is transferring, making sure you have a quality preservation copy, and then as the use of it from researchers diminishes quality or just time does, you make a new copy in the technology of the time. So that will not change. In terms of deciding what we keep that goes back to the scheduling issue and working out with the agency and stakeholders' participation whether or not those images or those records in electronic form should be kept for 50 years, 100 or as we say permanently in the archives, accessioned into the National Archives.

Mr. PUTNAM. Well, my time is long since run out, and I need to yield to my very able Vice Chair. I would just say that I hear what you're saying, but the fact that we have made copies and preserved the content of the declaration of independence doesn't mean that we quit preserving the actual parchment.

Mr. CARLIN. That is correct.

Mr. PUTNAM. And that same drive to preserve will apply to other things. As we go through this hearing, my interest is in finding out the processes used for determining what things are worthy of keeping for posterity.

People may not have thought it was terribly important to keep a list of the procurement orders that Washington ordered for his men for the winter at Valley Forge 225 years ago. Today an archaeologist looks up what they bought and how many buttons and where the buttons came from, and they can now find exactly where the campgrounds were because of what they threw out the back or what fell off their uniform. It probably didn't seem terribly important to save all that back then, but it has tremendous uses today. Now in this information age, in this allegedly paperless society, we're generating a whole lot more stuff, and it's in a format that is more tricky, because of obsolescence, to keep, and so those decisions become more and more important.

I'll yield to the vice chairwoman, Mrs. Miller.

Mrs. MILLER. Thank you, Mr. Chairman. I find your questions so interesting, I think it really does go to the whole heart of this, is what are we trying to archive? What is really necessary? I think that is what we're all struggling with, what are we actually—what is necessary and what is not, I suppose. You know, actually, I think in this case—well, in many cases, but certainly when we think about archiving, let me just—a personal example, first of all. We're thinking about everything that we should be archiving or shouldn't be archiving. I think Members of Congress need to set our own example and our own standards of what we are trying to do as well as we're questioning everybody else and all of the other agencies of what they are doing, and the Members themselves sometimes could be more cognizant, I think, of what is important and what is not.

We've tried to just, within my own office what happens when you—I'm a freshman. When you become a new freshman Member,

you inherit everything from your predecessor. The most up-to-date piece of equipment I had was a 1986 Macintosh, if you can imagine.

So we purchased a new computer system. We have no paper in our office to speak of. We put our whole data base into the thing. Any correspondence that we get from any constituent. We have every registered voter in my district, their name and their address, if they fax us or call us or what have you. We have a paper trail on it. We have a workflow that we can generate their name, and if they were asking us about a Social Security issue or a Veterans' benefit. That has been a helpful tool.

Here we are as a Federal Government telling you to get away from paperwork, and then sometimes we have way too much paperwork. What is important? We are looking for a little assistance from you about what is important.

I would like to ask a question of Ms. Koontz. I thought your testimony was very interesting. You outlined a number of things in regards to NARA that you thought probably could be done better. And I am just wondering, should there be a function with GAO when you are doing your audit process that be part of your auditing process with the various agencies in regards to records management? Is it already a part? And, for instance, if there are some agencies that are doing well, you might want to point that out to other agencies. DOD is apparently ahead of the curve. If there are other agencies that are not doing well at all, would this be identified as a material weakness in their audit? And how can you assist from your perspective these other agencies?

Ms. KOONTZ. I would not say that a look at records management would be a routine part of our information technology audits, but I think that is an interesting idea. And one of the things that we have looked at lately is how to, perhaps, get more recognition of records management in some of the institutional processes that we are trying to get in place across the Federal Government.

I think Mr. Cahoon talked earlier about the Federal Enterprise Architecture. Another way to go is to try to embed records management concerns in the investment process that agencies go through in order to make investment decisions about new IT systems. As they decide to acquire and monitor a system over time, that would certainly be a good way for an agency to be asking over time, how are you handling the records issues? And putting additional emphasis on them. Those are processes that we are trying to get institutionalized in agencies. But I think your idea about making it a more routine part of the audit is interesting.

The other thing I would mention, too, is that this last year we contracted with Grant Thornton to start doing some work that was best practices-oriented on records management across the Federal Government, and the principal investigator on that was Tim Sprehe, and he will be talking on the second panel. So he has developed some, shall we say, preliminary models for records management and identified a number of best practices that I think could be transferred to other agencies.

Mrs. MILLER. I think that is very commendable that you would be doing that with best practices with the private sector. The reason I asked you that question, you also mentioned that there is a

possibility of very important information being lost in those kinds of things. Have you found that? How can you make that kind of a statement? Are you finding these kinds of things very important information that has been lost by the various agencies? Are you finding that during your course of your audits? And if you do find that out, who do you tell about it?

Ms. KOONTZ. From an audit perspective, once it has been lost it is hard to find. And from that perspective, I don't think we have independently documented the loss of the information. What I think we are talking about there is that the risk that information is lost is dramatically increased. Although, I think maybe the NARA representatives might be able to talk about something like this in more detail that I can at this point.

Mrs. MILLER. I am wondering if part of the hesitancy that many of the agencies, and maybe this is a question for Mrs. Riofrio with the Department of Defense, but perhaps much of the hesitancy on behalf of all of the agencies about archiving and that kind of thing, particularly e-mails, you mentioned about e-mails. Now, in this last theater, we saw all of the troops that were using e-mails to communicate with their families, and what a wonderful thing that was. Obviously, you do not want to archive people's personal correspondence. But many people may feel hesitancy about giving up their records of e-mail in the normal course of business in the various agencies, and that I would suspect, makes it more difficult to determine what you are going to archive, as the chairman was mentioning about General Washington and those kinds of things. If he could have sent an e-mail, perhaps he would have done it, and now it makes some sense.

Ms. RIOFRIO. We have been working records management from a very decentralized approach for a long time. And it has not been particularly visible, I think. We concur with GAO's statement. But we have found that the focus, the new focus that we have taken with records management applications really transforms the way people think about electronic records, focusing on what is it that we are going to preserve and what is it that we are not. It seems that the people who are performing the function have the best understanding of what is most important.

But they were not necessarily, at least in the past, thinking about records management. But we were talking about now records management applications, sound like software. "Why don't you go put software in?" But you can't do that. You have to rethink your function. You have to rethink your schedules. You have to plan again. And then everyone in the organization becomes a records manager of sorts, and becomes conscious of records, and starts to make those decisions.

And the records management standard—we were going through this in detail yesterday—requires a great deal of flexibility and decisionmaking on the part of the organization and then of the individuals. And I think that is the beginning of some of the answers that visibility, not just the capability, but the visibility and the understanding.

So we are very excited about the large efforts in the Navy to see how that evolves. Thank you.

Mrs. MILLER. I am interested as you mentioned about the flexibility, and I know Mr. Cahoon mentioned in his remarks about trying to give all the various agencies a great degree of flexibility, and in concept, I think that sounds very good. But as we are trying to get to standards, if you give everybody that much flexibility, you could find yourself in a situation with particular format that is not compatible with proprietary systems. I am sure you are finding that all the time. Should NARA be establishing the standards and defining the standards?

You mentioned during your testimony, I think you mentioned about the GIS. We are all onto this GIS. We had a very interesting hearing a couple of weeks ago about GIS, and all the local municipalities have these wonderful GIS systems, but nobody at the county level can use them, and the State can't use them, and the Federal can't use them. And as we think about homeland security, how does that all trickle down? And in this way, I think, perhaps as well, is the Federal Government responsible for setting some standards as you are archiving different things? How do you work with the States and the other levels of government?

Mr. CAHOON. I think, Vice Chairwoman Miller, I believe that NARA has been very cautious about the formats that it has allowed to be transferred to the permanent archives. In fact, we have been so conservative in that up until recently, we only allowed relational and flat-file formatted data bases that would be readable independent of technology, to be accessioned into the archives and only records that were in ASCII format; those were the only formats that we accepted. And, of course, with that fairly limited set of records, there were a lot of records being created that did not fit neatly into that set of standards.

So as a result, we have opened up the possibilities for additional formats. E-mail with attachments. We have a standard associated with that. Tagged-image file formats for images, and PDF for textual documents can now be sent to us. But those are the only formats currently that are accessionable into the archives. And so we will continue to set standards for these other kinds of records as time goes on. That is part of our major initiative with the Electronic Records Archive.

And we are constantly involved in very collaborative efforts that include the States in thinking about not only standards and the records management applications that have been spoken of, but we collaborate with NASA on the set of standards, particularly the Open Archival Information System Model and the standards associated with it. We do believe that NARA is in a very important position to set standards for records that would be accessioned into the National Archives.

Mr. CARLIN. May I add something to that? You folks helped us a lot under the leadership of Chairman Davis on the e-government legislation because of a couple of aspects of that are going to be very helpful to us in making OMB a real player, and obviously, helping us as one agency working with the entire Federal Government, not only in the structure within OMB that will have a direct responsibility, but the committee that is set up. We will be a direct participant, including the Deputy Archivist who is going to be a major player on that.

I think in terms of setting the standards, it is not just like, yes, we have a responsibility. I also think we have been given additional tools now to convert that into reality.

The other thing I would quickly add, as a former State official as well, I understand very clearly the incredible amount of Federal money that flows to State and local. And whether it is Homeland Security or whatever, if we do not work with State and local to make sure those records are preserved, we will not have, from a Federal point of view, the records to really deal with accountability, whether the program really worked. So there are huge incentives for us to work with State and local. And one of our main avenues, of course, is the National Historic Publications Records Commission, which is our modest grant-providing division that supports and helps State and local governments, and certainly as we look to the coming years, working with them, passing along, allowing them to try some things. You, in the State of Michigan, have done some really great things working with the NHPRC to move us all forward as we share up and down the channels.

Mrs. MILLER. Thank you, I think my time has expired, Mr. Chairman, I appreciate it.

Mr. PUTNAM. Thank you, Mrs. Miller. Ms. Riofrio, how do you handle the issues of classified versus nonclassified documents as part of your electronic records management?

You can answer the question if you would like. We can swear you in real quick if that will save time. Would you like to do that? Could you identify yourself and your position for the record into the mic?

Mr. MATSUURA. Steven Matsuura. I am a senior electronic engineer from the Joint Interoperability Test Command, part of the Defense Information Systems Agency.

Mr. PUTNAM. Very good, please rise and raise your right hand.

[Witness sworn.]

Mr. PUTNAM. The witness responded in the affirmative. If you could please answer the question.

Mr. MATSUURA. First of all, classified and unclassified records have to be segregated. So they are not handled the same way because there is extra security procedures and safeguards that are required for classified records.

Additionally, the standard has additional requirements for information when a classified record is declared as a record. In addition to the normal information that a user has to input to document the records, he has to add information such as the reason for classification, what the declassification schedule is, and, of course, what level of classification it is.

Mr. PUTNAM. We collect, in Mr. Rumsfeld's terms, an "unknowable amount" of information from around the world through various sources and methods. It is then filtered and screened through a variety of appropriate personnel and agencies as part of our intelligence briefings and intelligence plans. How would someone know, and at what point in that chain does someone know, that they should preserve either the raw data or the interpreted data or the refined data? At what point does that become a record and subject to the process as a classified document that you just outlined?

Mr. MATSUURA. Well, it is somewhat subjective, but as mentioned before, each agency has to have a schedule, and if a record falls into a certain category or series, he knows that—or a document or a piece of information—he knows that it needs to be saved as an official record.

Mr. PUTNAM. Are the schedules designed by each individual agency, or are the schedules designed by NARA and handed down to the agency?

Mr. MATSUURA. Well, NARA, of course, develops the General Records Schedule, which agencies use if they are applicable or where applicable. But actually, agencies propose schedules, depending on what their business line is, and submit them to NARA for approval and that gives them the disposition authority.

Mr. PUTNAM. Ms. Koontz, would you like to comment on this whole process? There seems to be a general difficulty in describing to the subcommittee how the decision is made, what records are worthy of archiving.

Ms. KOONTZ. Well, I am not sure I can elaborate a whole lot. NARA has a whole process that they have used over many years where they use experts to appraise—that is, you know, look at the value of a record and determine whether that record could possibly be a permanent record or not. They are certainly more expert in that process than I am, but they do have a very longstanding process to decide which records will be kept and which records can be disposed of. That is what the whole of records management is, knowing what to keep and to get rid of the rest.

Mr. PUTNAM. But in your testimony, you said that their architecture for archiving electronic records has raised some concerns about costs and progress on that system.

Ms. KOONTZ. Right, the electronic record archiving.

Mr. PUTNAM. On the electronic side.

Ms. KOONTZ. Right. We are talking there about the development of their electronic archival system, which does what Mr. Carlin and Cahoon talked about preserving electronic records independent of software, so that you don't have to worry about software obsolescence, you can maintain the content over time, and part of the strategy is to migrate the media over time as well.

Mr. PUTNAM. And as all agencies move forward purchasing new IT infrastructure, and we spend $60 billion a year on that, does OMB require a component of that enterprise architecture to be electronic records management? Is that part of it?

Ms. KOONTZ. I think Mr. Cahoon talked with this earlier, about the relationship of records management to one of the component service areas in the Federal Enterprise Architecture. I know less about it than he does. He is working on that particular aspect.

Mr. PUTNAM. Well, we have had a lot of hearings in the subcommittee this year on the inadequacy of our IT procurement and acquisition programs, the lack of information security, the fact that even within a department, especially DOD, there is an awful lot of different rabbit trails that are being pursued with regard to IT. So it would surprise me if they all had their act together on electronic records management? Is that the case or not?

Mr. CAHOON. You needn't be surprised, Mr. Chairman. I would have to say that we are all working to get our act together. The

Federal Enterprise Architecture specified in four different places records management as a key component of the business reference model of that architecture. In the service component architecture, there are identified a set of service types, and digital asset management is one of those service types. And within that service type is electronic records management identified specifically as a component in the Federal Enterprise Architecture. And that electronic records management then has a series of elements to it that, over time, when implemented, will provide not only guidance but specific software, component software, that can be used by agency applications that will help us move a long way toward standardizing how records are managed, how they are described, the kind of information that is available about those records how those records relate to one another, how they form collections, information about their disposition, how long they could be kept and for what purposes they should be kept. All of that is part of this architecture.

And for me, that provides a sustainable, long-term way of bringing electronic records management into the forefront of information technology application development, because the components will be there, available for people to use and to integrate into their applications rather than having to figure out everything for themselves. So our requirements and standards, Mr. Chairman, can be built into those components themselves.

Mr. CARLIN. Mr. Chairman, could I add some, because you have a question that still has not been answered yet, and I would like to try one more time.

Mr. PUTNAM. OK.

Mr. CARLIN. First of all there is a basic fundamental that the format itself does not impact whether a record is kept 1 year, 2 years or permanently, whether it is on paper or electronic form. So when we shift to more electronic, that does not suddenly force a whole new set of schedules, generally. Now, obviously, there are some individual aspects that require some changes, and we are creating records electronically that we did not even have before in the paper form, but the format is not a factor. When we work to schedule records, to make those determinations, we want to document the decision and what the decision is will determine to a great extent how long it's kept.

Reference was made to General Records Schedules by our friend from the Department of Defense. Those are the schedules that are across the board for every agency, for them to use efficiently to deal with very short-time, important for the moment, but certainly not important in protecting rights and entitlements, not important for accountability, certainly not important for telling the national experience, but just for that particular business transaction for a moment. Those are General Records Schedules that deal across the board, that will apply as well electronically.

And so it's not the media that is the factor that necessarily determines how long a record will be kept. Because we are now electronic and have so much more volume, does not mean that we will go from 2 or 3 percent accessioned in the archives to 50 percent. It may go to 5 percent because of the ease with which dealt with it and our capacity to provide access to those records efficiently. But we don't want to cloud the whole electronic storecase with end-

less records that are not needed. If they were not needed in paper, probably they aren't needed electronically any longer than they were when they were a paper format.

Mr. PUTNAM. That is a fair point. My concern is as much as anything on the map, GIS, geospatial photography side of things, which are much more voluminous than they were in the past, and frankly, easier to store because you can put all of this stuff on a CD-ROM or whatever. But the other issue and as the Technology Chairman, I am probably going to say something I shouldn't, but I will make a confession, I was on Air Force 1 on September 11th. And until we shut down all of our electronic devices, I sent out e-mails to my wife for one. Well, I made a point of saving those e-mails, but I also printed them out because I did not trust the computer. One day I would be, you know, I would hit the wrong delete button. I just did not trust it, so I printed them out and saved them.

I suspect you all probably do the same thing. That, you know, as good as your electronic capacity is, if there is something really special in there, you probably drop back to the good old-fashioned paper form, which is what all of us do, which is why this paperless society cuts down more trees than we ever have.

And so the root of my questions about what is worthy of being saved and all of that, and how do we save it as the technology changes so that you don't have to keep a TRS8 to be able to read this 10 years worth of documents, and a 286 to read this, and a Pentium to read this decade's worth of documents. You know, at the end of the day, as that technology changes so quickly, we have to be very careful about how we preserve these things. And so that is my concern as an amateur history buff, that is my issue in trying to decide how you guys decide what to save.

And, unquestionably, you are the best in the world at it. I mean, I don't propose to tell you how to decide what to save. But if you are just the systems administrator in the field office of the Farm Service Agency somewhere in Kansas, and you are trying to decide what is worthy of saving, I am trying to figure out before we go cracking heads who gives them some direction, some clear-cut policy on what they should keep and what they can throw out the back-door. And I have taken this hearing—the staff is going nuts, I have taken this way out into the historical side and not in the enterprise architecture, which is where we are supposed to be.

I recognize Mrs. Miller for another round of questions, and then we will move on.

Mrs. MILLER. In the interest of time, I will pass. Thank you, Mr. Chairman.

Mr. PUTNAM. One last question. Is there—and this gets to what I was saying about I printed out my e-mails—is there, say, a backup nondigital system for the storage of vital information should that information be lost?

Mr. CARLIN. Mr. Chairman, what you described previously is very real. Until we get records management applications to the quality that we are comfortable, a lot of agencies today, and with our encouragement, are printing out to paper. The record copy is paper, not the digital form that was originally created. So the practical issue you addressed is one that agencies do address today.

That decision out there in Kansas that you made reference to, first of all, a schedule is going to determine and that schedule will say whether the electronic or the paper copy is the record copy. And probably out there, I would guess, it would be paper today because we are still at the early stages of developing—when I say we, I mean globally, we are testing and learning, but the private sector, in producing, has made a lot of progress. And we may have one there today, but is it available or been tested to the extent that everybody across the board, we are saying go forward and use it. So your experience is one that has been experienced by agencies all across the Federal Government.

Mr. PUTNAM. Very good. Would any of the witnesses wish to add anything? Something that we did not ask about? Something, if you would like to clarify anything? This is your opportunity before we seat the second panel. Yes, sir?

Mr. CAHOON. Mr. Chairman, I would like to add, with respect to this, your critical question. NARA looks very carefully at the business processes that each agency undertakes to do their business. And, based on an analysis of those business processes, determines what records are necessary to document the decisions and the transactions that business process creates.

Based on that analysis, a record schedule is developed. That schedule is then reworked and evaluated and tuned up and then presented to Mr. Carlin for his signature. And it's based on that understanding of the business process and what is necessary to document the rights and entitlements of individuals, the decisions and actions of Federal officials and what would be important to know for the national experience, all become factors in the decision of what we should keep.

Risk of those records being lost, and the value of those records to the business of the agency all factor in to the decision that is made as to what we ought to keep and what can be disposed of after the normal course of business. Thank you.

Mr. PUTNAM. Very good, anyone else?

Mr. CARLIN. Mr. Chairman, I would only add that we appreciate very much this opportunity. We appreciate your leadership, along with your vice chair and your interest. And that we would hope that this would be an ongoing communication over time, because we realize you need to know what we are doing. You need to have your questions answered. And as I said to you privately, we appreciate the relationship with GAO and their active involvement in what we are doing in all aspects of electronic records, and we thank you very, very much for this opportunity today.

Mr. PUTNAM. You are welcome, and we look forward to seeing your facility next week. The record will remain open for us to submit further questions that were not addressed in the hearing, and we would ask your cooperation in responding to those. At this time, we will excuse the first panel and seat the second. So the committee will go in recess for a minute and a half.

Thank you.

[Recess.]

Mr. PUTNAM. The subcommittee will reconvene. We want to welcome the second panel. And as was the custom with this sub-

committee and as you saw with the first panel, I would ask you to rise and raise your right hands, please, for the oath.

[Witnesses sworn.]

Mr. PUTNAM. Note for the record that all of the witnesses responded in the affirmative.

We will immediately go to witness testimony.

Our second panel begins with Dr. J. Timothy Sprehe. Dr. Sprehe is president of Sprehe Information Management Associates, a firm offering consulting services and information resources management to government agencies and private firms since 1991. He specializes in issues such as strategic planning for information resources management, public access to government information, electronic records management and electronic collection and dissemination of information.

He has conducted consulting studies for many U.S. Federal agencies and private firms doing business with the Federal Government. He retired from OMB in 1991, where at OMB he was the principal author of the original 1985 OMB Circular, No. A–130, the Management of Federal Information Resources. This government-wide information policy directive established comprehensive policy on managing information and managing information systems and technology.

Mr. Sprehe received an M.A. in 1963 and an Ph.D. in 1967 in sociology from Washington University in St. Louis. We welcome you to the subcommittee, Dr. Sprehe, and look forward to your testimony. You are recognized for 5 minutes.

STATEMENTS OF J. TIMOTHY SPREHE, PRESIDENT, SPREHE INFORMATION MANAGEMENT ASSOCIATES; ROBERT F. NAWROCKI, CRM, DIRECTOR, RECORDS MANAGEMENT AND IMAGINING SERVICES DIVISION, LIBRARY OF VIRGINIA; CARYN WOJCIK, STATE GOVERNMENT RECORDS MANAGEMENT, MICHIGAN; AND DR. RICHARD LYUSAKOWSKI, DIRECTOR, COLLABORATIVE ELECTRONIC NOTEBOOK SYSTEMS ASSOCIATION [CENSA]

Mr. SPREHE. Thank you, Mr. Chairman. It is a pleasure to appear before this subcommittee to discuss the subject of electronic records management [ERM].

The National Archives' 2001 survey on current Federal recordkeeping practices, of which I was a coauthor, stated that the chief paradox of today's Federal records management is the disconnect between paper and electronic recordkeeping. Many agencies do competent paper records management; only a handful do competent electronic records management.

One reason for this condition is the cultural chasm between the records management community and the information technology community. Generally, records managers do not understand IT, and IT managers do not understand records management.

A Federal record is not just something saved by a computer. Rather, a record is maintained as evidence in pursuance of legal obligations or in the transaction of business. It's something you can take into court where you must be able to prove that the record has authenticity, reliability, integrity, and usability. IT professionals seldom understand this meaning of "record." It's common to say

today that most records are born electronic. Where do they live and where do they grow up? In IT systems. And nobody is talking to the IT people about records.

A second reason that ERM is sadly wanting in Federal records is because the National Archives has been painfully slow to address ERM, let alone get out in front of it. Bold out-in-front policy guidance on ERM comes slowly in NARA's culture of archivists and historians, compounded by the agency's extreme, and perhaps well justified, fear of adverse litigation. You might note that for all of its commendable initiatives in the electronics record management area, NARA itself does not use electronics records management.

A third reason for the slow spread of ERM is the obstacles the agencies face, namely, lack of funding for ERM and indifference on the part of senior management. Too many agency heads and chief information officers consider ERM a back-burner low priority. They believe their top priorities are more immediate and include items such as security and risk management, not appreciating that ERM improves security and lowers risk.

Thus, as the volume of electronic records multiplies exponentially in the agencies, progress in ERM creeps along inch by inch. Disasters such as happened at the Bureau of Indian Affairs and the Federal Bureau of Investigation are the tip of a large iceberg. Many more records management disasters are out there waiting to happen.

On the bright side, a few agencies have achieved significant advances in making ERM an integral operating component in their information architectures. Agencies such as the Nuclear Regulatory Commission, the Office of Civilian Radioactive Waste Management—otherwise known as the Yucca Mountain Project—in the Department of Energy, and the Office of Comptroller of the Currency in the Treasury Department. These are agencies that take records very seriously by the nature of their missions, agencies where top management has issued a mandate that ERM shall exist enterprise-wide.

Agencies reaching full ERM implementation discover that an electronic records repository is an asset with many beneficial applications beyond records management. ERM gives the agencies instant access to institutional memory starting yesterday. They can leverage this asset to provide economies and efficiencies to other business functions. For example, the Nuclear Regulatory Commission's ADAMS, which stands for agencywide document access and management system, is a combined electronic documents and records management system. NRC has made ADAMS a core component of its public Web site, so that ADAMS not only performs records management, it also performs other important public information functions.

Even more, NRC has coupled ADAMS with electronic signature capability, so that ADAMS can receive electronic submissions from nuclear reactor licensees in the general public, and, hence, NRC has leveraged its ERM investment to carry out not only its records management responsibilities, but also the Government Paperwork Elimination Act and improve its performance under the Government Paperwork Reduction Act.

My single recommendation regarding Federal ERM is this: The Office of Management and Budget should change exhibit 300 on capital planning and budgeting in its annual budget directive, Circular No. A–11. The section of exhibit 300 that deals with IT systems should state that no new IT system will receive funding for development and acquisition unless the justification for the new system adequately explains how the system will provide for records created by or passing through the system. NRC and the Yucca Mountain Project are already implementing such a policy.

We can only hope that the many agencies planning today for ERM will receive the funding and leadership they deserve. From my research and consulting I have come to believe that ERM is the bedrock of what is known today as enterprise content management or what we used to call information resources management. Without ERM, enterprise content management is incomplete and hollow at its core. Thank you for inviting me to testify. And I would be happy to answer any questions you have.

Mr. PUTNAM. Thank you very much, Dr. Sprehe. I wish the first panel had hung around.

[The prepared statement of Mr. Sprehe follows:]

WRITTEN STATEMENT SUBMITTED BY J. TIMOTHY SPREHE BEFORE THE HOUSE SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY, INTERGOVERNMENTAL RELATIONS, AND THE CENSUS ON THE SUBJECT OF ELECTRONIC RECORDS MANAGEMENT JULY 8, 2003

My name is J. Timothy Sprehe and I am president of Sprehe Information Management Associates, a consulting firm based in Washington, DC, and in recent years specializing in electronic records management. It is a pleasure to appear before this Subcommittee to discuss the subject of electronic records management or ERM.

1. Lack of Electronic Recordkeeping. The National Archives' 2001 report on current federal recordkeeping practices stated:

> **Government employees do not know how to solve the problem of electronic records – whether the electronic information they create constitutes records and, if so, what to do with the records.** Electronic files that qualify as records—particularly in the form of e-mail, and also word processing and spreadsheet documents—are not being kept at all as records in many cases and are frequently not being scheduled. Employees lack guidance and knowledge concerning how to identify electronic records and what to do with them once identified. Technology tools for managing electronic records do not exist in most agencies. The agency information technology environments have not been designed to facilitate the retention and retrieval of electronic records. Despite the growth of electronic media, agency records systems are predominately in paper format rather than electronic. Virtually every agency visited indicated that the official policy is that their records will be maintained in paper format. Yet the agencies recognize that most records are now created in an electronic environment—in word-processing documents, spreadsheets, databases, and the like. The predominant e-mail policy is to print out e-mails that are considered records and to save the paper copies. The chief paradox of today's Federal RM is the disconnect between paper and electronic recordkeeping.

2. Records and the Cultural Chasm. One reason for this condition is the cultural chasm between the records management community and the information technology community. Generally, records managers do not understand IT and IT managers do not understand records management.

2.1. Records within records management culture. Within U.S. federal law, the term "record" is defined as:

> All books, papers, maps, photographs, machine readable materials, or other documentary materials, regardless of physical form or characteristics, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations, or other activities of the Government or because of the informational value of data in them. (44 U.S. Code Sec. 3301).

For purposes reaching beyond the U.S. Government, a simpler and more general definition of records is found in ISO 15489-1, the international standard for RM issued by the International Organization for Standardization:

> **Record**: Information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business.[1]

Within this meaning, the term "record" denotes information to which special characteristics adhere.

- **Authenticity**. An authentic record is one that can be proven
  - a) To be what it purports to be,
  - b) To have been created or sent by the person purported to have created or sent it, and

[1] International Organization for Standardization, *ISO 15489-1, Information and Documentation -- Records Management: Part 1: General,* First Edition, September 15, 2001, Geneva, Switzerland, p. 3. Recently, NARA has announced its intention to move toward the ISO standard. In its Proposal for a Redesign of Federal Records Management, July 2002 (http://www.archives.gov/records_management/initiatives/rm_redesign.html), NARA states that in its redesign of RM guidance and training NARA would: "incorporate the concepts in the ISO records management standard 15489 to the extent possible under Federal law. Specifically, incorporate into our theoretical framework the idea of trustworthy records that have the qualities of authenticity, reliability, integrity, and usability over time sufficient to meet the needs of stakeholders."

c) To have been created or sent at the time purported.

- **Reliability.** A reliable record is one whose contents can be trusted as a full and accurate representation of the transactions, activities or facts to which they attest and can be depended upon in the course of subsequent transactions or activities.
- **Integrity**. The integrity of a record refers to its being complete and unaltered.
- **Usability**. A useable record is one that can be located, retrieved, presented and interpreted.[2]

2.2. Records within information technology culture. Within the realm of computers and IT, "record" has a very different meaning.

In database management systems, a record is a complete set of information. Records are composed of *fields*, each of which contains one item of information. A set of records constitutes a *file*. For example, a personnel file might contain records that have three fields: a name field, an address field, and a phone number field.[3]

By way of explanation, a database management system is a collection of computer programs that enables one to store, modify, and extract information from a database; a database is an organized collection of information such as an electronic filing system. A field is the space allowed for a particular item of information, the smallest unit of information the user can access within the database.[4] The above meaning of "record" is a relational concept in the sense that the term has no meaning except in relation to other terms.

In the area of document technologies, "record" has other meanings, some related to the above and some not related.

> **Record** -- (1) Group of one or more words containing information about a common subject. One or more records make up a file. (2) To copy or set down information in some form for future use. (3) Any information that is stored by any device. (4) Number of fields that relate to a single item within the database. (5) In workflow, collection of individual items of data treated as a unit. Refers to items in a database. Each item is represented by a record that consists of one or more fields.[5]

For personnel trained in IT specialties, the above series of definitions is most likely to be the meaning they assign to the term "record."

2.3. Conclusion. The IT meaning of "record" bears little resemblance to the RM meaning of "record." The IT record carries none of the evidentiary, legal, and transactional connotations associated with the RM record. To say that a given piece of information is an IT record is to say nothing about its authenticity, reliability, integrity, and usability, defining characteristics of a record. Differences in the meaning assigned to the same term by various disciplinary specializations represent cultural and linguistic barriers to common understanding. RM and IT professionals have difficulty understanding one another when using such common terms such as "record."[6]

3. NARA Has the Institutional "Slows." A second reason that ERM is sadly wanting in federal agencies is because the National Archives and Records Administration has been painfully slow to address ERM, let alone get out in front of it. Bold, out-in-front policy guidance on ERM comes very slowly in NARA's culture of archivists and historians, compounded by the agency's extreme – and perhaps well justified -- fear of adverse litigation.

One among many examples of NARA's institutional slowness is the subject of websites as records. In 1998, a study sponsored by the National Historical Publications and Records Commission dealt with the subject of guidance

---

[2] *Ibid.*, p. 7.

[3] Source: *Webopedia*, URL http://www.webopedia.com/TERM/r/record.html (December 2, 2002). For other definitions see, for example, the High-Tech Dictionary at URL http://www.computeruser.com/resources/dictionary/, or TechWeb at URL http://www.techweb.com/encyclopedia/, all as of December 2, 2002.

[4] *Ibid*, following the hyperlinks found therein for "database," "database management system," and "field."

[5] Association for Information and Image Management International, *Glossary of Document Technologies*, an ANSI Technical Report prepared by AIIM International, ANSI/AIIM TR2-1998, p. 76. "ABSTRACT: This technical report provides definitions for the terms used in the micrographic, electronic imaging, and workflow fields. This technical report also includes telecommunications/Internet terms related to these fields." From the title page.

[6] "Archive" is another term that carries very different meanings in RM and IT. To the records manager, the word "archive" is a noun only, not a verb, and denotes an agency, place, or set of records intended for permanent long term preservation. To the IT professional, "archive" is both a noun and a verb and refers to the copying of files to a storage medium for backup of indeterminate but certainly not permanent duration.

for ERM on state and federal agency websites.[7] NARA posted a hyperlink to the study on its website, but not until January 2002, four years later, did NARA circulate a **draft** of its own guidance for website records management. Today, eighteen months later, NARA's guidance has still not been issued publicly. In fairness to NARA, we should note that the Office of Management and Budget, which enjoys but a superficial understanding of records management in general and ERM in particular, acts as a constant brake on NARA initiatives. Regardless of which political party is in power, OMB is consistently very conservative about records management policy, seeing always large dollar signs as the budgetary consequences of new records management policy initiatives.

Another major example of NARA's institutional "slows" is the Electronic Records Archives program.[8] Although NARA is now putting major resources into this program for the long-term preservation of digital information, it is my opinion, shared by many others in the field, that NARA is at least twenty to twenty-five years later in seriously addressing this issue. For decades, NARA's Center for Electronic Records has limped along on inadequate funding and low management priority.

4. Lack of Funding for ERM. A third reason for the slow spread of ERM is obstacles the agencies face, namely lack of funding for ERM and indifference on the part of senior management. Too many agency heads and Chief Information Officers consider ERM a back-burner low priority. They believe their top priorities are more immediate and include such items as security and risk management, not appreciating that ERM improves security and lowers risks.[9]

Thus, as the volume of electronic records multiplies exponentially in the agencies, progress in ERM creeps along inch by inch. Records management disasters such as happened at the Bureau of Indian Affairs and the Federal Bureau of Investigation are the tip of a large iceberg; many more records management disasters are out there waiting to happen.

5. The Plus Side of ERM. On the bright side, a few agencies have achieved significant advances in making ERM an integral operating component in their information architectures. The most notable examples are the Nuclear Regulatory Commission, the Office of Civilian Radioactive Waste Management, or Yucca Mountain Project, in the Department of Energy, and the Office of the Comptroller of the Currency in the Treasury Department. These are agencies that take records very seriously by the nature of their missions, agencies where top management has issued a mandate that ERM shall occur enterprise-wide.

5.1 Leveraging ERM Investment to Benefit Other Business Functions. Agencies reaching full ERM implementation discover that an electronic records repository is an asset with many beneficial applications *beyond* records management. ERM gives the agencies *instant access to institutional memory starting yesterday.* They can leverage this asset to provide economies and efficiencies to other business functions.

For example, the Nuclear Regulatory Commission's ADAMS, which stands for Agency-wide Document Access and Management System, is a combined electronic document and records management system.[10] NRC has made ADAMS a core component of its public website so that ADAMS not only performs records management functions but also provides important public information functions.

Even more, NRC has coupled ADAMS with electronic signature capability so that ADAMS can receive electronic submissions from nuclear reactor licensees. Hence, NRC also has leveraged its ERM investment to carry out the Government Paperwork Elimination Act and improve its performance under the Paperwork Reduction Act.

5.2 More Limited Example. ADAMS is an enterprise-wide solution that NRC has leveraged significantly. Other agencies have been able to multiply their investment dollars on even more limited applications. The Defense Advanced Research Projects Agency has a narrowly focused ERM application in its Comptroller's

---

[7] Charles R. McClure and J. Timothy Sprehe, *Guidelines for Electronic Records Management on State and Federal Agency Websites*, Report of a research study sponsored by National Historical Publications and Records Commission, January 1998. http://www.jtsprehe.com/newpage12.htm

[8] See http://www.archives.gov/electronic_records_archives/index.html.

[9] In a reorganization several years ago that resulted eventually in the Office of Environmental Information, the Administrator's office in the Environmental Protection Agency decided to abolish records management on the grounds it was an unnecessary function. When lower management protested loudly, the Administrator's office then decided to place records management in a unit with facilities maintenance and physical security (or "cops and mops," as the agency wags put it). Eventually, wiser heads prevailed but the episode is indicative of the abyss of ignorance from which many top management officials view records management. EPA, it should be noted, is an agency that in recent years has had one of the federal government's very best records management programs, motivated in substantial part by the fact that the agency experiences a high volume of litigation and Freedom of Information Act requests.

[10] See http://www.nrc.gov/reading-rm/adams.html.

Office. The application automatically copies all "ARPA Orders," a key research funds disbursement document, into an ERM system. DARPA's Executive Information System draws data from numerous agency sources, including the electronic records repository, to keep top management abreast of agency performance on a current basis. Note that the records management function is **not** the same as executive information functions, yet DARPA has arranged to have the former support the latter in an efficient and effective manner.

6. ERM Recommendation. My single recommendation regarding federal ERM is this.

> The Office of Management and Budget should change Exhibit 300 on Capital Planning and Budgeting in its annual budget directive, Circular No. A-11. The section of Exhibit 300 that deals with IT systems should state that no new IT system will receive funding for development and acquisition unless the justification for the new system adequately explains how the system will provide for records created by or passing through the system.

NRC and the Yucca Mountain Project are already implementing such a policy.

Exhibit 300 already mentions electronic records in conjunction with the Government Paperwork Elimination Act. But GPEA affects only a handful of the tens of thousands of federal IT systems, namely, those systems that pertain to the collection of information from the public. The Department of Defense, for example, collects very little information from the public and yet has thousands of IT systems that create and/or process federal records.

All agencies have records management policies and many have electronic records management and email records management policies. The problem is that the agencies establish no linkage between ERM policies and IT system design, acquisition, and management policies. Having an ERM policy is meaningless unless the policy is linked to policies governing the IT systems in which the electronic records reside.

7. Conclusion. We can only hope that the many agencies planning today for ERM will receive the funding and leadership they deserve. From my research and consulting, I have come to believe that ERM is the bedrock of what is known today as Enterprise Content Management. Without ERM, Enterprise Content Management is incomplete and hollow at its core.

Thank you for inviting me to testify.

J. Timothy Sprehe
Sprehe Information Management Associates
4201 Connecticut Ave., N.W., Suite 610
Washington, DC 2008
Telephone: 2202-364-5300
Fax: 202-364-5302
E-mail: jtsprehe@jtsprehe.com

Mr. PUTNAM. Our next witness is Mr. Bob Nawrocki.

Mr. Nawrocki is the director of the Records and Imaging Service Division in the Library of Virginia. Previously he was the electronic records coordinator for the library. He is responsible for dealing with electronic records issues facing the State of Virginia. Previously he spent 2 years as a contractor with the Air Force working on their integrated digital environment.

His experience involves working for State and Federal Governments, as well as for several law firms and private industry. Mr. Nawrocki is a Certified Records Manager and has a Masters in Science in Library and Information Sciences from the Catholic University of America. He is an adjunct faculty member with Catholic University teaching Information Systems for Libraries and Information Centers and Organization of Information.

Welcome to the subcommittee. You are recognized.

Mr. NAWROCKI. Thank you, Mr. Chairman. I will briefly be speaking today about the Library of Virginia and what we are doing here in Virginia dealing with electronic records.

The Library of Virginia has responsibility for the Commonwealth's records management and archival programs. As such, we work closely with State and local agencies to assist them with the management of their records. Over the past 5 years, electronic records have become a significant issue. In 2000, the Library was able to hire their first electronic records coordinator to develop polices, guidance, and provide advice on the management of electronic records.

The Library faces many of the same problems that NARA does, even though they are on a much smaller scale. In the 1990's, a number of Virginia counties began to use scanning and electronic storage for the recording of deeds. While this reduced the amount space required and improved access, it did not provide the long-term stability and preservation required of such records. In fact, we had a county that actually lost about 30 or 40 deeds when they had a computer crash. Luckily, they were able to recover those over a short period of time.

Changes to the Code of Virginia have made digital images legal substitutes or replacements for original documents. The Library wants to encourage the use of microfilm as a backup to digital images, since there are no recognized, permanent media in the digital arena. When stored properly, microfilm can last up to 500 years. Through a philanthropist, the Library was able to obtain a Kodak Archive Writer to convert digital images to microfilm. We use an encrypted Internet connection for scanned images to be sent to the Library's State Records Center, where they are stored on a RAID device until data would be written to microfilm.

The microfilm was processed under a quality control and the resultant negatives were stored in our media vault. And this project continued into the fall of 2002 when most of the staff was laid off due to extreme budget cuts. By that time, private vendors were able to step in using the same technology to do the same actions at the same or lower costs.

Over the past 3 years, the Library has worked to educate State and local agencies about electronic records and the best methods to preserve them. Guidelines for transferring electronic records into

the Library are under development. Our greatest concern is that in the short term we will lose access to these records before a reliable method of archiving them is developed. We have found that the single most important tool for raising awareness, though, is education through presentations and articles.

Second, we believe that the development of guidelines that provide advice and suggested tools for electronic records creators are also useful. We also find it is important not to develop guidelines and tools that are too narrowly drawn since the same solution is not always applicable in all cases.

Rapid technological change results in records only a few years old being unrecoverable because the hardware or software does not exist. While there is significant, extensive research being done long-term on electronic preservation, we need to remember that we have to manage these records in the short term, so that when the proverbial magic bullet is created, there will be older electronic records to be preserved.

The current effort to create a Portable Document Format-Archival is an excellent example of this thinking. The use of hybrid analog/digital technologies is another. We need more collaborations like this which utilize existing technology and software to provide records managers, archivists and librarians with the tools needed to manage today's electronic records. We also need to continue to use the tools, both analog and digital, which allow us to provide access to existing electronic records and preserve them at the same time.

There is much work being done around the world in the field of digital preservation and sharing of information. This is necessary and must continue if we are to solve this intractable problem. Thank you.

[The prepared statement of Mr. Nawrocki follows:]

Statement of Robert Nawrocki

before the

Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census

Good morning. My name is Robert F. Nawrocki and I am currently the Director of the Records and Imaging Services Division of the Library of Virginia. I have been in records and information management for over 30 years. I am a certified records manager and a past president of ARMA International.

The Library of Virginia has responsibility for the Commonwealth's records management and archival programs. As such we work closely with state and local agencies to assist them with the management of their records. Electronic records have become a significant issue in the last 5 years. In 2000 the Library hired its first Electronic Records Coordinator to develop policies, guidance and to provide advice on the management of electronic records. I was fortunate to have filled that position from 2000 until April 25th when I become Director of the Division.

In the late 1990s a number of counties began to use scanning and electronic storage for the recording of deeds. While this reduced the amount of space required and improved access it did not provide the long term stability and preservation required of such records.

Changes to the Code of Virginia made digital images legal substitutes or replacements for original documents. The Library wanted to encourage the use of microfilm as a backup to digital images since there are no recognized permanent media in the digital arena. When stored properly microfilm can last almost 500 years. Through a philanthropist LVA obtain a Kodak Archive Writer to convert digital images into microfilm. Using an encrypted Internet connection scanned images were sent to the Library's State Records Center where they were stored on a RAID device until the data could be written to microfilm. The microfilm was then processed, underwent quality control and the resultant silver negatives stored in the State Records Center's media vault. This

project continued until the fall of 2002 when most of the microfilm staff was laid off due to extreme budget cuts. Private vendors will now provide the same conversion that the library did at about the same cost. The microfilm negatives are stilled stored in the records center's media vault. We consider that the project was valuable since it showed the benefit of a hybrid system which allowed counties to provide faster, distributed access with electronic records while preserving those records, at a lower cost, using a proven technology, microfilm.

Over the past three years the Library has worked to educate state and local agencies about electronic records and the best methods to preserve them. Guidelines for transferring electronic records into the archives are under development. Our greatest concern is that in the short term we will lose access to these records before a reliable method of archiving them is developed. We have found that the single most important tool for raising awareness is education through presentations and articles in state publications. Secondly we believe the development of guidelines that provide advice and suggested tools for electronic records creators is also useful. It is important to not develop guidelines and tools that are too narrowly drawn since the same solution is not applicable in all situations. Flexibility is the key to as long as the proper result is achieved. Additionally records managers and archivists must work closely and cooperatively with their Information Technology associates. We need to collaborate with them when new information systems and the required information architectures are being developed to insure that the appropriate records management functions are included. We have found that by working with IT and making them aware of records management and archival needs and requirements enables everyone to perform their jobs better.

. It is important to remember that during the entire process of creating and preserving electronic records we must continue to preserve the evidentiary aspect of the electronic record. This cannot be done unless the appropriate audit trails and safeguards have been built into the information architecture. If the electronic records are not shown to be reliable and have integrity their evidentiary value will be zero and the work to preserve them will be wasted. This can only be

done through the collaboration and work of all parties involved in the creation, use, maintenance and preservation of the electronic records.

There is much hard work to be done now and in the future. This work can be only done through the collaboration and involvement of all whether they are in the federal, state or local governments or private industry.

Electronic records present a serious problem for everyone. Rapid technological change results in records only a few years old being unrecoverable because the hardware or software needed does not exist. While there is significant and expensive research being done long-term on electronic preservation we need to remember that we have to manage these records in the short-term so that when the proverbial "magic bullet' is created there will be older electronic records to be preserved. The current effort to create a Portable Document Format-Archival is an excellent example of this thinking. The use of hybrid analog/digital systems is another. We need more collaborations like this which utilize existing technology and software to provide records managers, archivists and librarians with the tools needed to manage today's electronic records. We also need to continue to use the tools, both analog and digital, which allow us to provide access to existing electronic records and preserve them at the same time. There is much work being done around the world in the field of digital preservation and sharing of information. This is necessary and must continue if we are to solve this seemingly intractable problem.

Mr. PUTNAM. At this time, I would like to turn the gavel over to our distinguished vice chair and native of Michigan, former Secretary of State, Mrs. Miller, for the next introduction.

Mrs. MILLER [presiding]. Thank you, Mr. Chairman. It's my pleasure to do so, to introduce Karen Wojcik. She has worked for the State Archives of Michigan since 1996. She is State Archives' primary liaison with the executive branch, the legislature, the Supreme Court and our Court of Appeals. She is responsible for appraising public records for their historical value and for developing electronic records management and preservation strategy.

Mrs. Wojcik received both her bachelor's degree in history and her master's degree in information and library science from the University of Michigan, Ann Arbor. And we wonder who she roots for on football days between the Michigan State and the University. She is a Certified Archivist. She has participated in several grant projects funded by the National Historical Publications and Records Commission, including the Michigan RMA Pilot Project, the PERM Project, and she served on the advisory board for the SDSC Archivists' Workbench Project. She is the Chair of the National Association of Government Archives and Records Administrators, on the Committee on Electronic Records and Information Systems, and has been nominated to serve as secretary on this board as well, and we are certainly pleased to have her here as well.

Ms. WOJCIK. Thank you. Mr. Chairman, thank you for inviting the Michigan Department of History, Arts, and Libraries to testify at this hearing. I am honored to be speaking to you today about the management and preservation of our government's records, particularly electronic records. And, of course, I especially want to thank Congresswoman Candice Miller for inviting us to be here today.

The Michigan Department of History, Arts, and Libraries is committed to enriching the quality of life for Michigan residents by providing access to information, preserving and promoting Michigan heritage, and fostering cultural creativity. The Division of Archives and Records Management within the Michigan Historical Center is responsible for ensuring that Michigan's government records are properly managed and preserved throughout their entire lifecycle.

Public records are essential for protecting our legal rights, for documenting the actions of our government, and for understanding the society in which we live, and, therefore, it cannot be sufficient to have our history preserved by accident.

As a general rule, we estimate that less than 5 percent of all government records possess permanent historical value. Professional archivists and records managers have worked for years to develop systematic processes that identify which records possess historical value and to provide for their preservation. Unfortunately, many government agencies do not follow these procedures, and they fail to protect the irreplaceable items in their custody.

In the traditional paper-based world, valuable records have managed to survive despite this neglect. However, electronic records will not survive long enough to be used by future generations without active investments in their ongoing preservation and access.

Computer technology and electronic records create many challenges and opportunities for the records management and archival professions. This technology is a moving target and we need long-term solutions. Certainly, the National Archives and Records Administration has served as a leader in this field of research.

The Michigan Division of Archives and Records Management began discussing electronic records issues more than 25 years ago. With the advent of desktop computing, we became particularly concerned about how we could manage and preserve e-mail and word processed documents.

The U.S. Department of Defense issued the first version of its standard 5015.2 for Records Management Applications in 1997, and several commercial products are on the market that comply with this standard. We wanted to test the RMA software in State government offices to determine if it would address the record retention problems we were trying to solve.

We decided to apply for and we received a 2-year grant beginning May 2000 from the National Historical Publications and Records Commission to conduct a pilot project using a DOD-certified Records Management Application. At the time, the State Archives was an agency within the Michigan Department of State led by Secretary of State, now Representative, Candice Miller. The project had three goals: One, to assess the ability of an RMA to classify and manage electronic records and execute retention requirements, including the identification and segregation of archival records; two, to analyze the cultural impact that RMAs have on agency staff, information technology personnel, records managers and archivists; and three, to conduct a business process analysis and evaluate the potential for RMAs to be used in an enterprise-wide setting.

Our pilot project demonstrated that RMA software works. Electronic records that are created by common desktop programs can be organized and stored in a centralized repository that automatically implements the appropriate retention period for the records. However, we found people will resist changing the way they file and access their electronic records, because less than a third of our project participants adopted the RMA software as their primary tool for storing electronic records. The RMA features need to evolve to make the filing and retrieval of electronic documents appear transparent to users.

We also learned that business process improvements can be derived from using RMA software, especially when the business process change involves the transformation from a paper-based process to an automated process. When these improvements are adopted by the agency, RMA use and satisfaction does increase.

This project demonstrated that management support for change is essential to the success of an RMA. Managers must establish expectations and consequences for not following established procedures for electronic recordkeeping. Encouragement by management needs to focus on the positive benefits to the individual and the agency. The Department of History, Arts and Libraries is continuing to support the limited use of the RMA software now that our pilot project has ended.

However, RMAs are record retention tools, not preservations tools. The electronic records that are stored in the RMA's centralized repository remain in their native format. The RMA is not capable of ensuring that they remain accessible as underlying technology changes. Therefore, a methodology must be developed for preserving archival electronic records and those with long-term retention requirements. This is why the NHPRC initiated a partnership between the San Diego Supercomputer Center and the State of Michigan to address long-term and permanent preservation of electronic records. In November 2001, we were awarded a 2-year grant from NHPRC for what we call the PERM Project. This collaborative project is developing functional requirements for preserving electronic records that are stored in RMA repositories so they remain accessible. These functional requirements were published in January 2003, and currently researchers in San Diego are developing a prototype to test these functional requirements.

These and other projects are essential for ensuring that our documentary heritage remains accessible to future generations. Thank you again for inviting me to testify before you today.

[The prepared statement of Ms. Wojcik follows:]

Statement of Caryn Wojcik
before the
Subcommittee on Technology, Information Policy,
Intergovernmental Relations and the Census

Testimony: July 8, 2003

Mr. Chairman, thank you for inviting the Michigan Department of History, Arts and Libraries to testify at this hearing. I am honored to be speaking to you today about the management and preservation of our government's records, particularly electronic records. I especially want to thank Congresswoman Candice Miller for inviting us. My name is Caryn Wojcik, and I am the Electronic Records Archivist for the State Archives of Michigan.

The Michigan Department of History, Arts and Libraries is committed to enriching the quality of life for Michigan residents by providing access to information, preserving and promoting Michigan heritage, and fostering cultural creativity. The **Division of Archives and Records Management** within the Michigan Historical Center is responsible for ensuring that Michigan's government records are properly managed and preserved throughout their entire lifecycle.

Public records are essential for protecting our legal rights, for documenting the actions of our government, and for understanding the society in which we live; and therefore **it cannot be sufficient to have our history preserved by accident**. As a general rule, we estimate that less than 5% of all government records possess permanent historical value. Professional archivists and records managers have worked for years to develop **systematic processes** that identify which records possess historical value, and to provide for their preservation. Unfortunately, many **government agencies** do not follow these procedures, and they **fail to protect** the irreplaceable items in their custody. In the traditional paper-based world, valuable records have managed to survive despite neglect. However, **electronic records will not survive** long enough to be used by future generations without active investments in their ongoing preservation and access.

Computer technology and electronic records create many challenges and opportunities for the records management and archival professions. This technology is a **moving target**, and we need long-term solutions. Certainly, the National Archives and Records Administration (NARA) has served as a leader in this field of research.

The Michigan Division of Archives and Records Management[1] began discussing electronic records issues more than 25 years ago. With the advent of desktop computing we became particularly concerned about how we could manage and preserve e-mail and word processed documents. The U.S. Department of Defense (DoD) issued the first version of its standard 5015.2 for **Records Management Applications** (RMAs) in 1997, and several commercial products are on the market that comply with this standard.[2] We wanted to test RMA software in state government offices to determine if it would address the record retention problems we were trying to solve.

We decided to apply for, and we received, a **two-year grant** (beginning May 1, 2000) from the National Historical Publications and Records Commission (NHPRC) to conduct a **pilot project** using a DoD-certified Records Management Application.[3] At the time, the State Archives was an agency within the Michigan Department of State, led by Secretary of State, now Representative, Candice Miller. The project had three goals: (1) to assess the ability of an RMA to classify and manage electronic records and execute retention requirements, including the identification and segregation of archival records; (2) to analyze the cultural impact that RMAs have on agency staff, information technology personnel, records managers and archivists; and (3) to conduct a business process analysis and evaluate the potential for RMAs to be used in an enterprise-wide setting.

Our pilot project demonstrated that RMA software works. Electronic records that are created by common desktop programs can be organized and stored in a centralized repository that automatically implements the appropriate retention period for the records. However, we found that people will resist changing the way they file and access their electronic records, because less than a third of our project participants adopted the RMA software as their primary tool for storing their electronic records. RMA features need to evolve to make the filing and retrieval of electronic documents appear transparent to users. We also learned that business process improvements can be derived from using RMA software—especially when the business process change involves the transformation from a paper-based process to an automated process. When these improvements are adopted by the agency, RMA use and satisfaction increases. This project demonstrated that management support for change is essential

---

[1] The Records Management Services recently joined the State Archives of Michigan as part of the Department of History, Arts and Libraries, Michigan Historical Center. The two programs will be merged into a single Division of Archives and Records Management.

[2] The DoD 5015.2-STD and the list of certified RMAs can be found at http://jitc.fhu.disa.mil/recmgt/.

[3] For more information about the project, visit the project website, http://www.michigan.gov/hal (suggested search term "records management application").

to the success of an RMA. Managers must establish expectations and consequences for not following established procedures for electronic recordkeeping. Encouragement by management needs to focus on the positive benefits to the individual and the agency. The Department of History, Arts and Libraries is continuing to support the limited use of the RMA software. This way, we will be ready to make sound recommendations to Michigan government agencies when they are ready to use these products.

However, RMAs are record retention tools, not preservation tools. The electronic records that are stored in the RMA's centralized repository remain in their native format. The RMA is not capable of ensuring that they remain accessible as underlying technology changes. Therefore, a methodology must be developed for preserving archival electronic records and those with long-term retention requirements. This is why the NHPRC initiated a partnership between the San Diego Supercomputer Center (SDSC) and the State of Michigan to address the long-term and permanent preservation of electronic records that are captured and managed using RMA software. In November 2001 we were awarded a two-year grant from NHPRC for the "PERM Project". This collaborative project is developing functional requirements for preserving electronic records that are stored in an RMA repository so they remain accessible. These functional requirements were published in January 2003, and currently the researchers in San Diego are developing a prototype that will test these functional requirements.[4]

The Michigan Division of Archives and Records Management appreciates the support that the National Archives and its grant-making body, the NHPRC, have given us as we seek to find practical tools to ensure that electronic records are properly managed and preserved. These, and other projects, are essential for ensuring that our documentary heritage remains accessible to future generations.

Thank you again for inviting me to testify before you today.

---

[4] The PERM Project website is http://www.sdsc.edu/PERM/

Mrs. MILLER. We certainly appreciate your testimony today.

Our next witness is Dr. Richard Lysakowski. He is the executive director and president of the Collaborative Electronic Notebook Systems Association and the Global Electronic Notebook Systems Association [GERA] as well. It is a professional standards association built on government-industry partnership with a mission to create program, policy, and procedural and technology standards for quality electronic records programs. Dr. Lysakowski has 25 years of experience working with automation systems in various scientific software, engineering and project management roles in both the public and the private sectors.

We welcome you to the subcommittee and look forward to your testimony, sir.

Mr. LYSAKOWSKI. Thank you, Ms. Chairwoman.

A little bit of background on CENSA, the Collaborative Electronic Notebook Systems Association. Our primary mission is to create totally electronic replacements for inventors' notebooks and regulatory records, record books that are used in a research development quality assurance manufacturing other enterprise business environments. These are vital records that indicate the quality of a project that manufacturers are shipping. And the industries we cover include pharmaceutical, chemical, food and beverage, oil and gas, consumer products, government and many other industries.

Replacing the inventor's notebook has been one of the greatest challenges because frequently in a patent interference it will come down a single page in a single record book that makes or breaks a multibillion dollar a year, 20-year monopoly for product revenues. So we are talking about serious commercial impact.

Regulatory impact, if records are either fraudulent or inaccurate to the point where product quality is poor, can frequently shut down a facility—manufacturing facility for as long as 6 months, losing more than $1 million per day in potential revenue. So serious consequences can result from poor quality recordkeeping.

And so for these records, making them electronic has been quite a challenge. And we worked with NARA to find out what the best practices are out there now and in the near future and worked with many vendors. Our role as a market development association is to create multiple competing products for all the necessary components that make up a complete electronic system for records creation, management, archiving, preservation, retrieval and access. So looking at the whole life cycle of the electronic record has been quite important for us.

And NARA's focus has been, at least in its current funding, focusing on records management. We need it to include archives management specifications.

So some summary points that I needed to make: We are appreciative of NARA in having funded the Global Electronic Records Association to create the first quality electronic records practice standards. That work is just being complete now.

But looking more globally, I would say that NARA is underfunded to face the grand challenge of electronic records for the full lifecycle for all agencies that it serves and industry and the private sector in general.

Government needs other funding sources. NARA's budget for fiscal year 2003 is $268 million. That is less than the annual revenue of even one of our member corporations. So one issue is that NARA is insufficiently funded to achieve the goals for all agencies and the private sector too.

If NARA is to help the agencies and the citizens it serves, it should provide blueprints, mandate standards and requirements for policies and technology systems that implement records management and archives management. Don't forget the archives for each agency. The DOD standard does a nice job of covering records management but does very little on archives.

Third point is the OMB exhibit 300. I would concur with the first witness on this panel, that two parts need to be added to that exhibit. One is on the preservation of business assets. The information assets that are created by any IT system must be demonstrated to be completely exported and reconstructible in another system before any system is purchased. So that procurement process must specify interoperability between at least three to five systems to guarantee alternatives. So the procurement process itself needs to change. These demonstrations of reconstruction of an archival information packet need to be shown before any money is provided to a vendor.

The last point is subscription-based software pricing is one of the greatest threats we have to records and information assets. This is a new pricing model whereby you pay every 12 months for use—continuing use of your software to access your property. Software is a tool to create intellectual and business assets. Some vendors are pushing subscription-based pricing instead of perpetual-right-to-use license for the software. So unless this software purchase includes purchase of the format and some type of free viewer that goes along with it is made available, then subscription-based pricing shouldn't be pursued by the Federal Government for any system containing important business assets.

Mrs. MILLER. Thank you. You know, I might just pick up on that subscription-based pricing. I am just trying to understand what you were saying there. Because it would seem to me, with all the different—just the government demand itself, that there would be a software program that would be very viable in the marketplace, we are talking about setting standards and that kind of thing. Are you aware of such a thing, a governmental software for records management archiving? Mr. Sprehe had mentioned that as well.

Mr. LYSAKOWSKI. I am not aware of any software program freely from the government.

Mrs. MILLER. Mr. Sprehe, are you aware of any software that the government agency would utilize for standards uniformity?

Mr. SPREHE. I am not.

Mrs. MILLER. You mentioned, doctor, as well about NARA being underfunded, and I think you mentioned the amount of what their funding actually was in that. It would seem that it is very important that we identify in any funding request, when we talk about dollars here, from the congressional standpoint certainly, that we put in place in a business plan—as a plan that archiving for every agency, that records management, all of this would be a part of a

business plan before appropriators would talk about funding particular projects for any agency as well.

And Mr. Sprehe used the example of Yucca Mountain. I think he used two examples. I forget the other one that you mentioned. That they had actually included that in their business plan.

Is that a recommendation that you would make to the Congress that in any business plan that they should put those kinds of critical portions in there?

Mr. SPREHE. Most certainly, yes.

Mrs. MILLER. Do you have any comment, Doctor?

Mr. LYSAKOWSKI. Yes. I was advocating that you actually add at least one new section to exhibit 300 that requires during the acquisition process that vendors—at least three to five vendors demonstrate interoperability of archival information packets before you make any single purchase. Thereby, you set up an insurance policy for all congressional systems.

We have successfully used a strategy in industry for specific analytical techniques, mass spectrometry in particular is a good example. Dow Chemical wanted to purchase a mass spectrometry system, an instrument and the data system that goes along with it, $150,000 purchase, but, as a condition of the procurement, they short-listed the top three vendors, and they said, OK, we want you all to show us interchange of the mass spectral data sets among all three of you before we buy from any one of you.

Well, it took 2 weeks to write the code by one programmer, 2 weeks to distribute it to the other two vendors and embed it in their products, and the demonstration happened within 30 days. Now, that was for $150,000 procurement.

There was a specification that already existed, but it was not a standard and the vendors did not implement it. But if there is a specification, it is trivial for vendors to implement it; and on the last page of my testimony document, I state that vendors will use the lack of a specification as an excuse for creating proprietary file formats and buyers give up their power at the most critical moment by not demanding interoperability demonstrations.

Mrs. MILLER. I would be interested to know from the witnesses in the middle there, Ms. Wojcik from Michigan and also Mr. Naworocki from Virginia, do you—when you're making your funding requests to your appropriators, to your legislature for your work, do you put in as part of your business plan the kinds of funding that is necessary for records management and for archiving? Do you make that a big part of your business case? Do you find that there is enough attention being paid to what is necessary to actually archive these kinds of things?

Ms. WOJCIK. The State of Michigan does not require that record retention be addressed in business plans. We did draft a policy that's published in the Department of Management and Budget's administrative manual that requires that record retention be identified during the procurement of new systems, but, unfortunately, that has not been implemented. We do see that—if it was implemented systematically, it would certainly help us identify how long records need to be retained and which ones need to be preserved.

Mr. SPREHE. If I could add to my comment, exhibit 300 of the OMB circular does mention records but only with respect to the

Government Paperwork Elimination Act. The problem with that is GPEA is a very narrow act. It only affects information collected from the public. If you think of the Department of Defense, which has probably tens of thousands of IT systems, DOD collects very little information from the public, and all those other systems that may be producing or processing records, they don't pay any attention to records management.

Mrs. MILLER. Not to interrupt, but there I think is where you mentioned the cultural chasm, I think is the way that you had put it, between historians, perhaps, maybe librarians and IT, where you have people who are historians wanting to make sure that everything is archived properly and some IT people perhaps with the view that it happened, get over it, and not wanting to keep it.

So how you bridge those, that may be another question. I'll let Mr. Naworocki respond to the first one first.

Mr. NAWOROCKI. Sure. Within the State of Virginia, working very closely with the various IT agencies in order to try to embed records management into that, it hasn't always been successful. I think our biggest problem has been simply the fact that Virginia has had a severe budget crisis over the last 2 years, and that's overwhelmed it. But by working very closely with IT both in local agencies and on the statewide version in order to try to push that—I think it's very true there's a cultural disconnect. People assume, because it's electronic, it's not a record, or they just think it's too much trouble.

I think that possibly one of the areas that we need to work in is not to work on very specific projects but rather to take a look at how we can leverage the power of the States and the Government—the Federal Government in order to talk to some of the software manufacturers about embedding it.

I'm just thinking, well if we can convince Bill Gates it was worth a few bucks to him to embed records management and to award every other product he developed, we wouldn't have a problem—or archival, actually. That's literally when it comes down to. Whenever you have to graft something on, it becomes a kludge and just doesn't work. It needs to be embedded in the original activity in order to make it function, really.

Mrs. MILLER. Do you have any comments on how we might bridge that cultural chasm there with—how can the Federal Government assist in that kind of a thing in setting out standards? As you mentioned, it should be embedded in the original technology. Is there something that we can do to bridge that?

Mr. NAWOROCKI. You're the 800 pound gorilla. Whatever you guys decide, they are going to follow.

Once again, the most important thing is that the Federal Government has to demonstrate a concern. They have to demonstrate a commitment to that, and then that flows from that. I think it also needs to be remembered that, really, it should not be just strictly what the Federal Government does, because what may work for you may not work for us, but rather if you look at it as a collaboration among private industry, the States, local governments and the Federal Government, I think certainly, as you continue your work in this project, you should take a look at what is being done in those other areas in order to determine what is the

best practice. Very simply, just because it's been done in NARA for the last 50 years doesn't always mean it's probably the best thing to be done.

Ms. WOJCIK. You know, this is a technology problem, and it's going to need some type of technology solution, and the people who create that technology are part of that solution. Having tools is just one part, but there's also that cultural issue, and I think we need to see more accountability from the top down, where, you know, administrators, managers are accepting that responsibility, acknowledging it, are fully aware of the consequences of not taking responsibility for both records management and records preservation. Working form the bottom up isn't going to solve anything.

Mrs. MILLER. Working from the top down and setting that kind of standard I think is very important.

I just mention a personal example that I have within my own office, but I'd like to address this question to the State librarian. Are you familiar with—libraries I think with the advent of the Internet and all this electronic, people thought, well, geez, no one will read books anymore, that libraries would lose their value, that they would go by the wayside. But in fact I think libraries have been very much on the leading edge of using electronics and reaching out to the community and in so many different ways.

Are you familiar with the digital library research that's being conducted by the Library of Congress? And, if so, how does the Library of Congress—I'm sure you have your national association—how does that filter through the libraries as far as, again, a mix of cooperation between the libraries and the historians?

Mr. NAWOROCKI. I'm not the State librarian. I work in the records, but I understand the library process. I think what happens also is there's a number of different government agencies for different ways, and the Library of Virginia just as the national—the Library of Congress, we also have our own digital library project, and that is a very important part because we have very unique items, and through the Internet, through digitization, we're able to spread that information, make it available. So that someone who is doing genealogical research in their jammies in Australia—they don't have to make that trip.

I think it's very important that both the archival and the records management world also understand what is being done in the digital library arena, because they're doing some very important work on access and how to obtain access and indexed material that often we don't think about on our side. I think sometimes I really would love to just take all the research that is being done, throw it in a blender and hope that the answers come out, but I think oftentimes we stovepipe but not always talk to each other, and we need to do more talking.

But, yes, it's filtering down. I think many State libraries, local libraries are working on digital projects and, you know, preserving and making accessible information that otherwise would be totally inaccessible, except for those few folks who could make it to the library.

Mrs. MILLER. Well, as we're getting into the lunch hour, this will be my final question for the panel. You were all here for the first panel, and I think we are just trying to get to the crux of the prob-

lem here of trying to determine what kinds of records really need to be archived and what is the best practices. I remember in specific Ms. Koontz from the first panel had made a comment I think about Mr. Sprehe, that you had developed some preliminary models for recognizing the best methods for records management. Perhaps you could elaborate a little bit on that, if you would, sir.

Mr. SPREHE. Thank you. I put a little bit of that into my testimony.

One of the things that I learned and that I guess the team learned in this study of transitioning from physical to electronic records management—and we have applied the metaphor of a capability maturity model but only as a metaphor. The strategic advantage was that it speaks to the IT community. They know what that is. My view is that we have placed stress to a great degree on the bad things that will happen to you if you do poor records management and that the costs and risk—and that has limited appeal to top leadership in any enterprise, including Federal agencies.

What we discovered is—and which I tried to present with the concept of leveraging the records management investment is that if you do good records management, very good things happen to you beyond records management. It helps you do other business functions far better, more efficiently and effectively. That grabs the attention of top leadership.

In a recent presentation commending the ADAMS system at the Nuclear Regulatory Commission, the recipient from NRC said, you know, it took us a long time to get our electronic records management system into the shape it is in today. Now what's happening is that, as other systems are developed in our agency, those IT managers are saying, and we want a system just like ADAMS, because they understand that if you can grab instantaneously yesterday's records and all the way back electronically, it's an enormous asset that they should build into their application.

Mrs. MILLER. Thank you.

Well, I certainly want to thank you all for attending today. It has been fascinating. It really is quite a fascinating subject that we all struggle with, with our changing world and changing technology and how we best utilize it for everyone.

Is there anything that any one of you four would like to add for the record that we didn't ask you or questions or input that you would like to put in as part of the record?

Mr. NAWOROCKI. If I might just briefly to touch on two points. One is accountability. And I think working with the GAO and within the State of Virginia we're looking to work with our accounting agency in order to develop that accountability. But, also, we need to understand that everyone will have to become their own records manager, and it's going to require a lot of training and a lot of understanding on behalf of folks to do that, and I think it's a long road ahead, but it really is going to require a great deal of education as well as the technological capabilities.

Thank you.

Mrs. MILLER. Thank you all so very, very much for coming.

The record will be held open for 2 weeks for submission of additional testimony or questions, and we'll be happy to have those given to the subcommittee as well.

At this time, the hearing will be adjourned. Thank you.

[Whereupon, at 12:12 p.m., the subcommittee was adjourned.]

[Additional information submitted for the hearing record follows:]

CENSA © Copyright 2003 – All Rights Reserved Internationally

Testimony from Dr. Richard Lysakowski, Executive Director of the Global Electronic Records Association (GERA), President of the Collaborative Electronic Records Association (CENSA)

House Government Reform Committee, Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census

June 8, 2003

Important Points to Cover:

NARA Alone Cannot Solve eGovernment's Problems with Electronic Records.

Electronic Records are not uniquely NARA's problem. They are inherent to all of OMB's eGovernment goals. Electronic Records are inherent to all work processes in all federal and state agencies. Every time that private industry has approached OMB to develop common government and industry-wide solutions for electronic records, OMB has repeatedly pushed all responsibility for electronic records onto NARA's shoulders as the "Managing Partner for the Electronic Records Management (ERM) e-Gov Initiative." NARA's ERM Initiative stated goals are to provide a government-wide policy framework and guidance for electronic records management. However, NARA's mission, scope, focus, size, budget ($268M FY2003), and capabilities do not permit it to solve all local electronic records problems of federal, state, industrial, and nonprofit organizations. NARA's primary mission remains to "ensure continuing access to essential evidence that documents: 1) the rights of American citizens, 2) the actions of federal officials, 3) the national experience." NARA as our federal archives takes stewardship of only a very small percentage of federal and other historically important records for long-term preservation and access "until the end of the republic."

Industrial and private sector problems with electronic records are driven by many other factors not directly related to NARA's mission. For example, the private sector are driven by many concerns simultaneously, e.g., USPTO and intellectual property protection concerns, EPA and environmental protection and quality concerns, FDA and healthcare product safety and quality concerns all drive, SEC and accurate financial reporting and integrity concerns. These are just a few of the many concerns that are immediate, but have long-term impact on the private sector and industry.

Problems of electronic records are multifaceted – driven by administrative, legislative, legal, regulatory and business concerns – and implemented via well-orchestrated organizational programs, qualified people, policies, quality standards and procedures, and technology systems. While the needs for keeping high quality evidence of activities are universal throughout the public and private sectors, the scope, scale, and reasons for recordkeeping vary so widely that NARA's solution and approach cannot and will not work for all.

For eGovernment Reform regarding Electronic Records to work, a more universal, distributed approach is needed that does not put NARA at the center of everything, hence causing one of the smallest federal agencies to become a bottleneck in moving to Electronic Records as the basis of recordkeeping in eGovernment. The universal, distributed approach can be coordinated and aligned with NARA's assistance, however, some fundamental strategy shifts must occur so that NARA does not constrain the process it seeks to facilitate.

CENSA © Copyright 2003 – All Rights Reserved Internationally

NARA's ERM eGov Initiative – Where's the eRecords Archive in each eGov Agency?

NARA's current strategy focuses on assisting other federal agencies with Electronic Records Management Initiatives by providing a methodology for determining agency-unique requirements on top of the US DoD 5015.2 RMA Standard. While this approach takes good advantage of the excellent work of the DoD Version 2 standard – which NARA had significant input in developing – it does not address the need for each federal agency to set up and run its own Electronic Records Archives. It does not lay down the standard functionality or format standards for Electronic Records Archives to successfully deal with the 99% of agency records that NARA never receives.

NARA's ERA Solution Does Not Solve Other Agencies' ERA Problems

NARA's Electronic Records Archive (ERA) Solution is a NARA-specific solution that is being designed first to solve NARA technological needs for a system that will solve NARA's internal problems that are growing exponentially in scale in time. The NARA ERA is by design a "custom" system designed to NARA's specifications. NARA's ERA is not a general-purpose, "Configurable-Off-The-Shelf" (COTS) product designed to scale up to large agencies the size of the Department of Defense or Homeland Security, or down to much smaller agencies at the federal and state levels.

Procurement Processes Are Not Designed to Protect Agency or Taxpayers' Property

See slides and handouts for more information.

Buyers' Behaviors Must Change to Require Full Export and Interoperability of Their Data

See slides and handouts for more information.

Subscription-Based Software Pricing Is the Biggest Threat to Record Security and Assets

Software products are tools used to generate personal and business property (data, information, and knowledge). Like all tools, software depends on having quality product manufacturers. Software vendors would love to move software to become a service, utility, or commodity, like water, electricity, gas or oil that is consumed and must be re-paid for periodically – either monthly, annually, or some other period. Large vendors like Microsoft have been the first to make this move to subscription-based pricing. It shocked many companies who got their first software "subscription bill" for the following year. If they did not pay their bill, their software would have been "deactivated", shut off, or somehow made unusable, thereby denying them access to data and documents they own.

I have not taken the time to locate the exact federal and state legal codes, but I know there are federal and state laws against denying one access to one's property. So I question the legality of this method of selling software without giving buyers a no-cost exit (easy migration) strategy.

Advocates for subscription-based software licensing (software vendors) argue that this eliminates the many hassles with continuously having to update software packages, because this can be done automatically as part of the subscription. They also argue that it prevents piracy because if software "times" out or becomes "deactivated" automatically if the software vendor doesn't receive its regularly scheduled payment, then software pirates are unable to use stolen software. These are valid arguments

CENSA © Copyright 2003 – All Rights Reserved Internationally

for software environments where dynamically updated software is done during the active subscription, and concerns about stability and "system quality" and validation are not of tantamount concern.

However, when it comes to business assets, a business' access to its own property must NEVER be terminated because they did not make their monthly or annual payments to the vendor of the software. This is similar to the extortionary tactics used by Mafioso that shut down businesses and take people's property hostage unless regular payments are made. Records are corporate or government property; one's access to them must NEVER be at risk of "deactivation." Subscription-based pricing models being foisted upon unwary buyers in government, industry, and private sector consumers give free license to software vendors to abusive or unethical business practices. Open data and record portability standards whose full rights are purchased at the time of software purchase are the only way around this potential abuse or loss of access to one's property. What we need more than software "utilities" is higher-quality software product manufacturers.

Lack of Government-Mandated Record Storage Format Standards Is Holding Back Progress and Innovation

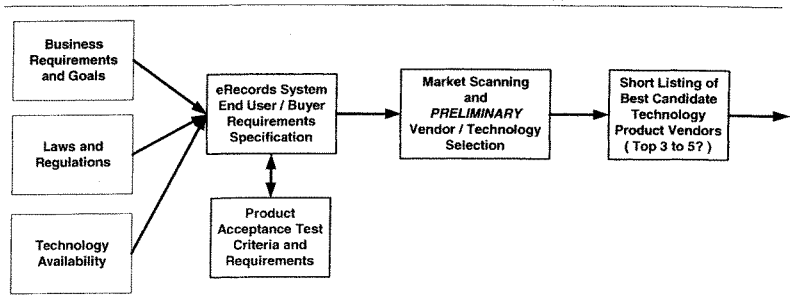
The slow formal acceptance and adoption of the Portable Document Format (PDF) as an official record storage format by the US Federal Government as a whole was a shining example of slowing down innovation. It has been accepted and mandated by law in other countries, including Australia, Italy, Holland, and Germany. I may elaborate on this during testimony.

Risk Assessment in Section 300 of OMB Circular No. A-11 (2002) Does Not Sufficiently Detail the Process to Protect Assets Over Long (Archival) Periods of Time

OMB must specify a process for procurement of eRecords systems that requires demonstration of complete interoperability and supplier independence (full capabilities to migrate away from one vendor to another) BEFORE purchasing. With electronic records, this is the only way to provide a minimal insurance policy against rapid planned obsolescence by computer hardware and software vendors.

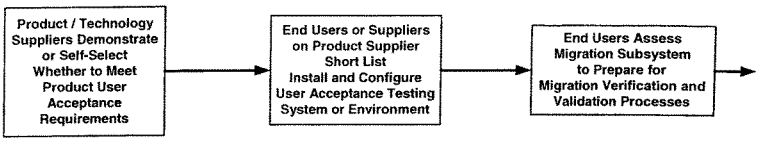
CENSA developed a simple, logical recommended process for protecting business' assets that includes four phases, illustrated below. Supplier candidates must pass each phase in succession.

Phase 1 Articulating Requirements and Finding Qualified Suppilers and An User/Buyer Driven Process with Input From Suppliers

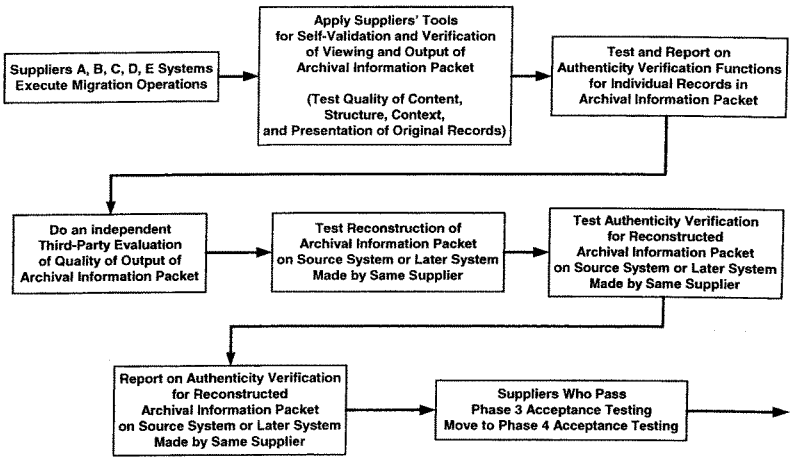


CENSA © Copyright 2003 – All Rights Reserved Internationally

Phase 2 - Work With Qualified Suppliers to Prepare User Acceptance Testing Systems / Environment

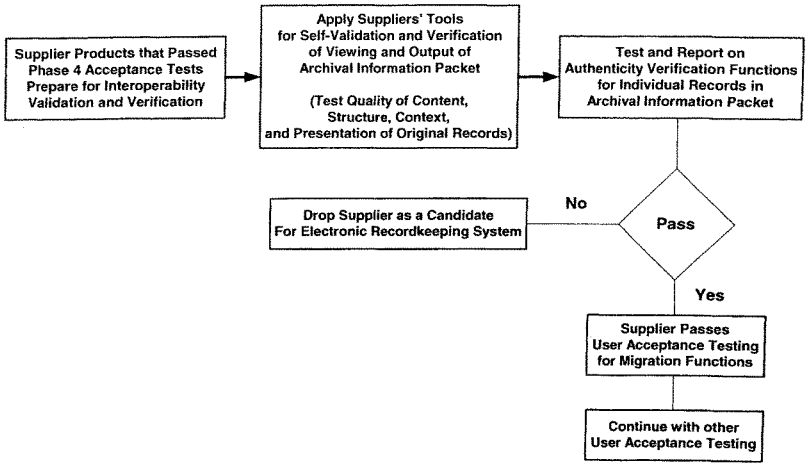


Phase 3 - Test Quality of Output of Migrated Archival Information Packets



CENSA © Copyright 2003 – All Rights Reserved Internationally

Phase 4 - Test Quality of Interoperability of Migrated Archival Information Packets



Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

"LOOMING INFORMATION AGE CRISIS EXPECTED TO CAUSE TRILLION-DOLLAR LOSSES OVER NEXT 20 YEARS"

Also known as

"TITANIC 2020 - A CALL TO ACTION"

by Rich Lysakowski and Zahava Leibowitz

Research Publications

1 EXECUTIVE SUMMARY 2
2 THE ORIGIN OF TITANIC 2020 3
3 FOCUS OF THIS REPORT 4
4 THE US NATIONAL ARCHIVES: EXPERIENCED SEA CAPTAINS IN NEED OF BETTER SHIPS 5
5 MODERN RECORDS: THEY DON'T BUILD THEM LIKE THEY USED TO 5
6 THE GROWTH OF ELECTRONIC RECORDS 6
7 RAPID PLANNED OBSOLESCENCE: THE HEART OF THE TITANIC 2020 PROBLEM 7
8 A "DAY-IN-THE-LIFE" HASSLES OF MIGRATION AND INCOMPATIBILITY 8
9 ECONOMIC MODEL OF THE USERS' COST TO UPGRADE FROM OFFICE'95 TO OFFICE'97 10
10 BREAKING THE STATUS QUO: GIVING DATA CONTROL BACK TO ITS RIGHTFUL OWNERS 12
11 AVOIDING THE ICEBERGS: THE ITINERARY HAS BEEN CHANGED 14
12 THE DIE HAS BEEN CAST FOR 21ST CENTURY INFORMATION TECHNOLOGY 17
13 THE SHIP HAS LEFT THE DOCK 17
14 BACKGROUND ON CENSA 18
15 BACKGROUND ON THE AUTHORS 18
16 FOR MORE INFORMATION 18

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

"TITANIC 2020 - A CALL TO ACTION"

by Rich Lysakowski and Zahava Leibowitz

1 Executive Summary

The potential electronic crisis predicted to hit the Western world upon entering the Year 2000 was the subject of global concern for many months, if not years. The airwaves were filled with the speculations of everyone from diskjockeys to systems analysts to politicians on whether public utilities, banks, or nuclear weapons would fail. Y2K Plus, a company specialized in solving Year 2000 computer problems, estimated that the cost of prevention and repair of Y2K technical problems will total approximately $750 billion dollars. With legal preparations and litigation the total cost will likely be over 1 trillion dollars.[1] Intense preparations for years, replacing old computers and rewriting millions of software programs spared the world from major problems.

While it is no wonder that Y2K inspired thoughts of a coming apocalypse, a much larger and very real crisis is on the horizon now -- one we call "Titanic 2020." It has been brewing for a long time and grows larger everyday. Titanic 2020 results from software business practices that make new products quickly obsolete, cause major hassles for users, and put critical assets at risk. Comparing Titanic 2020 to Y2K, Y2K was like the tip of an iceberg. Like the passengers on the Titanic on her maiden voyage, many of our most valuable records stored by the ship of computers will perish.

Icebergs are dangerous because they do not reveal their size until we get very close to them.Unlike the iceberg that sank the original Titanic, the iceberg looming in our future is man-made. Titanic 2020 is not just one iceberg or one definite deadline like Y2K -- it will be many icebergs scattered across many future horizons. By then it is too late because the real danger is below the horizon.

Simply stated, Information Technology (IT) of the early 21st Century -- and the organizations that use IT -- are generally ill prepared to prevent damage or loss of valuable electronic records or data. Every time we migrate electronic records or data stored in computer systems to the next generation of technology, we potentially face perilous collisions. Some collisions will lose priceless personal data and records. Until some important changes occur in Information Technology itself, large losses will mount. Every organization that uses computers for storing critical data or legal records will have Titanic collisions and full-blown disasters from irrecoverable data and violated records. Countless personal and professional lives will be lost in failed rescue attempts. Costly lawsuits to recover damages for lost properties will tie up courtrooms for years.

This report identifies the primary causes of the impending disasters we will all face as we rush into our digital future. It provides an economic analysis of a recent, extremely costly software migration that resulted from incompatible data formats. The example is at once practical, symbolic, and symptomatic of the problems for technology users caused by the software industry at large choosing a strategy of rapid planned obsolescence. The report also admonishes users to take back ownership of access to their data because today, guaranteed long-term access to records requires a continual subscription with an unknown cost-of-ownership price tag. Lastly, this is a Call to Action to reduce the risks and costs of ownership of Information Technology to finally support permanent electronic records by changing business strategies and buying behaviors. The Titanic 2020 problems are well understood, and will persist by design, until this Call to Action is heeded.

---

[1] See Ted Smalley Bowen. "Y2K Legal Wrangling Escalates." InfoWorld, November 29, 1999, Cover Story.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

2 The Origin of Titanic 2020

The Titanic 2020 icebergs started to nucleate at the start of the Information Age, roughly in the mid-1960s, when we started to rely on electronic computers for business records. The icebergs have been growing rapidly in size and number since then. IT systems of the early 21st Century create and store "fragile" electronic records and data, meaning their integrity is easily broken when they are removed from the technology that was used to create them. Once broken, an electronic record's integrity cannot easily be restored. At first, electronic records were mostly historical in nature, but increasingly we create them for critical business and scientific matters too. For society to move all the way into The Information Age, we need to replace paper and microfilm as the default permanent storage media for data created electronically. We must still take some important transitional steps.

The Y2K problem was the result of using a shorthand, two-digit year code instead of a full four-digit date code. However, for modern electronic records, this date code is only one item in a long list of parameters that must be kept to ensure integrity and access in the future. Like the Y2K problem, which is the result of a lack of concern over the past 30 years with a known parameter problem, the problems leading to Titanic 2020 are also known, but little to no action is being taken to correct them. We continue to produce digital records in formats with short life spans that are increasingly difficult and sometimes impossible to access years later.

"Titanic 2020" is a moniker for the serious problem of widespread software design flaws in today's IT systems. These design flaws will affect every organization that must keep today's electronic records well into the 21st Century. Every organization is building its own fleet of Titanic ships in the form of electronic records technology systems and data systems. Many of the records these vessels carry will be shipwrecked on the dozens of icebergs growing in the future seas of IT. Each migration event is a point in time when records must change carriers and preserve their full integrity and accessibility. In front of each migration event on the horizon floats a potential man-made iceberg with a size and shape unknown at this time.

The designs of today's IT systems simply do not permit facile or reliable migration of their contents, at a predictable or reasonable cost. Unwittingly, by our own design, we keep creating the ships without lifeboats to carry us safely into the future. Not only are we as a society failing to take note of the problem looming over us, like the builders of the Titanic, we are continually taking steps to seal our fate.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

3 Focus of This Report

This report focuses primarily on legal records; however, in most cases it applies equally well to less formal or secure compilations of "data." Legal experts widely agree that electronic records and data are admissible in court as evidence. They also agree that electronic data can have a legal weight equivalent to formal, signed legal records -- as long as the data integrity has been preserved beyond a shadow of a doubt since the time of its creation to the time of its surrender in legal proceedings.

This report focuses on matters of critical concern to individuals, businesses, and governments.

Individuals may think a report about electronic records does not apply to them, since they may use personal digital devices to create data for their eyes only. However, this report applies to any collection of any information with integrity that you may want to preserve and access years later.

At the individual level, personal digital devices make our lives easier on a daily basis. We increasingly rely on them in our daily lives. However, digital devices do not yet provide reliable and easy ways to move our digital results across generations of technology. Digital records sometimes change irreversibly as we move them across different devices and software packages over time -- sometimes altering critical characteristics of authenticity, reliability, or accuracy that allow them to be used as historical records later. We increasingly entrust priceless sentimental works and events to digital-only devices. If a record's content is not entirely lost in migration, sometimes it can be corrupted or damaged irreversibly, a fact which may only be discovered after great personal loss of time trying to recover it. Such time and effort spent to reliably migrate digital data through software versions over time needlessly encumbers and wastes our personal lives.

At the levels of business and government, records are the lifeblood of accountability for past actions, and data are the basis for current actions. Data can be useful for very long periods of time even though it changes from one data processing software to the next, and from one computer system to the next. At some point processing stops, conclusions are reached, and a permanent record is created. In the rush to become ever more efficient and effective, manage knowledge and leverage intellectual capital, we have overlooked the vessels to carry electronic records of action far into the future for both legal and historical reasons. In any endeavor, seeds for tomorrow's success come from today's intellectual property and other recordings of knowledge. We are killing those seeds by using vessels that can easily sink and destroy the integrity of the seeds. At the same time, we are creating unknown legal and economic problems and losses for ourselves that will be nearly impossible to estimate until they are upon us.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

4 The US National Archives: Experienced Sea Captains in Need of Better Ships

As the keeper of the government's records, the United States National Archives and Records Administration (NARA) was among the first groups to experience, on a relatively small scale, the danger that lies in front of us all. Its buildings in and around Washington, D.C. are packed with as many records as they can physically hold, including many original audio and video recordings of historic moments in time, all on various forms of outdated media. In order to access these media, the original audio players and video projectors must be reconstructed and then each recording must be played -- sometimes in its entirety -- in order to catalogue its contents and separate out accessible from inaccessible information and key historical information from the unimportant. Due to a lack of funding and resources, many of these records will molder, crumble, or rust before they can be reviewed and catalogued by archivists. It is part of NARA's responsibility to decide which records submitted to them they will have the resources to preserve in an accessible form and which are beyond their resources (and may be lost forever.)[2]

Digital media such as digital audio, digital video, and other more complex digital data types (scientific, engineering, and business data) are rapidly proliferating. Digital media is quickly becoming the preferred recording method -- strongly favored over traditional paper and analog methods. Digital formats permit innovative computer processing, searching, mining, and reuse of all types of knowledge. However, accompanying this ease of use is "The Digital Time Bomb" - an indefinite period of time after which data is inaccessible because no one paid money or attention to ensure its accessibility. Complex digital records are increasingly difficult to deal with because there are few universal standards for long-term preservation and access. Each software supplier creates its formats and changes them as needed for a variety of reasons. Many people hope that the World Wide Web and technologies such as the eXtensible Markup Language (XML) will offer some stability and solve universal access problems. However at the current frenetic pace of innovation and market competition, chaos will remain the norm until enough end users and suppliers are organized and work together effectively to produce stability in the marketplace. Currently, nearly everyone using computers is forced into a defensive position of preserving records in the lowest common formats -- in the worst cases -- by printing to paper.

5 Modern Records: They Don't Build Them Like They Used To

The problem faced by the US National Archives is in many ways much simpler than Titanic 2020. The records stored at NARA represent only the history of the US Federal Government. NARA's repositories hold historical records of operations of all branches of government and public service since the United States was formed. NARA's records are the primary instruments of accountability of past public administrations. NARA typically does not receive records from a public administration until after it leaves office. While the records hold tremendous value with respect to our understanding of who the American People are and how they got to the present, loss of these records is unlikely to dramatically affect the day to day lives of most citizens. Such is not the case with most electronic records being created today in business or by individuals.

Most electronic records today are produced and relied upon in the normal course of everyday business. Individuals become most aware of how much they are affected by the integrity of these records when their bank statement or credit card bill has an error. Businesses are even more

[2] See Stille, Alexander. "Overload." New Yorker. March 8, 1999, p38ff.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

dependent on the integrity of their records, since record integrity is critical to defending patent rights, demonstrating compliance with government regulations, defending tax returns, etc. As the world population grows rapidly, and society becomes increasingly litigious, the number of records that individuals and businesses must create and depend upon is increasing exponentially.

6 The Growth of Electronic Records

Accurate models of the growth of electronic records are difficult to construct, but we know that several factors are important. The most important ones are population growth and the accelerating adoption of information technologies to achieve greater efficiencies of scale and sharing of knowledge than paper-based systems can provide.

Models of growth for the Earth's population provide the most compelling drivers of growth. Population experts predict that the global population will expand to around 11 billion by 2050 and potentially to 24.8 billion by 2150.[3] It is obvious that for each person an increasing number of vital records must be kept: certificates of birth and death; marriage and drivers licenses; mortgages, stocks and other property transactions, plus many other types of routine transactions that we need or want to record. With so many transactions of lasting legal value, it is easy to see that the number of records will continue to grow exponentially. In fact, most institutions today -- whether corporate, government, or even small businesses -- harbor not-so-secret fears that the volume of important records is growing so fast that they will soon become, if they are not already, impossible to manage.

NARA provides a good example of the explosive growth of electronic records. Between 1972 and 1989, NARA archived 5000 electronic record files. Between 1989 and 1999, it archived 100,000 electronic records. In 2000, it expects to archive at least 1,000,000 records. In 2002, NARA expects to archive at least 40,000,000 e-mail records from the Clinton Administration White House alone. Simultaneously, it is preparing to archive approximately 90,000,000 personnel records from other branches of government.[4] NARA does not archive state and local records, that is the duty of state and local archives. Nor does NARA archive the millions of records of experimental and medical research, tax records, or countless other types of confidential records that could be important to future citizens. Generally each institution must archive their own records, unless they are of particular historical importance to the public.

Jean-Pierre Wallot, the former National Archivist of Canada, stated in 1994 that "more records have been generated during the past ten years than in all of prior human history."[5] Because of the rapid adoption of computers for business in the Information Age, the time required to double the number of critical records is decreasing rapidly. At the current rate, the number of records will double in 5 years, and double again less than 3 years later. If the number of records continues to grow at the same rate of growth as that of the human population, the numbers are staggering; simple math tells us that ***within 10 years, the number of records produced on the planet could be doubling every 60 minutes.***

---

[3] Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat (1999). *Long-range World Population Projections: Based on the 1998 Revision.* (ESA/P/WP.153).
[4] Ken Thibodeau, Director of the Center for Electronic Records Archiving at NARA supplied these statistics.
[5] See Wallot, Jean-Pierre, The National Archivist of Canada. "Archival Oneness in the Midst of Diversity" American Archivist, Vol. 57, Number 2, Spring 1994, pg. 380.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

Luckily, several mitigating factors will constrain the problem somewhat. Good records management practices dictate that business are only required to retain records for the number of years mandated by law. Some types of records can be deleted soon after their creation. For example, electronic commerce transaction records are short-lived, requiring retention for only two to three years. The IRS requires that personal tax returns only be kept for 7 years.

The Titanic 2020 problem is compounded by the fact that longer retention periods are mandated by law for some types of business records. The Food and Drug Administration (FDA) requires records for demonstrating regulatory compliance to be retained for 25 years or more. Patent records must often be maintained for as many as 35 years (including filings for claim extensions). Trade secret records (e.g., formula for Coca-Cola, for example) must be maintained until it is made public, if ever. The FDA requires records created during development and manufacturing of implanted or replacement body parts (pacemakers, organ transplants, and, in the future, genetic therapy records) to be retained for at least as long as the affected patients are alive. The FDA also recently introduced a new law that requires electronic records to be kept in addition to paper records.[6] In the future, other federal agencies will make this a requirement as well. It is not enough for these records merely to be retained, they must maintain their integrity and be accessible. Preserving record integrity and accessibility can be particularly difficult since this involves the ability to view and print an accurate and complete visual rendering of the record at any time it is needed in the future.

7 Rapid Planned Obsolescence: The Heart of the Titanic 2020 Problem

With regard to records created over the past 10 years, The National Archives faces a relatively small problem that is the result of old technologies becoming obsolete and unavailable. Individuals and businesses that create and drive our economy, our government, and our educational and research institutions face a similar problem of Titanic proportions. Whether or not data integrity is compromised during a migration (a potential financial burden in itself), a tremendous financial burden accompanies almost all data migration. Format incompatibility results in reduced productivity, which becomes a constant and costly problem. Even though modern technology allows electronic records to be created, accessed, and searched more quickly than records in any other form, because of planned obsolescence and the rapid decrease in time to obsolescence, this same technology ensures that these records must be regularly migrated to the newest available technologies (which are frequently unable to read the older formats). Miss some migrations and you may lose access to critical records entirely or regularly migrate data and absorb the associated financial loss. Neither choice is one we can afford.

[6] See www.censa.org/21CFRpart11.html or www.21CFRpart11.org for extensive information on this regulation.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

8 A "Day-In-The-Life" Hassles of Migration and Incompatibility

File incompatibility is a frequent problem between old and new versions of the same software. The two examples given here, Microsoft Office and Lotus SmartSuite, were chosen only because they are familiar to many people. The problems described here are so commonplace that almost any other software for creating or managing data or records could have been used to illustrate the problem.

In the case of Microsoft Office, a file incompatibility exists between the Office95 and Office97 versions of MS-Word, which for simplicity we will call Word'95 and Word'97.[7] Although the price of the software upgrade was only $80, the incompatibility creates many hidden costs in the version upgrade. The hidden costs arise from the loss of productivity as users unsuccessfully attempt to migrate documents from Word95 to Word97. More hidden costs arise from the need to train employees on all the differences between the two product versions, as well as strategies for migrating documents. Many other hidden costs exist depending on the context of the upgrade, e.g., personal versus business. Hidden costs associated with software upgrades are usually many times greater than the actual cost of the software.

For example, when Word'97 was released, documents created and saved in a Word97 format could not be accessed via Word'95. Thus, when a Word'97 user sends a Word'95 user a document, the latter is unable to open, view, or print the Word'97 file. If the Word'95 user is unable to speak directly to the Word'97 user, all work stops until she may do so. When the Word'95 user and the Word'97 user manage to connect, they may quickly discover the problem is due to the fact that they are using different versions of Word, if they are experienced or sophisticated users. The Word'97 user then resaves the file in the older format and resends it. If the users are inexperienced, it may take them a long time to discover what the problem is and how to solve it. Additionally, even though the Word'97 user saves the document in the older format, if the document contains new (and incompatible) formatting commands, some parts will look jumbled when it is opened in Word'95. The Word'95 user must again contact the Word'97 user to determine whether or not the document's appearance is correct. If one or both of the users are experienced or sophisticated, they may quickly realize that the new formatting will either need to be removed or manually reproduced in the old format. If they are not sophisticated users, they may not be able to recognize the source of the problem, or they may conclude that the document file is corrupt and must be entirely regenerated from a much older version, or even from scratch. In any case, manually reformatting even part of a document can be tedious and time consuming if the document is complex and uses many styles of headings, tables, and figures. When the Word'95 user believes the reformatted document is finally complete and correct, the document is generally faxed to the Word'97 user for a final review. At this point, after spending anywhere from a few minutes to several hours on solving problems due to incompatible format issues, many users will choose to fall back on the reliability of the older technology.

Incompatibility problems were very common for the first two years after the release of Office'97. Many large organizations have a policy of not upgrading to newer versions of software for many months or even years after new versions of software are available in order to avoid bugs and other

[7] The version of Word released with Office95 was actually Word version 7.0. We refer to it as Word95 to avoid confusion between MS-Word version released with Office95 and MS-Word version released with Office97.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

costly problems inherent in popular software. Unfortunately, the longer a company waits to upgrade, the longer they and the organizations they work with must suffer because of incompatibility problems. Sometimes vendors will create point solutions to these problems but fail to announce that such solutions are available. For example, about 18 months after the release of Office'97, one of the researchers of this article called Microsoft to find out if a Word'97 to Word'95 translator was available for Word'95. After paying $35 for the technical support call, the author was informed that a file translator was recently posted on the Microsoft Technical Support website. To the authors' knowledge -- confirmed by polls of fellow users of Word'95 -- no notification of the availability of this file translator was ever sent to customers.

Often, some people assume that with enough finagling, it is possible to work around the incompatibility problems between older and newer formats, but this is not always the case. Complex documents, particularly those that include graphics and tables may convert very poorly, some not at all, and some may actually become corrupted. For example, it is not uncommon for pointers (invisible electronic markers that indicate what should be displayed where in a document) to be misinterpreted, damaged, or destroyed in the process of opening a file in a different format. Such corruption often results in the total loss of formatting and/or content.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

9 Economic Model of the Users' Cost to Upgrade from Office'95 to Office'97

The following economic model is based on a population of 100,000,000 users at work. One hundred million was chosen because in 1997, Microsoft claimed to have sold at least 80 million copies of MS-Office. After only two years, this would add up to more than 100 million users.[8]

For calculations, it is assumed that each user work 37.5 hours per week, 48 weeks per year.

During the first year after the software upgrade's availability, it is assumed that 10 minutes was lost per week per user. This is an extremely conservative number; many users reported losing much more time before recognizing and learning to solve the problem -- even fairly sophisticated users.

Year 1 - Time Lost			Description
Minutes		10	Minutes Lost on Average per Week
	(10 x 48 weeks) =	480	Minutes Lost Annually per User
	(480 x 100,000,000 users) =	48,000,000,000	Minutes Lost Annually by Total User Base
Years	(48,000,000,000 ÷ 108,000[▽]) =	444,444	Work-Years Lost by Total User Base

[▽] 108,000 is the number of minutes in a work-year (37.5 hrs/wk * 48 wks/yr)

In the second year after the software upgrade's availability, it is assumed that time lost was reduced 50% because people learned to recognize and solve the problem faster, and more people have upgraded to Office'97. Thus, during the second year each user lost an average of 5 minutes per week. This means that 222,222 work-years were lost by the total user base in the second year.

Similarly, in year 3 after the software upgrade's availability, it is assumed that the time lost due to incompatibility problems was only 1 minute per week per user, and 99% of users have upgraded. As a result, 44,444 work-years were lost by the total user base during the third year.

The cumulative time loss derived from adding the conservative numbers above is astounding. For every 100 million users, a conservative estimate of ***711,110 total work-years***[9] were lost during the first three years the software upgrade's availability. To translate these work-years into United States dollars, the total work-years is multiplied by cost per employee per year. A cost of $110,000 US dollars fully-loaded annual headcount cost was selected.[10] Thus ***the total labor cost due to this particular file incompatibility problem was $42 billion dollars.***

This cost of lost labor does not include the upgrade purchase price, which is also substantial. To calculate the cost of the upgrade itself, the model assumes that for a group of 100 million users, it assumes 25% of these users upgraded during year one, 75% of these users upgraded by the end of

[8] A full spreadsheet model with the variable model parameters is available for download at www.censa.org.

[9] A work-year equals the amount of work one person would do in one year when working. In this example, we use 37.5 hours per week, 48 weeks per year.

[10] Fully-loaded annual headcount cost means the total annual cost of an employee to an organization. This include salary, benefits, other compensations, and other expenses on a per employee basis (utilities, rent/mortgage, office supplies, etc.). Fully-loaded costs range from $70,000 for low-wage earners in small organizations to over $225,000 for high-wage earners in large organizations. A median of $110,000 was chosen as a conservative annual headcount cost.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

the year two, and 99% of these users upgraded by the end of the year three. At a cost of $80 per upgrade, $6 billion dollars were spent the first year, $2 billion dollars were spent the second year, and $80 million dollars were spent the third year after availability. This means that the total cost of the software upgrade to the user base was about $8 billion dollars.

Adding the lost labor costs and upgrade purchase costs, for 100,000,000 users, ***the upgrade from Office'95 to Office'97 cost the user community $51 billion dollars.***

Unfortunately, the *actual* cost of incompatibility is much higher. For example, this model assumes each user lost 10 minutes per week the first year, 5 minutes per week the second year, and 1 minute per week the third year. Many users report having lost three or more days in a week on multiple occasions to resolve incompatibility problems. Similarly, although this model is based on the estimation that 99% of users have upgraded by the end of the third year, recent surveys still suggest that many large companies (as many as 7%) still have not upgraded from Office'95 to Office'97.

Additionally, while the model indicates a lost productivity cost of roughly $43 billion dollars, this does not reflect other related costs such as retraining (which is more difficult to estimate). Retraining costs are a significant cost of using software upgrades, sometimes requiring less sophisticated users to take new training courses. To include training costs, we would need to include course costs, training time costs, and lost productivity costs. Since the total cost of training includes course training costs, training time, and lost opportunity costs, the actual costs could total several times $43 billion dollars.

End users lose this time and money in business and at home. To avoid or solve the incompatibility problems, users are often forced to buy the latest release of a product. One would hope that such incompatibility problems are never intentional. Yet upgrades provide substantial revenues for vendors. So, from a vendor's point of view, incompatibility and rapid planned obsolescence are both ways to drive customers to upgrade.

Although in this research report, Microsoft Office is used as the example for modeling costs of planned obsolescence, similar analysis of any other commonly used software would yield similar results. For example, Lotus SmartSuite also produces documents for business communications. A Lotus representative reported that at least 25,000,000 copies of Lotus SmartSuite had been sold by late 1999.[11] Lotus claims that the Lotus SmartSuite market share is 10%, which would mean at least 250 million office automation suites total have been sold in recent years, including Microsoft Office and Corel's PerfectOffice. Another commonly used software, Lotus Notes, has sold at least 40 million copies total, with an annual sales "run rate" of at least 10-15 million more per year.

Taken cumulatively, the numbers we input to test the economic model above are very conservative. The model was developed primarily to show how to establish some minimum costs of rapid planned obsolescence to society. The total number of word processors sold is probably in the many hundreds of millions, so the actual costs are undoubtedly larger than those given here. When the other types of software packages are counted, the costs of rapid planned obsolescence could skyrocket to trillions of dollars by the Year 2020.

---

[11] Collected from a Lotus Worldwide Public Relations representative on 12/15/1999.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

Many software proponents would argue that instead of focusing on lost productivity and lost opportunity costs, users should really focus on the productivity gains possible from having more powerful office automation tools. Productivity gains from these automated tools are unquestionably great. It is possible to develop models to calculate financial benefits gained from improved office automation tools. Indeed, such economic models were developed years ago for word processors and spreadsheets, and have been relegated to the annals of business history No one would never think of going back to manual typewriters, scissors and tape, and correction fluids instead of automated word processing tools. The models and facts given here are meant to inspire critical consideration of business practices that routinely thrive on rapid planned obsolescence. It is important to expose some of the actual costs of planned obsolescence. There should be more and better automation -- not less -- with planned preservation available as a standard option without extreme additional cost.

10 Breaking The Status Quo: Giving Data Control Back to Its Rightful Owners

While moving data between two different versions of the same software can present migration issues, this problem pales when compared to the problems encountered when data must be moved between different products or applications. An outdated mentality persists among software vendors, that of "keeping customers by holding their data hostage." In order to create and maintain captive market share, vendors frequently define their products as the center of the data universe. In a capitalist economy, this is one way to create a successful business. Software vendors whose products are not the center of the data universe strive to compete and re-center the data universe around their own products. One way for a vendor to achieve or maintain this goal is to make it easy to import data into their software, but make it difficult to export complex data to other systems. Data export and external access to important system-level data structures such as file structures and database indexes are usually left as an "afterthought" for the next major software version (if they are added at all). However, more and more businesses increasingly rely upon many different systems to accomplish their business processes and goals. Thus, standards for data sharing and access across system space and time boundaries are becoming urgent priorities.

Obviously marketing reasons exist to prevent sharing data too easily. If data may be easily exported, then customers can more easily migrate to competitors' products. For this reason, vendors rarely dedicate their limited resources to solving export problems. Instead, niche vendors and integrators respond to these customers by providing costly custom data import and export and conversion utilities. This creates another industry that provides a continuing stream of revenue for smaller vendors and integrators as the larger vendors introduce regular upgrades and enhancements. Because most of these data integration and conversion solutions are not simple for end users to deploy or use, users must pay highly-trained, usually expensive, consultants to integrate their applications and ensure that users understand how to use them. The result is a set of endemic software industry practices that ensure that while the data created, manipulated, and analyzed by computer software belongs to the consumers, the long term access to that data is controlled by the vendors who sell the software (thereby locking up the customer base).

These problems are enhanced by the fact that vendors regularly leave the marketplace in one way or another: businesses fail, old product lines need to be retired, or a new product line is created without a practical and easy way to migrate from the old to the new product. Other vendors, in an effort to grow powerful in the marketplace, produce newer, more innovative systems before data is migrated from existing systems. Then they will provide easy import capabilities, with poor export

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

capabilities. In these cases, time and money must be spent developing custom application and data migration applications, which are quickly out of date, and may only be useful for one migration. The problem is increasingly serious because we as a society are increasingly dependent on computer systems that fail to address major record keeping issues.

The alternative today is if you want records that you create and own to be accessible long-term without continual expense, it is your financial burden to design or buy your own system to make your data system-independent, even though it was your data to begin with. However, by taking the records out of proprietary formats, the integrity of the data is frequently lost, so the risk is high that you could lose the purpose of your investment anyway.

Historically, it is easy to see the trends that led to these practices. Enormous amounts of time, money, and expertise were necessary simply to solve the technical problems standing in the way of meeting basic business needs. Today, many of the original problems have either been solved or new tools exist to help solve existing problems more quickly. A more effective way of building and keeping a large consumer base is to begin to focus on the grander needs of consumers, including long-term data access. Vendors can develop successful marketing strategies around providing the best customer quality and value, rather than proprietary tricks to hold customer data hostage.

With better infrastructure technology standards, data access will get easier and more open. The World Wide Web has the potential to revolutionize attitudes and behaviors about data ownership more then anything else in history. Many infrastructure standards exist such as the eXtensible Markup Language (XML), Java, JavaBeans, Lightweight Directory Access Protocol (LDAP), Common Object Request Broker Architecture (CORBA) services, and the proprietary Microsoft standards Common Object Model (COM) and ActiveX, MAPI, Open Database Connectivity (ODBC), and others. The problem is that these protocols change very quickly relative to the time scales of large organizations.

With the newer protocols, such as XML and Java, the "Browser Wars" are still raging. The Browser Wars are where vendors such as NetScape, Microsoft, Oracle, IBM, and others have created their own proprietary extensions to the HTML and Java languages ostensibly to enhance them. The net results have been improved protocols, with minor incompatibilities that make users' lives difficult, make website developers' lives miserable, and make HTML and Java far too unstable to be suitable for long-term data preservation and access applications.

Vendors' most often-cited reasons for not supporting particular standards are:

1) the cost (It takes scarce time and resources away from competitive features to define and implement standards. Standards are always a cost center, never a profit center for vendors.),
2) poor current acceptance (no one wants to be first to say goodbye to their proprietary technology),
3) short lifetime of most standards (especially if no one really commits to them!),
4) the need to "enhance" standards ("our ideas are better" or "there were always some unacceptable compromises made"),
5) the Not-Invented-Here syndrome (everybody wants to be the standard, nobody wants to follow someone else's standard.)

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

These reasons are sometimes legitimate; frequently they are simply competitive smokescreens. The persistence of these reasons put the onus on large corporate customers to push vendors to support existing and emerging standards for long periods of time. Buyers must finally make design for long-term preservation and access and interoperability into high-priority, baseline requirements for all new IT products for electronic records or knowledge management. Part of this Call To Action is for large organizations -- government and industrial -- to push for standards by "voting with your feet" to stop purchasing IT products that do not support long-term preservation and access of electronic data and records in application-neutral, open and published formats.

11 Avoiding the Icebergs: The Itinerary Has Been Changed

The Titanic 2020 problem is likely to get much worse before it gets better. Universities do not routinely offer courses or curricula on how to design digital systems for long-term preservation and access to data or records.[12] An entire generation of computer science and engineering professionals needs to be taught basic designs for preservation, not just innovation at the expense of history. Some records are too valuable to entrust to today's computer systems. Commercial products sadly address only a "long-term" horizon of a few years. They most certainly do not include built-in capabilities for their own replacement. The mind shift called "paper minds to electronic records" must still take place to move us all the way into the Information Age.

Practicing information technologists must develop a fuller appreciation of the disciplines and practices of record managers, archival scientists, and library and information scientists. These professionals are the guardians of corporate accountability. They are also important gatekeepers for intellectual capital, and levers for getting the most out of it. Information technologists must design acceptable technology solutions that meet the needs of these communities of business professionals.

There is an exploding need for universal digital archiving systems and formats. Universal formats for each data type (e.g., motion pictures, audio, digital certificates) are badly needed by many communities. The World Wide Web Consortium[13] is sponsoring work to harmonize semantics for web-based information that will apply generally to as many communities as possible. However, competitive wars are expected to stretch out real semantic unification for information for many more years. In the mean time, we badly need universal information container formats that allow high-quality publishing of collections of dynamic information collected from many sources. An electronic legal record is compilation of data that is recorded, signed, and sealed in an immutable form, typically in the form of documentary evidence. Thus, one of the greatest needs for electronic records has been format standards for publishing compound documents that can include text, graphics, numerical data, tables, etc.

The storage format that comes closest to a universal standard for high-quality publishing is Adobe System's Acrobat Portable Document Format (PDF).[14] PDF is a "container format" that permits capture, publishing, sharing, and preservation of complex compound documents with embedded

[12] Research and codification of the rules of system design and engineering for long-term preservation and access to electronic records are underway now in an international collaborative research project called the InterPARES Project. Courses on this subject are being developed and taught in a few leading universities and industry now. See www.censa.org/industry-interpares, www.interpares.org, or www.is.gseis.ucla.edu/us-interpares for more information. It will be a few years before these principles are routinely embedded in higher education.

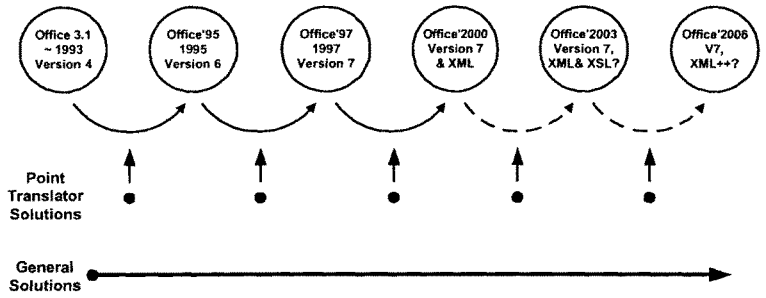
[13] See www.w3c.org for more information on the World Wide Web Consortium.

[14] For more information on Acrobat and PDF, see www.adobe.com/products/acrobat/adobepdf.html.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

digital audio, video, and other dynamic data types. PDF permits communication of these published documents across all major types of computer platforms (MS-Windows, Apple Macintosh OS, and many flavors of UNIX, including Linux) and across time. Although it does not work for all types of data, it is good enough for archiving of many types of high-quality digital documents and data. Extensions to PDF or other products still need to be developed to support permanent archiving of large commercial databases, spreadsheets, and other specialized scientific and technical data sets.

An explanation of the terms "forward compatibility" and "backward compatibility" will help to clarify the following discussion. "Forward" or "backward" always means relative to a software version or to a file format version. Backward compatibility of software means a newer software version can read and process files created by an older software version. Forward compatibility of software means an older software version can read and process files created by a newer software version. People most often desire backward compatibility of the software to read and process older files, since the software is continually updated. Forward compatibility of the software is less frequently needed, only by people who fall behind in software versions, e.g., trying to read an MS-Word'97 file with MS-Word Version 7.0 ("Word'95).



The diagram above illustrates the difference between long-term, general solutions and "point solutions" in time. Point solutions rely on translators to convert from an older version to a newer version of a file format. MS-Office includes file converters to translate nearly all of its older file formats into the latest file format, and vice versa, obviously with a loss of features that did not exist in the older formats.

General file format solutions can be read directly with any newer software versions, without using a translator, and if they are high-quality general solutions, without generating errors. General solutions are more reliable than version-to-version file translators, because version-to-version incompatibilities can cause critical information to be lost, especially embedded graphics and links, table formatting, and other software version-specific details. However, general solutions are much harder to create, because IT changes so fast and competition is so fierce. More important, it requires a tremendous amount of foresight and creativity to create products that evolve gracefully in such a fast-moving world, plus a firm business commitment to long-term compatibility. Relative to lower-tech industries, the IT industry per se is very young, immature, unstable, and competitive.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

This makes it difficult to create general solutions with high-level of fidelity and interoperability that the leading IT vendors agree to support for long periods of time. With a few notable exceptions, vendors have perceived file and data formats as a competitive weapon to use against competitors and customers.

Adobe Systems (San Jose, CA) is an example of a company with a truly unique philosophy and commitment regarding long-term product compatibility. Adobe has a broad and public commitment to provide 100% backward compatibility of the PDF format *for at least the next 25 years*. No other software vendor has such a forward-looking philosophy and commitment to its customers. An important business reason for this commitment is that Adobe is beholden to the publishing industry, which has standardized on Adobe's PostScript language and many of Adobe's other multimedia publishing products. PDF is a new and improved but compatible version of PostScript, that also supports high quality publishing on the World Wide Web. The publishing industry simply cannot afford to retool every few years. Publishers must be able to access, view, and print their archived documents and books on demand for many decades to come.

The publishing industry's practical need for a general solution that has stability and backward compatibility was recognized early by the United States Internal Revenue Service, the Food and Drug Administration, the Environmental Protection Agency, the Patent and Trademark Office, and many other government agencies. *More than 140 government agencies worldwide have adopted PDF as their standard for document submission and archiving.* Many industries have followed the lead of these government bodies and adopted PDF, primarily because it makes it easier to do business internally and with the government.

The Collaborative Electronic Notebook Systems Association and its end user member companies, which include several large pharmaceutical corporations, adopted PDF in 1997 as the preferred format for document communication and publishing and permanent storage of electronic records for three specific reasons:

1) PDF's is the defacto government standard for document publishing and archiving;
2) PDF's high quality rendering capabilities for multimedia publications; and
3) Adobe's long-term commitment and long track record for successful and complete backward compatibility, going all the way back to the version 1.0 of PostScript released in 1984.

The third point is the most important from a practical point of view. PostScript printers of 2000 can still faithfully render files created in PostScript Version 1.0 format. This track record of 15 years speaks volumes in a world where Web years are measured in weeks or months on the calendar.

With extensions to support XML, PDF can apply to all types of data published on the web and elsewhere, and for exchange and archiving of more specialized scientific and technical data types. The most recent release of PDF added security features to support signing and witnessing of PDF documents, allowing users to create self-contained, portable electronic records. Finally, an army of third-party software developers is using the open and published PDF format to extend and apply PDF to different applications and niche markets. PDF has a long future ahead of it.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

12 The Die Has Been Cast for 21st Century Information Technology

Adobe has proven to business leaders, IT professionals, and governments worldwide that stability by design, coupled with rapid innovation are not only possible, but profitable too. Most important, with Adobe's commitment to long-term stability and support, the die has been cast for end user's expectations of the IT industry in the 21st Century. This report does not mean to single out Adobe for benevolence or philanthropy -- clearly the company has benefited economically from its long-term commitments. In every age of technology there have been companies that have built fortunes by getting into new areas of business early. In the near future, suppliers who commit to stable, flexible *and* open technology designs for electronic records and data systems will have a huge competitive edge. They will profit over those who choose to maintain a perpetual blind eye to the need for stability in at least one part of their technology design -- the data or records owned by their customers. In the future, the stability of digital records, whether corporate, government, or personal, must be measured in decades to centuries rather than web years.

Innovation does not have to come at the expense of the integrity of our data or records over time. Information technologists are among the smartest and most creative people on the planet, clearly capable of creating flexible technologies that are extensible far into the future *and* backward compatible. Stability, flexibility, extensibility, and backward-compatibility are not opposing or incompatible requirements but simply additional business constraints on present and future electronic record and data storage technologies, products, and systems.

13 The Ship Has Left the Dock

The Titanic was designed to be the biggest, most technologically advanced ship of its time, and indeed it was. No one would argue that the ship failed to meet the expectations of the passengers who paid enormous sums of money to stay in her staterooms. But the fact that the ship appeared to set a new standard against which all other ships would be measured inspired too much confidence in it creators. They did not believe it necessary to make sufficient provisions for the safety of their passengers. Of all the questions about the Titanic's tragic demise, a particularly haunting one stands out: would the passengers have booked passage if they had known the boat had no sufficient exit strategy?

Like the passengers of the Titanic, it is easy for users of computer systems to believe in the marketing messages of vendors and become overconfident in the safety of their data and records. However, experience has given most modern users the chance to understand what passengers of the Titanic could not, that within the current, albeit often-impressive solutions lie dangerous pitfalls that cannot be ignored. Perhaps the market dynamic that once allowed Red Star Line to create the Titanic and now allows computer businesses to create systems used by millions of users shouldn't be broken. But the disaster that visibly looms in our future cannot be ignored. If the Titanic had had enough lifeboats for all her passengers, she would perhaps still be remembered as the biggest and most advanced ship of her time. She would not have become symbolic of the inability of man to recognize and plan for his own fallibility. Like the Titanic, current computer systems are in part a result of man's unwillingness to recognize and address the shortcomings of an industry that has mushroomed both technologically and financially. If this problem is not addressed now, in approximately 20 years, the Titanic may be replaced with a new icon of disaster.

Copyright © 2000 - CENSA, The Collaborative Electronic Notebook Systems Association
800 West Cummings Park, Suite 5400, Woburn, Massachusetts 01801 USA - Web: www.censa.org

14 Background on CENSA

CENSA, the Collaborative Electronic Notebook Systems Association, is a market development association focused on catalyzing markets for advanced automation for research and development organizations. CENSA uses unique market development programs that involve end users, suppliers, and government agencies working together to identify, specify and create solutions to market problems faster than normal methods. With over 800,000 employees in CENSA member organizations, CENSA has a broad impact across many sectors, including software, high-tech, consumer products, chemical, pharmaceutical, biotech, healthcare, government and many others. CENSA emphasizes regulated industries where intellectual property protection is key to success.

CENSA's current market development programs focus on systems for long-term preservation and access of electronic data and records, electronic notebook authoring applications, architectures and tools for component-based application and systems integration, and high-priority scientific, technical, and business components. CENSA's deliverables to its members are specifications for complete systems, component software and hardware technologies and products, and extensive specialized knowledge for setting up and using the systems and tools successfully.

CENSA sponsors the Global Industry Interagency Group on Electronic Recordkeeping and Digital Archiving Systems (GIIG on ERDAS), a non-profit organization of government agencies and industries dedicated to defining and implementing acceptable systems for long-term preservation and access of electronic data and records in industry and government applications.

15 Background on the Authors

Rich Lysakowski, Executive Director, CENSA

Rich Lysakowski is the Executive Director of the Collaborative Electronic Notebook Systems Association (CENSA). He has 20 years experience in various scientific, engineering, marketing, and project management roles in the private and public sectors. He has a Ph.D. in Physical and Analytical Chemistry, with a specialization in R&D and laboratory automation. He has done over one hundred presentations at conferences and workshops, authored over 30 publications, and edited two books. He also teaches regularly for the American Chemical Society and other professional organizations. He holds two patents on software technologies.

Zahava Leibowitz, M.A., ABD, Researcher and Writer, CENSA

Zahava Leibowitz is a researcher and writer for CENSA. She specializes in medical and technical writing, technology research, regulatory compliance and validation issues and documentation, document management systems, and project management. She has numerous publications to her name, and has been the senior editor for two books. She has over ten years experience in teaching writing and editorial strategies.

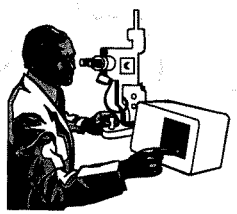
16 For More Information

For more information on CENSA and solutions to these problems, please contact:

CENSA, 800 West Cummings Park, Suite 5400, Woburn, MA 01801 USA
Phone: 781-935-9600; fax: 781-935-3113; E-mail: info@censa.org; website: www.censa.org.

The Electronic Business Asset Protection Program

Opening the Road to Long-term Asset Protection



CENSA © Copyright 2003 – All Rights Reserved Internationally

- Software buyers / users don't fully "own" the data assets they pay for...
- "Asset Access" expires due to:
  - Closed data formats used as a primary method for vendor "account control strategy"
  - Rapid planned obsolescence / change to keep revenue stream flowing to vendors
  - Product retirements
  - Vendor failures
  - Lack of diligence in migration (symptom of lack of open formats)
  - Subscription license to use software "expires"

CENSA © Copyright 2003 – All Rights Reserved Internationally

- Computers and software have been designed *only to increase productivity*
  ... not protect information assets
- Buyers don't exercise their control before purchasing software and equipment
- Government or Corporate buyers don't think they control evolution of markets
- If no standards specs exist, vendors use that as an excuse for proprietary formats

CENSA © Copyright 2003 – All Rights Reserved Internationally

- PDF provides strong insurance against information asset decay or loss
  - PDF / A gives high-quality, full fidelity rendering
  - Many agencies, laws, countries mandate PDF for recordkeeping and document submissions
- XML provides reusable knowledge components
  - XML DTDs, XML Schemas, specifications, and tools
  - Component, information, semantic models(ontologies)
- XML provides portable information objects
  - moved across applications over space and time
  - easily imported into other suppliers' software
  - Allow full reprocessing of data where needed
  - faster software design and implementation

CENSA © Copyright 2003 – All Rights Reserved Internationally

- Don't buy systems that don't properly comply with ISO PDF / Archive and XML / XSD Standards
- Our laws must mandate e-records system interoperability demonstrations in procurement process
- Withhold purchasing just long enough to secure interoperability (multiple sourcing insurance)
- Don't submit to "urgent" need to buy a system
- Take back your purchasing power in this renewed "age of accountability"

CENSA © Copyright 2003 – All Rights Reserved Internationally

- Electronic Records Systems Are Different from Traditional Systems

- If you don't plan (and budget) to preserve it...
  you won't be able to access it !!

- Thus...

"Design for Preservation"

is a new critical systems design center

CENSA © Copyright 2003 – All Rights Reserved Internationally

CENSA © Copyright 2003 – All Rights Reserved Internationally

Additional Material for the Congressional Record from Dr. Richard Lysakowski, Executive Director of the Global Electronic Records Association (GERA), President of the Collaborative Electronic Records Association (CENSA)

Witness on Expert Panel

House Government Reform Committee, Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census, Meeting on July 8, 2002

Submitted July 28, 2003

1 Followup Written Points for the Congressional Record:

I was asked to provide additional clarification points, research results and further recommendations to the Subcommittee. I respectfully submit the following points for consideration by Congress.

1.1 NARA Alone Cannot and Should Not Be Expected to Solve the USA's Problems with Electronic Records.

Electronic records are not uniquely the problem of the US National Archives and Records Administration (NARA). Electronic records are integral to all eGovernment work processes in all federal, state, and local agencies. Several times over the past five years, private industry has approached OMB to develop or mandate common government-wide and industry-wide solutions and standards for electronic records. Everytime, OMB has pushed all responsibility for electronic records onto NARA's shoulders, now calling NARA the "Managing Partner for the Electronic Records Management (ERM) e-Gov Initiative." This has slowed badly-needed progress.

NARA's ERM Initiative's stated goals are to provide a government-wide policy framework and to provide guidance for electronic records management. NARA's mission, scope, focus, size, budget ($268M FY2003) do not permit it to solve all electronic records problems of federal, state, industrial, and nonprofit organizations. NARA's primary mission remains to "ensure continuing access to essential evidence that documents: 1) the rights of American citizens, 2) the actions of federal officials, 3) the national experience." NARA takes stewardship of only a very small percentage of federal and other historical records for long-term preservation and access.

NARA is underfunded and understaffed to develop scalable solutions and standards for electronic records for all federal, state, and local government agencies. Thus industry is left to wait for open or de facto standards to emerge or to create its own. More often than not it is a waiting game because industry standards must be compatible with government regulations for recordkeeping. At times, industry creates de facto standards that governments adopt, like AT&T's UNIX, IBM's PC, and Adobe's Portable Document Format (PDF). However, at this point IT innovation is being stifled across the board by the lack of standards for electronic records and data formats. If federal agencies were to simply adopt or drive archival formats and interoperability standards, industry would adopt them, implement them, and drive faster creation of robust markets and technical services to support them. It does not matter who leads, but it is important that someone does.

A great example of what the US Government can do for itself and industry with a moderate investment is the US DoD's 5015.2 Standard for Records Management Applications. Once DOD

CENSA © Copyright 2003 – All Rights Reserved Internationally

created and mandated this standard, it was adopted and used as a baseline for standards throughout many industries, including pharmaceutical, chemical, food, medical devices, aerospace, automotive, and others. It remains a baseline against which all good records management technology systems are measured and certified. The total budget to create DoD 5015.2 was estimated to be under $5M USD and it took less than three years time to develop. Industry spent at least ten times that amount of money to develop DoD-compliant solutions. Such a large industrial investment would have been needed to meet any specification where technology changes were needed.

DoD 5015.2 standard did not go far enough and specify standards for the entire lifecycle of electronic records, including requirements and specifications for records creation, records management, archival functions, preservation functions, retrieval and reproduction functions, and file format standards for individual document-based records and collection of records. This is excusable because the scope had to be limited so that DoD could deal with known problems for technological solutions existed in 1995.

We should take advantage of the fact that The Department of Defense has the power to create and mandate standards faster than any other government body and there is frequently spillover of benefits directly to the civilian sector and the commercial markets.

1.2 Government Administrative and Regulatory Bodies Drive Industry, Not NARA.

Industrial and private sector problems with electronic records are driven by factors not directly related to NARA's mission. For example, the private sector is driven by many concerns simultaneously, e.g., the USPTO drives concerns for intellectual property protection, the EPA drives concerns for environmental protection and quality, the HHS / FDA drive concerns for healthcare product safety and quality, the SEC drives concerns for accurate financial reporting and integrity. For transportation, the FAA flight safety regulations mandate long-term availability of design, test, calibration and maintenance records for commercial and military aircraft in service for longer than 20 years. For automotive transportation safety, records are needed for product design, testing, maintenance and recalls for two or more decades. In the life, medical, disability insurance industries records retention periods are anywhere from 15 to 90+ years. Mortgage and banking industry mandate retention 15, 30 years (or longer). DOL ERISA-compliance and Social Security Administration pension benefits plan record retention periods cover the working life of all citizens and beyond. These are just a few of many immediate concerns with long-term impact on the private sector. All these industries are trying now to use electronic records whenever possible.

Problems of electronic records are multifaceted – driven by administrative, legislative, legal, regulatory and business concerns – but solved via well-designed and implemented organizational programs, qualified people, policies, quality standards and procedures, and technology systems. CENSA and GERA have partnered industry and government agencies to create the set of standards called "the Quality Electronic Recordkeeping Practices" (QERP) standards that articulate all the required program and technology elements to design, staff, operate, audit, and maintain electronic evidence over any time frame. While the need for keeping evidence of activities is universal throughout the public and private sectors, the scope, scale, values and accountability reasons for recordkeeping vary so widely that NARA's solution and approach will not work for all.

CENSA © Copyright 2003 – All Rights Reserved Internationally

1.3 NARA's ERM eGov Initiative – Where's the eArchive in each eGov Agency?

NARA's current strategy focuses on assisting other federal agencies with electronic records management initiatives by providing a methodology for determining agency-unique requirements on top of the US DoD 5015.2 RMA Standard. While this approach takes good advantage of the excellent work of the DoD standard – which NARA provided significant help to develop – it does not directly address the need for each federal agency to set up and run their own electronic records archive. It does not lay down the standard for electronic records archives to successfully deal with the 99% of agency records that NARA never receives.

NARA's Electronic Records Archive (ERA) solution is being designed to solve NARA's immediate technological needs for a system. The NARA ERA is a "custom" system designed to NARA's specifications. NARA's ERA is not being designed to be a general-purpose, "Configurable-Off-The-Shelf" (COTS) product designed to scale up to large agencies like the Department of Defense or Homeland Security, or scale down to much smaller agencies at federal, state, and city levels. From my point of view, this is a wasted opportunity to mandate standards and build markets.

1.4 A More Coordinated, Multi-Agency and Industry Approach is Needed

For eGovernment to work, a more universal, distributed approach is needed that does not put NARA at the center of everything. Sole reliance on NARA for guidance and standards could cause a bottleneck in moving to electronic recordkeeping in eGovernment. A universal, large-scale approach can be developed with NARA as a leader, follower, and participant. However, some fundamental strategy shifts must occur so NARA does not constrain processes it seeks to facilitate.

I recommend that a small coordinated team of US agencies and industry with a large stake in preservation of digital information assets be assembled to make progress much faster than thus far. I recommend NARA, LOC, The Smithsonian Institution, GAO, NASA, DoD, HHS, and industry work together. NARA's expertise is crucial and it cannot solve everyone's problems.

NARA is too small to cause the kind of change that we need within a reasonable timeframe. Its small budget (est. $289M FY-2004) is seriously constraining. Other US agencies have resources that greatly exceed those of NARA. For example:

- The Library of Congress (est. $416M FY-2004)
- The Smithsonian Institution (est. $478M FY-2004)
- General Accounting Office (est. $474M FY-2004)
- NASA (est. $15B FY2004)
- DoD (est. $389B FY2004)
- HHS (est. $539B FY-2004)

These other agencies have their own internal needs for total lifecycle solutions for electronic records (including long-term preservation and access) that are similar to NARA's internal needs.

Several agencies working together would push for faster and better solutions to the digital records problem much better than NARA alone. A focused team of large US government departments and agencies with the most at stake should be assembled to solve this problem. The Library of Congress' Digital Library program includes digital preservation and is succeeding. NASA has preservation needs for massive digital data sets that record the weather, the condition of the earth,

CENSA © Copyright 2003 – All Rights Reserved Internationally

space science, and defense-related issues. HHS scientific and medical research, development, and regulation are being slowed by the lack of standards for digital data and record interchange. *These few, but powerful agencies and departments could collaborate as a focused team to create and mandate solutions in three to five years, rather than 20.*

1.5 Leverage The Success of Other Countries

Other countries have succeeded by working on a scale appropriate to the size of their agencies.

Australia in particular has much we can learn from and adopt. The Australian success has led US experts to acknowledge that scalable, long-term solutions can be built using today's technologies. They solved electronic records management and archives problems on a smaller scale than the US's problems, but one that can be extrapolated for the USA. It too is meant to serve as a core building block of their eGovernment initiative. The system was developed to first fit the State of Victoria (11M citizens) and then fit the Whole of Australia (26M citizens). The solution is commercially available now and can adapted to fit a wide range of needs of agencies, states, and industries. US agencies and industry worldwide should look to leverage as much as possible from Australia. I have studied their system in detail and visited Australia twice to see it in action. It is clear that their system is fully compatible with US NARA's proposed ERA system.

The State of Victoria Electronic Records Strategy (VERS) Program started in about 1995. The Victoria State Department of Infrastructure led a cross-disciplinary team of Australia's leading government archivists, records managers, computer scientists and software engineers, and leading software vendors. It uses XML and the Adobe Portable Document Format (PDF) for electronic record archival collection storage and retrieval. Their initial goal is to retrieve and reproduce authentic electronic records in an XML/PDF format for 100 years minimum (with or without Adobe Systems, Inc. in existence.) The VERS program includes organizational program, policy, technology system design standards, training, and government-wide dissemination projects. An important point to note is that they reproduced the PDF Viewer software completely on their own (with no help from Adobe) from the written specification as a "digital archaeology experiment" to prove to themselves that could reliably reproduce PDF-based documentary records.

The VERS team of experts researched requirements, prototyped systems using available commercial software as the starting point, tested the solution in pilot projects for many months, and put the final system into production in 2001. The Victoria Government has now mandated compliance by law for all state, city, and local agencies for the software system functionality, PDF documentary record file format, and XML electronic record collection format. File formats other than PDF can be embedded with the PDF preservation copy if full reprocessability is required.

The VERS Program Office has established a Centre for Excellence for Electronic Records that is charged with disseminating the technology solution and training on programs, policies, and procedures to all the other government departments. They have budgeted money to certify enough competing vendors so that they have a set of flexible, scalable technology solutions that will fit different government department needs.

An important point is that the Australian Government removed a barrier to market development (and their own success) by paying for vendor and product compliance testing and certification themselves. In the US, to get a product certified for DoD 5015.2 standard compliance each vendor

CENSA © Copyright 2003 – All Rights Reserved Internationally

must pay $10,000 or more for each major product version. Many vendors have complained that this is excessive when it must be repeated for each major product release. In the case of Australia they can keep costs of re-testing compliance and recertification by evolving their record and file collection archive formats prudently using XML's schema versioning and version management tools.

By 2007, the plan is to disseminate the program and technology solution to all departments of the State of Victoria, and likely to the Whole of Australian Government. The total program budget from 2001 through 2007 is expected to be $90M Australian Dollars. It is startling that Australia could make so much progress so fast, starting after but finishing before the USA, and using less than 1/10 the money. Australia is about 1/10 the population of the USA.

1.6 Avoiding Great Waste and Expense to Society

We must create (or adopt) and mandate standardized solutions soon. We will continue to buy and sow the seeds of our own destruction until then. In a December 1999 research report[1], subtitled "Titanic 2020" I gave examples of the hidden costs to society for the lack of standards. This report is available at " http://www.censa.org/html/Publications/Titanic2020.htm ". Using the example of one generation of word processing documents alone, I *conservatively* estimated the cost of the lack of electronic document interchange standards to society was $52 billion USD. I did not assess economic impact for other documentary formats. Without standards for information asset preservation across technology generations, every new system that we install becomes the nucleus of an iceberg that we will either hit and destroy assets, or pay excessive costs to avoid asset loss.

1.7 Subscription-Based Software Is A Huge Threat to Record Security and Assets

Software products are tools used to generate personal and business property (data, information, and knowledge). The usual model of software product sales is a perpetual, right-to-use license on a limited number of computers for a fixed price. Like all tools, software depends on having quality product manufacturers. Software vendors would love to change the way that software is purchased to transform software into a service or utility like water, electricity, gas or oil that is consumed and must be re-paid for periodically – either monthly, annually, or some other period. Microsoft was the first major large software vendor to push for subscription-based pricing models. Advocates for subscription-based software licensing (mostly vendors) argue that it eliminates hassles with software updates, because updates can be done automatically as part of the subscription.

The problem with this is method of software purchasing is that if you do not pay your bill, your software can be "deactivated", "time out", shut off, or somehow made unusable, thereby denying you access to data and documents that you created.

There are federal and state laws that prevent denial of access to one's own property. I question the legality and ethics of this method of selling software without giving buyers a no-cost exit strategy in the form of easy migration. Having to pay a subscription to access your own property is similar to extortionary tactics used by the Mafia to shut down businesses or take people's livelihood and property hostage until regular "payments" are made. Records are corporate, government, or

[1] "Looming Information Age Crisis Expected To Cause Trillion-Dollar Losses Over Next 20 Years" by Dr. Richard Lysakowski and Zahava Leibowitz, CENSA, Inc., December 1999.

CENSA © Copyright 2003 – All Rights Reserved Internationally

individual property – one's access to them must never be at risk of "deactivation." Subscription-based pricing models being foisted upon unwary consumers give vendors freedom to use abusive pricing or unethical business practices.

The best way around this potential loss of access to one's property is to require and buy open data and record portability standards with perpetual, right-to-use licenses for data schemas and file formats. There are too many cases of lost data assets because undocumented, proprietary file formats were not migrated before they became obsolete, the vendor "upgraded" the file format to include incompatible features, the vendor dropped that product line, the vendor's business failed, or the business was taken over by another company that stopped supporting the format. Buying perpetual rights to use data schemas and file formats avoids all of these problems because it permits easy migration by the property owner, the original vendor, or any third party vendor.

We do not need privatized "software utilities" or software sold as a service. We need more high-quality software product manufacturers that do not use file formats as a way to control their customers' property. We are in this quandary because buyers have not insisted on open, published formats as an absolute condition of purchase of software.

1.8 Lack of Government-Mandated Record Storage Format Standards is Holding Back Innovation and Costing Taxpayers Big Money

Many US government bodies adopted PDF as a document transmission and preservation standard years ago (FDA, the US Courts, EPA, FAA, IRS, and others). Slow formal acceptance of PDF as an official record storage format by the US Federal Government as a whole harms innovation.

Some compelling statistics about PDF are:[2]

- 2400+ government agencies worldwide use PDF for document interchange.
- 500 million freely-licensed copies of PDF Reader have been downloaded by public and private sector individuals. Last year over 50,000 copies were being downloaded per day.
- Over 2,000,000 official publications exist on US government external websites.
- Over 675,000 links exist worldwide to download the PDF Reader from sites all over the world.
- Over 1800 PDF tools developers are creating add-ons and tools that compete directly with Adobe.
- Adobe Reader comes preloaded on PCs from the top 10 PC manufacturers.
- The Government of Germany just signed a country-wide deal for PDF and Acrobat and has mandated the use of PDF for interchange format for interagency communication.
- Holland, Italy, and Australia have written PDF and XML into law as mandatory for archival submissions of high quality documentary records.
- Japan and the European Union (EU-EMEA) have specified PDF for all electronic Common Technical Document (eCTD) supporting documents; an XML backbone is used to relate all files together for new applications for approval of new medicines and medical devices.

[2] Obtained the market research department at Adobe Systems, Inc.

CENSA © Copyright 2003 – All Rights Reserved Internationally

While the numbers are large, they make an important point, that when a technology that is necessary and sufficient exists, adoption as a de facto standard proceeds quickly. Adoption speeds innovation, saving government and industry huge amounts of money.

Money-savings tied to the adoption of PDF is well illustrated by the US Bankruptcy Courts. The US Bankruptcy courts used PDF for all court records for the Enron, Worldcom, and Global Crossing bankruptcies. The US Courts admitted that they did not have to add one new clerk, but estimated that if they had to use paper records, they would have had to add 50 new court clerks.

1.9 Risk Assessment in Section 300 of OMB Circular No. A-11 (2002) Does Not Detail or Score the Asset Protection Requirements

OMB must specify a procurement process for eRecords systems that requires demonstration of interoperability and complete independence of supplier data and file formats before purchasing. Product independence means full capabilities to migrate one's property away from one vendor's system to another without additional expense. With electronic records and data formats, this is a good way to provide an insurance policy against rapid planned obsolescence by computer hardware and software vendors.

OMB Circular No. A-11 Section 300 (2002) Planning, Budgeting, Acquisition, and Management of Capital Assets, is incomplete and insufficient for information asset protection. This OMB Circular A-11 must separately articulate and quantitatively score all US government IT system procurement processes for "interoperability" and "preservation capabilities." These critical aspects of IT systems are buried in the discussion on "Risk Inventory and Assessment (All Assets)" in Part 1, Section F on page 20 and 21. These should be separately scored as "Asset Protection and Preservation" and "Interoperability and Information Interchange."

Without an explicit scoring system for information technology system capabilities for "asset protection and preservation" and "interoperability", US Government procurement processes will not protect government, industry, or taxpayer property (records). We continue to put everyone's assets at high risk, and pay more than necessary to retain ownership and access to them.

1.10 The US Needs Better Procurement Processes for Protecting Record Assets

Given that OMB's A-11 Circular does not explicitly measure and quantitatively score IT system information asset preservation, protection, or interoperability, one can conclude that procurement processes will not sufficiently protect agency or taxpayer assets over long periods of time. This is not unique to OMB, but is a prevalent buying practice throughout government and industry.

All IT system buyers' behavior must change to require as a pre-condition of purchase:

1) full portability and reuse of records, and
2) permanent access to data and records generated by the buyer from all vendor's software.

As long as government and industry do not insist on these pre-conditions for all purchases, they will continue to lose valuable assets. This has been the status quo for decades.

See slides in Appendix for more information on why the market status quo is the way it is now.

CENSA © Copyright 2003 – All Rights Reserved Internationally

1.11 A Simple Buying Strategy Exists to Protect Business Assets

From watching some of the smartest companies buying software products, we recommend a simple, logical strategy and process for buying IT systems so that business owners' assets are protected. The process is explained here and illustrated in the next section. The Appendix is a handout that explains why the status quo exists and what to do about it.

The strategy is simply "Require suppliers to demonstrate full data interoperability before you buy." Put another way "only buy software products from suppliers that will ensure your access to your information assets without them in the future."

This buying strategy is not very popular with suppliers that like to use closed file formats as a way to control customer migration. However, this brute force method of account control results in many lost customer assets, because migrations are much more difficult. More responsive vendors are competing at levels above file formats, i.e., on product quality, functionality, and customer service.

It is important to distinguish between closed, proprietary formats and open, documented proprietary formats. Closed formats are not freely available for anyone to read and write. Proprietary formats are not bad to use for records as long as they are open, documented, perpetually and freely available for anyone to use without license fees. With the advent of XML, and widespread usage of PDF for documentary records, there is no longer any excuse for buying software that does not come with open, documented formats.

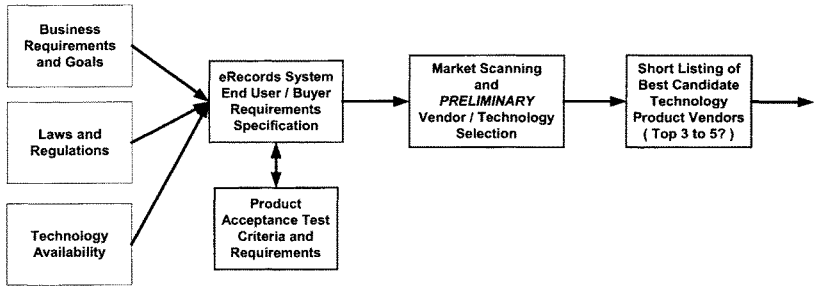
A well-designed electronic records archive is a fixed collection of information structured, stored, and secured in such a way that permits easy reconstruction and retrieval in another records management system, whether current or future system. If a system cannot do this, then the system designers not include long-term asset preservation and access as a design center. This decision by designers is either an oversight or made for competitive reasons. To achieve full electronic records ownership and access one must be able to export "Archival Information Packets" that preserve the full content, structure, context, and presentation of records. One must also be able to verify the authenticity of the entire collection of records and individual records within the Archival Information Packet after it has been migrated to another system.

CENSA © Copyright 2003 – All Rights Reserved Internationally

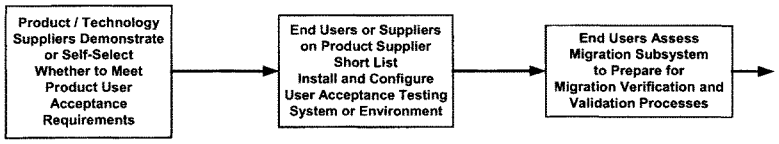
1.12 The Process for Buying IT Systems To Protect Business Assets

The four phases of this recommended process are illustrated below. "Supplier candidates" must pass each phase in succession. This process is particularly critical for Electronic Recordkeeping Systems, but is also recommended for buying any type of data or information system.

Phase 1 - Articulate Requirements and Find Qualified Suppliers.
Be sure to use a User/Buyer Driven Process with Input From Suppliers.

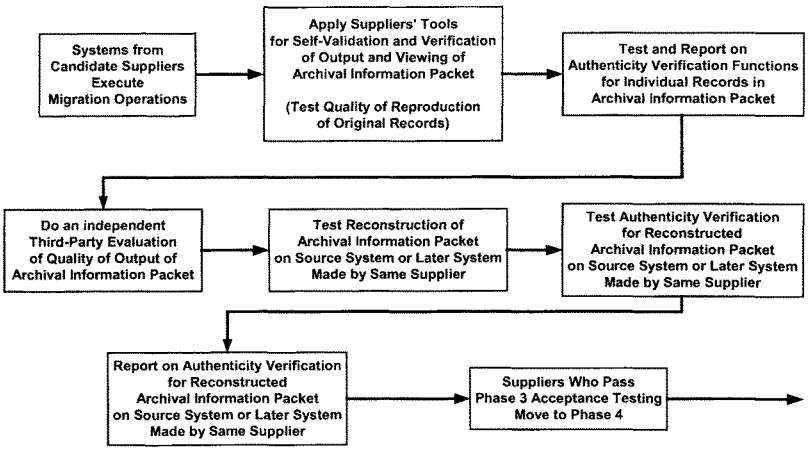


Phase 2 - Work With Qualified Suppliers to Prepare User Acceptance Testing Systems / Environment

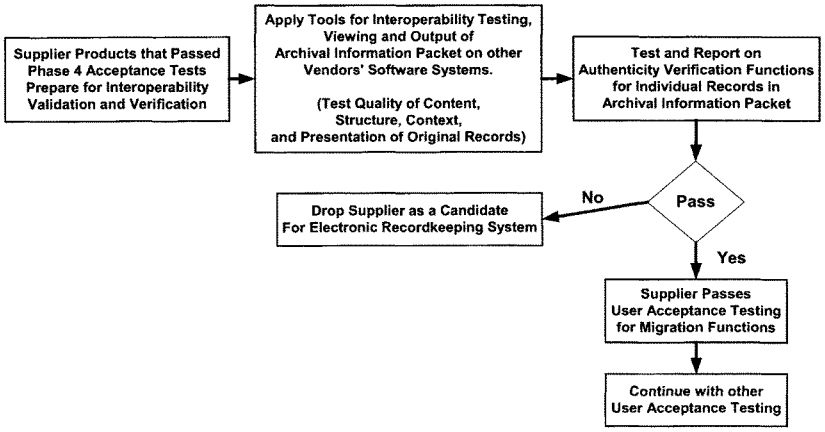


CENSA © Copyright 2003 – All Rights Reserved Internationally

Phase 3 - Test Quality of Output of Migrated Archival Information Packets and Authenticity of Records



Phase 4 - Test Quality of Interoperability of Migrated Archival Information Packets



CENSA © Copyright 2003 – All Rights Reserved Internationally

Appendix – Business Asset Protection Program

We have studied the problem of electronic business asset protection and have found it to be primarily the problem that buyers:

- do not make it a priority to own permanent access to file formats for all software product licenses they purchase,
- do not exercise their purchasing power,
- do not take responsibility for migrating their information assets.

The status quo is caused by other reasons too. However, changing procurement practices to include testing and scoring for asset preservation, migration, and interoperability will fix the problems. In some cases, they may be required or decide to pay extra to purchase the rights to permanent access to their assets. In the long run, however, it is worth the cost, because problems with asset loss will be greatly reduced.

It is up to buyers to change the status quo and set up their own Business Asset Protection Program.

The Electronic Business Asset Protection Program

Opening the Road
to Long-term Asset Protection

- Software buyers / users don't fully "own" the data assets they pay for...
- "Asset Access" expires due to:
  - Closed data formats used as a primary method for vendor "account control strategy"
  - Rapid planned obsolescence / change to keep revenue stream flowing to vendors
  - Product retirements or failures
  - Vendor business failures
  - Lack of diligence in migration (symptom of lack of open formats and design for preservation)
  - Subscription license to use software "expires"

CENSA © Copyright 2003 – All Rights Reserved Internationally

- Computers and software have been designed *only to increase productivity* ... not protect information assets
- Buyers do not exercise their control before purchasing software and equipment
- Government or Corporate buyers do not think they control evolution of markets
- If no standard format specifications exist, vendors use that as an excuse for closed, proprietary formats

- PDF provides strong insurance against information asset decay or loss
  - PDF / A gives high-quality, full fidelity rendering
  - Many agencies, laws, countries mandate PDF for recordkeeping and document submissions
- XML provides reusable knowledge components
  - XML DTDs, XML Schemas, specifications, and tools
  - Component, information, semantic models(ontologies)
- XML provides portable information objects
  - moved across applications over space and time
  - easily imported into other suppliers' software
  - Allow full reprocessing of data where needed
  - faster software design and implementation

CENSA © Copyright 2003 – All Rights Reserved Internationally

- Don't buy systems that don't properly comply with ISO PDF / Archive and XML / XSD Standards
- Our laws must mandate e-records system interoperability demonstrations in procurement process
- Do not submit to "urgent" need to buy a system
- Withhold purchasing just long enough to secure interoperability (multiple sourcing insurance)
- Buyers must take back their purchasing power in this renewed "age of accountability"

- Electronic Records Systems Are Different from Traditional Systems
- If you don't plan (and budget) to preserve it... you won't be able to access it !!
- Thus...

"Design for Preservation"
is a new critical systems design center