The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project

Table of Contents

Acknowledgements

Introduction

PART ONE — establishing and maintaining trust in electronic records

Authenticity Task Force Report

Introduction
Basic Assumptions of the Research
Research Design and Methodology
Research Findings
Relationship between Conceptual Requirements for Authenticity and Existing Standards
Relationship of Findings to Other Research Initiatives
Conclusion
Areas for Further Research

PART TWO — choosing to preserve: the selection of electronic records

Appraisal Task Force Report

Introduction
Basic Assumptions of the Research
Research Design and Methodology
Research Findings
Relationship of Findings to Other Research
Conclusion
Summary of Recommendations
PART THREE — trusting to time: preserving authentic records in the long term

*Preservation Task Force Report*

- Introduction
- Research Design and Methodology
- Findings
- Products
- Relationship to Existing Standards
- Conclusion
- Recommendations

PART FOUR  i) — an intellectual framework for policies, strategies, and standards

*Strategy Task Force Report*

PART FOUR  ii) — contextualization

*National and Multinational Team Reports*

- Australian Research Team Report
- Canadian Research Team Report
- Chinese Research Team Report
- Global Industry Research Team Report
- Italian Research Team Report
- National Archives of the Netherlands Report
- United States Research Team Report

PART FIVE — InterPARES terminology

*Glossary Committee Report*

Conclusion

APPENDICES
Appendix 1—Template for Analysis
Appendix 2—Requirements for Assessing and Maintaining the Authenticity of Electronic Records
Appendix 3—Appraisal of Electronic Records: A Review of the Literature in English
Appendix 4—A Model of the Selection Function
   i) Model Diagrams
   ii) Activity Definitions
   iii) Arrow Definitions

Appendix 5—A Model of the Preservation Function
   i) Model Diagram Descriptions
   ii) Model Diagrams
   iii) Activity Definitions
   iv) Arrow Definitions

Appendix 6— How to Preserve Authentic Electronic Records
Appendix 7— Walkthrough Applying the "Preserve Electronic Records" Model

Glossary
Acknowledgements

The project director wishes to thank first and foremost the Social Sciences and Humanities Research Council of Canada, the University of British Columbia, the United States National Historical Publications and Records Commission, the Italian National Research Council, the Italian Central Office for Archival Properties (specifically the Documentation and the Publication Offices), the Italian Institute of Culture in Vancouver, the National Archives of Canada, the United States National Archives and Records Administration, and all the other institutions and organizations that have provided financial support to the InterPARES Project and that are identified in the Introduction. Special thanks go to the School of Library, Archival and Information Studies for providing the office space for the headquarters of the project. The director also wishes to give the most heartfelt thanks to the following InterPARES Project staff for their deep commitment to the research and the researchers, which was manifested not only in long working hours and constant enthusiastic availability, but also in a flurry of creative ideas that have contributed significantly to the good management of the project: coordinators Tahra Fung (University of British Columbia), Kevin Glick (University at Albany, State University of New York), and Jean-Pascal Morghese (University of British Columbia), former coordinators Peter Van Garderen (University of British Columbia) and Richard Sloma (University at Albany, State University of New York), and project assistant Sandy Orr. Special thanks go to the Post-doctoral Fellow, Brent Lee, who has provided invaluable managerial support to the direction of the project and intellectual contribution to the research, and whose expertise has opened the way to the second phase of the project. Of course, the heart of a research project is the research team. The project director wishes to thank profusely all InterPARES researchers for their dedication to the project and intellectual contribution to the research, and whose expertise has opened the way to the second phase of the project. Of course, the heart of a research project is the research team. The project director wishes to thank profusely all InterPARES researchers for their dedication to the project, and in particular the chairs of the task forces—Heather MacNeil, Terry Eastwood, and Ken Thibodeau—and of the Glossary Committee, Ken Hannigan, for their outstanding leadership: The Long-Term Preservation of Authentic Electronic Records: Findings of the InterPARES Project is their work and speaks by itself of the remarkable intellectual effort that everyone has put into the project and of the wealth of new knowledge so generated. It is not the practice of a project director to single out individual researchers, but the exceptional contribution of the archivists of the National Archives of Canada (NAC) cannot go without special recognition: the amount and quality of their work in carrying out case studies, disseminating research results, providing constant feedback, and enriching the research workshops with their experience and creative ideas have been outstanding, and their commitment to the international collaboration has been unfailing. The names of the NAC archivists who were official members of InterPARES are listed in the Introduction; the names of those who constituted the NAC internal InterPARES research team are: Cathy Bailey, Dave Brown, Jana Buhlmann, Marnie Burnham, Cara Downey, Andrew Horrall, and Paul Marsden. Finally, warm thanks go to the student research assistants whose enthusiasm and dedication have given the research that sparkle that scholarship by itself can never achieve: it is hoped their future careers and lives will be enriched by the work they did for InterPARES as much as the InterPARES research and researchers have been enriched by their unique individual contributions.

The individual task forces wish to thank the InterPARES Project staff for their unfailing support. Special thanks go to the following research assistants: Chaja Beck, Gijs Boon, Mirjam Brouwer, Henk Duits, Ingmar Koch, Jacqueline Stroet, Robert van Vuuren, Marianne Vos (Netherlands Institute for Archival Education and Research); Andrew Ashton, Rebecca Hatcher, Irene Kaplan, Michelle Powers, Mark Wolfe (University at Albany, SUNY); Tiran Behrouz, Lisa Beitel, Kathy Bossort, Deirdre Bryden, Kathleen Burns, Robert Edwards, Anna Gibson, Prisca Giordani, Elaine Goh, Robyn Hulley, Shadrack Katuu, Jacqueline Killawee, Marta Maftei, Ian McAndrew, Shauna McRanor, April Miller, Alex Richmond, Sion Romaine, Claire Vesseire, Jeff Whyte, and Lara Wilson, Joleen Wright, Jane Zhang (University of British Columbia); Francesca Marini, Eun Park, Marisol Ramos, Shelby Sannet, Kalpana Shankar, Ciaran Trace, Julio Vera (University of California, Los Angeles); Melissa M. Terras (Humanities Advanced Technology and Information
Institute [HATII], University of Glasgow; Flora Anastassiou, Vincenzo De Meo, Susanna Oreffice, Silvia Trani (University of Rome); and Monica Grossi (University of Urbino).

The Authenticity Task Force would also like to thank the following institutions, organizations, and their staff for their assistance in conducting the case studies: Academic Medical Centre, Utrecht; Banff Centre for the Arts; Bank of Italy; Canadian Intellectual Property Office; the Government of Canada's Department of Indian Affairs and Northern Development, Department of National Defence, and Department of the Solicitor General; Lepera & Ward Architectural Firm; Ministry of Housing, Spatial Planning and the Environment, Government of the Netherlands; Ministry of Justice, Government of the Netherlands; Municipal Environment Agency, Rome; Nanjing University; the National Archives of Canada; National Bank of Labour; New York State Workers’ Compensation Board; Province of North Holland, Haarlem; the Smithsonian Institution; the United States National Archives and Records Administration; the United States Patent and Trademark Office; the University at Albany, SUNY; the University of British Columbia; the University of California, Los Angeles; and the University of Glasgow.

The Authenticity Task Force acknowledges gratefully the following agencies and institutions for the resources they have invested in the project: Bank of Italy; Beijing Municipal Archives; the Central State Archives of Italy; the National Archives of Canada; the Netherlands Institute for Archival Education and Research; the Smithsonian Institution; the State Archives of China; the State Archives of Rome; the United States National Archives and Records Administration; the University at Albany, SUNY; the University of British Columbia; the University of California, Los Angeles; and the University of Glasgow.

The Appraisal Task Force would like to thank the following institutions for hosting its workshops or arranging meetings to communicate the results of its work and to solicit comments: the Hong Kong Public Record Office; the National Archives and Records Administration; the National Archives of Canada; the Netherlands Institute for Archival Education and Research; the State Archives of Rome and Italian Central Direction of Archives; the University of British Columbia, School of Library, Archival and Information Studies; and the University of Toronto, Faculty of Information Studies.

Special thanks go to Yvette Hackett of the National Archives of Canada for her help at the May 2001 meeting of the task force in Ottawa.

The Appraisal Task Force gratefully acknowledges the following agencies and institutions for their support: the East Asian Section of the International Council on Archives (EASTICA); the Faculty of Information Studies at the University of Toronto; the National Archives and Records Administration; the National Archives of Canada; the National Archives of Ireland; the Netherlands Institute for Archival Education and Research at the University of Amsterdam; the State Archives of China; and the State Archives of Rome and the Central Direction of Archives of Italy.

The Preservation Task Force would like to thank the Collaborative Electronic Notebook Systems Association; Curtin University of Technology; Georgia Tech Research Institute; Johns Hopkins University; National Archives of Canada; National Archives of France; National Archives of Italy; National Archives of Sweden; National Archives of The Netherlands; the Public Records Office, U.K.; San Diego Supercomputer Center; University at Albany; University of California, Los Angeles; University of British Columbia; and the U.S. National Archives and Records Administration. The task force would like to add Michael Skipper of the National Archives and Records Administration to the list of those deserving special mention.

The Strategy Task Force would especially like to thank research assistant Ian McAndrew (University of British Columbia).
Introduction

"The Long-Term Preservation of Authentic Records" project, commonly called InterPARES (International Research on Permanent Authentic Records in Electronic Systems), was launched as a result of the growing interest in the findings of the Preservation of the Integrity of Electronic Records research project, commonly called the UBC Project, which was carried out at the University of British Columbia’s School of Library, Archival and Information Studies from April 1994 to March 1997. The UBC Project defined the requirements for creating, handling, and preserving reliable and authentic active electronic records (i.e., records that are used daily by the body producing them in the regular course of business). The UBC Project's researchers, Terry Eastwood and myself, and the project's research assistant, Heather MacNeil, worked in close collaboration with the United States Department of Defense Records Management Task Force to identify requirements for records management application (RMA) software.¹

The research following the UBC Project had to address the long-term preservation of authentic inactive electronic records (i.e., records that are no longer needed for day-to-day business but must be preserved for operational, legal, or historical reasons). The rapid obsolescence of computing technologies was creating difficulties for those concerned with the long-term preservation of documents in digital form. This was especially true for documents that are records, since they serve as indispensable instruments of accountability, a means of protecting individual and corporate rights, and as sources of information for future generations. The preservation of records created in electronic systems was thus posing a critical challenge for present and future archivists.

As archival institutions around the world sought to develop strategies to effectively preserve the electronic records created by the governments, corporations, and other organizations that they serve, the potential need to migrate these records across technologies was raising questions related to the records' authenticity. On what basis can electronic records be presumed authentic when they come into archival custody? What steps need to be taken by their preservers to ensure that the means used to ensure their long-term preservation and accessibility do not compromise their authenticity? The immense scope and ubiquity of these and related issues surrounding the long-term preservation of authentic electronic records made evident the need for an interdisciplinary, international approach.

From January 1999 to December 2001, the InterPARES Project, led by myself, sought to address a broad range of questions surrounding the authentic preservation of inactive electronic records. The InterPARES researchers were organized in national and multinational teams for the double purpose of securing funding from separate sources and of acting as research units sharing a common juridical-administrative context, the responsibility of which was to refine the findings of the project—as they were developed—in light of the requirements of their own environment. The latter purpose was especially important in a project whose collaborative international multidisciplinary nature required that the groups formed to investigate the research domains comprise co-investigators of diverse cultural and scholarly backgrounds searching universal solutions to common problems. The teams were: the Australian Research Team, the Canadian Research Team, the Chinese Research Team, the European Research Team, the Global Industry Research Team, the Italian Research Team, and the United States Research Team. To determine the overall direction of the research and steer it as necessary, an International Team

was created comprising representatives of all the research units and all the participating institutions, and chaired by myself. The International Team met three times each year for the duration of the project. This face-to-face exchange of views and discussion of research findings promoted a highly collaborative research process, allowing for the findings to be the truly integrated results of a common effort, products profoundly shared and believed in by all InterPARES researchers.

The stated goal of the InterPARES Project was to develop the theoretical and methodological knowledge essential to the permanent preservation of electronically generated records and, on the basis of this knowledge, to formulate model strategies, policies, and standards capable of ensuring their preservation. To meet this goal, the research was organized into four domains of inquiry; a corresponding task force was formed to address the research questions specific to each domain.

The objective of Domain 1 was to identify the conceptual requirements for preserving authentic electronic records, and sought to identify the elements of electronic records that are necessary to maintain their authenticity over time. The original research questions to be addressed by the Authenticity Task Force, chaired by Heather MacNeil, were:

- What are the elements that all electronic records share?
- What are the elements that allow us to differentiate between different types of electronic records?
- Which of those elements will permit us to verify their authenticity over time?
- Are these elements for verifying authenticity over time the same as those that permit us to verify their authenticity in time (i.e. at the point at which they are originally used)?
- Can those elements be removed from where they are currently found to a place where they can more easily be preserved and still maintain the same validity?

Domain 2 was concerned with the appraisal of authentic electronic records. The Appraisal Task Force, chaired by Terry Eastwood, sought to determine whether or not the evaluation of
electronic records for permanent preservation should be based on theoretical criteria different from those applied to traditional records. Specific research questions originally included:

- What is the influence of digital technology on appraisal criteria?
- In what ways does appraisal differ depending on the type of systems prevalent in each phase of computing?
- How do the media and physical form of the records influence appraisal?
- How do retrievability, intelligibility, functionality, and research needs influence appraisal?
- Does the life cycle of electronic records differ from that for traditional records?
- When in the course of their existence should electronic records be appraised?
- Should electronic records be appraised more than once in the course of their existence and, if so, when?
- How are electronic records scheduled?
- Who should be responsible for appraising electronic records?
- What are the appraisal criteria and methods for authentic electronic records?

Domain 3 focused on records preservation. The Preservation Task Force, chaired by Ken Thibodeau, sought to identify the procedures and resources necessary for the long-term preservation of authentic electronic records, all the while considering the conceptual requirements for authenticity articulated by the Authenticity Task Force as well as the conclusions of the Appraisal Task Force. The research questions were:

- What methods, procedures, and rules of long-term preservation are in use or being developed?
  - Which of these meet the conceptual requirements for authenticity identified in Domain 1?
  - Which methods of long-term preservation need to be developed?
  - Which of these methods are required or subject to standards, regulations, and guidelines in specific industry or institutional settings?
- What are the procedural methods of authentication for preserved electronic records?
  - In what way can archival description be a method of authentication for electronic records?
  - In what way can appraisal and acquisition/accession reports be constructed to allow for the authentication of electronic records?
  - What are the procedures for certifying electronic records when they cross technical boundaries (e.g., refreshing, copying, migrating) to preserve their authenticity?
- What are the technical methods of authentication for preserved electronic records?
- What are the principles and criteria for media and storage management that are required for the preservation of authentic electronic records?
- What are the responsibilities for the long-term preservation of authentic electronic records?

Domain 4 was devoted to the articulation of a framework for developing policies, strategies, and standards for the long-term preservation of authentic electronic records. The Strategy Task Force, chaired by myself as project director, sought to develop such a framework, recognizing that effective organizational policies, strategies, and standards derive from the recognition that each cultural, juridical, and organizational environment has its own needs. An important aspect of the research within Domain 4 was the contextualization of the findings for each of the participating jurisdictions. The research questions for Domain 4 were:

- What principles should guide the formulation of policies, strategies, and standards related to the long-term preservation of authentic electronic records?
- What should be the criteria for developing national policies, strategies, and standards?
- What should be the criteria for developing organizational policies, strategies, and standards?
In addition to these four task forces, a Glossary Committee, chaired by Ken Hannigan, was struck to control the use of specialized vocabulary within the project.

InterPARES research was a truly collaborative effort, with the participation of academics from a range of disciplines, working archivists from a number of national archival institutions, and representatives from the global business community. The project was directed by Dr. Luciana Duranti (University of British Columbia), and included the participation of the following individual researchers: Jason Baron (U.S. National Archives and Records Administration), Elisabetta Bidischiini (Union for the Chambers of Commerce, Italy), Richard Blake (Public Records Office of the United Kingdom), Sergio Cardarelli (Bank of Italy), Barbara Cartocci (Chamber of Deputies, Italy), Paola Carucci (Central State Archives of Italy), Su-Shing Chen (University of Missouri-Columbia), Chen Wei (Beijing Municipal Archives), Michele Cloonan (University of California at Los Angeles), Barbara Craig (University of Toronto), Fabrizio De Martinis (Sogei), Du Mei (State Archives Administration of China), Terry Eastwood (University of British Columbia), Fynnette Eaton (Smithsonian Institution Archives), Philip Eppard (University at Albany, SUNY), Sharon Farb (University of California at Los Angeles), Vincenzo Festinese (Bank of Italy), Gigliola Fioravanti (Central Office for Archival Properties of Italy), Normand Fortier (National Archives of Canada), Lucilla Garofalo (Central State Archives of Italy), Anne Gilliland-Sweetland (University of California at Los Angeles), Linda Giuva (Central State Archives of Italy), Maria Guercio (University of Urbino), Yvette Hackett (National Archives of Canada), Babak Hamidzadeh (University of British Columbia), Ken Hannigan (National Archives of Ireland), P.C. Harirahan (The Johns Hopkins University), Ross Harvey (Curtin University of Technology), Hans Hofman (National Archives of the Netherlands), Torbjorn Hornfeldt (National Archives of Sweden), Peter Horsman (Netherlands Institute for Archival Education and Research), Livia Iacovino (Monash University), Caterina Isabella (Municipal Environment Agency of Rome), Agnes Jonkers (Netherlands Institute for Archival Education and Research), Alexandra Kolega (Central State Archives of Italy), Brent Lee (University of British Columbia), Rich Lysakowski (Collaborative Electronic Notebook Systems Association), John McDonald (National Archives of Canada), Ian Macfarlane (Public Records Office of the United Kingdom), Sue McKemmish (Monash University), Heather MacNeil (University of British Columbia), Roger Maxwell (Public Records Office of the United Kingdom), Mirella Mombelli (Special School of Archivists and Librarians, Italy), Antonella Mule (Central State Archives of Italy), Gianni Paolini (Special School of Archivists and Librarians, Italy), Christine Petillat (National Archives of France), Tom Quinlan (National Archives of Ireland), Marina Raffaelli (Special School of Archivists and Librarians, Italy), Enrico Rendina (Rome Research Consortium), Bill Rhind (Collaborative Electronic Notebook Systems Association), John Roeder (University of British Columbia), Seamus Ross (University of Glasgow), Claudia Salmini (State Archives of Venice), Maurizio Savoja (State Archives of Milan), Leon Stout (Penn State University), Mario Terranova (Authority for Public Administration Information Technology), Ken Thibodeau (U.S. National Archives and Records Administration), Bill Underwood (Georgia Tech Research Institute), Wai-kwok Wan (Hong Kong Public Records Office), Bruce Walton (National Archives of Canada), and Zhao Zhon Xiu (State Archives of China).

The InterPARES Project was funded as a Major Collaborative Research Initiative by the Social Sciences and Humanities Research Council of Canada, with matching funds from the National Archives of Canada; the University of British Columbia’s Hampton Fund, Vice President Research Fund, and Dean of Arts Fund; and the University of Toronto’s Faculty of Information Studies. The School of Library, Archival and Information Studies has provided the space for the headquarters of the InterPARES Project. Non-Canadian national and multinational teams have provided financial support to their own researchers, either by obtaining research grants or by securing for their individual members sufficient resources in money and in kind from their own institutions (i.e., universities and archives). In addition, they have funded some research activities of the whole project, specifically: for the American team, the National Archives and Records Administration of the United States, and the National Historical Publications and Records Commission funded four research workshops and a symposium; for the Italian team, the Italian Ministry for Cultural Properties and Activities, Central Office for Archival Properties, funded three research workshops and published the proceedings of the first Vancouver symposium on InterPARES and the final
findings, and the Italian National Archival Association funded one research workshop and an international conference on InterPARES in Cagliari; for the European team, the Open Society Archives of the Soros Foundation in Budapest offered a research workshop and an international seminar; and for the Asian team, the State Archives Administration of China funded a two-day symposium on InterPARES findings in Beijing. External supporters of InterPARES have been the Italian Institute of Culture of Vancouver, that has co-sponsored two symposia on InterPARES in Vancouver, once with the UBC European Institute, and once with the Italian Chamber of Commerce of British Columbia; and the East section of the International Council on Archives (EASTICA), that has sponsored a conference on InterPARES in Hong Kong.

The organization of the InterPARES findings reflects the organization of the project. Part One is devoted to the report of the Authenticity Task Force. Part Two comprises the Appraisal Task Force Report. Part Three is the report of the Preservation Task Force. Part Four includes the report of the Strategy Task Force as well as the reports of the various national and multinational teams. Part Five is the report of the Glossary Committee.

The appendices following the reports include several documents that are in themselves major products of the InterPARES Project, such as the Template for Analysis, the Requirements for Assessing and Maintaining the Authenticity of Electronic Records, the models of the “Selection” and “Preservation” functions, and some self-contained studies that may be of interest in their own right. Not all the key documents produced by InterPARES researchers in the course of their investigations are included in the InterPARES findings, primarily to contain their volume. However, those that are excluded are electronically published and accessible on the project’s Web site, <http://www.interpares.org>.

Luciana Duranti
InterPARES Project Director
Authenticity Task Force Report

Task Force Members
Heather MacNeil, University of British Columbia (Chair)
Chen Wei, Beijing Municipal Archives
Luciana Duranti, University of British Columbia
Anne Gilliland-Swatland, University of California, Los Angeles
Maria Guercio, University of Urbino
Yvette Hackett, National Archives of Canada
Babak Hamidzadeh, University of British Columbia
Livia Iacovino, Monash University
Brent Lee, University of British Columbia
Sue McKemmish, Monash University
John Roeder, University of British Columbia
Seamus Ross, University of Glasgow
Wai-kwok Wan, Hong Kong Public Record Office
Zhao Zhon Xiu, State Archives of China
Introduction

The goal of the Authenticity Task Force was to identify conceptual requirements for assessing and maintaining the authenticity of electronic records. Under the original InterPARES research plan, five questions were to be addressed within the domain of investigation assigned to this task force (Domain 1):

- What are the elements that all electronic records share?
- What are the elements that allow us to differentiate between different types of electronic records?
- Of those elements, which will permit us to verify their authenticity over time?
- Are the elements for verifying authenticity over time the same as those that permit us to verify their authenticity in time, that is, at the point at which they are originally created and transmitted?
- Is it possible to move the elements from their current position to a place where they can more easily be preserved, without affecting validity?

During the course of the research, these initial questions were considerably revised and refined, and new, unanticipated questions emerged.

Recognizing the need to delineate the full complexity of the issues associated with the authenticity of electronic records, the task force adopted two distinct yet complementary research approaches. The first was theoretical and deductive, based on contemporary archival diplomatics. It involved identifying and defining the elements of an ideal electronic record in general, and those that are relevant to a consideration of its authenticity in particular, using concepts and methods derived from diplomatics and archival science that in turn are based upon what is known about traditional records, juridical systems, and record-keeping practices. The second approach was inductive and empirical, and made use of selected case studies of extant electronic systems. Although these systems were in many cases far removed from the ideal electronic record as established through the first approach, they were able to elucidate the shifting boundaries of electronic records, emergent record-keeping processes, and new manifestations of traditional record elements. Both approaches were aimed at theory building, and the conceptual requirements for assessing the authenticity of electronic records emerged out of their triangulation.

The primary outcome of the work of the Authenticity Task Force has been the development of two sets of requirements: one includes requirements that support the presumption of the authenticity of electronic records before they are transferred to the preserver’s custody; the other includes requirements that support the production of authentic copies of electronic records after they have been transferred to the preserver’s custody. The research also resulted in several additional datasets and products, which are discussed in this report.

Basic Assumptions of the Research

Definition of record

A record is defined as any document made or received and set aside in the course of a practical activity. The interpretation of this definition in the context of electronic systems is discussed under the “Research Design and Methodology” and “Research Findings” sections of this report.
Definitions of authenticity, authentic, and authentic record
In common usage, the concept of authenticity is defined as “the quality of being authentic, or entitled to acceptance.”\(^1\) The term authentic means “worthy of acceptance or belief as conforming to or based on fact” and is synonymous with the terms genuine and bona fide. Genuine “implies actual character not counterfeited, imitated, or adulterated [and] connotes definite origin from a source.” Bona fide “implies good faith and sincerity of intention.”\(^2\) From these definitions it follows that an authentic record is a record that is what it purports to be and is free from tampering or corruption.

Rationale for establishing conceptual requirements for assessing the authenticity of electronic records
In both archival theory and jurisprudence, records upon which the creator relies in the usual and ordinary course of business are presumed authentic. However, records created and maintained in electronic form are continually at risk of inadvertent or intentional alteration, and such alteration may not be readily perceptible. The authenticity of electronic records is threatened whenever the records are transmitted across space (i.e., when sent between persons, systems or applications) or time (i.e., either when they are stored offline, or when the hardware or software used to process, communicate, or maintain them is upgraded or replaced). Requirements for assessing and maintaining the authenticity of electronic records that are preserved over the long term are necessary, therefore, to support the presumption that an electronic record is, in fact, and continues to be, what it purports to be and has not been modified or corrupted in essential respects. The interpretation of what constitutes “in essential respects” is explained under “Research Findings.”

Differentiating between authenticity and authentication
Because of the ongoing developments in the area of authentication technologies, it is necessary to clarify the distinction between authenticity—which is the focus of InterPARES—and authentication. In common usage, authentication is understood as a declaration of a record’s authenticity at a specific point in time by a juridical person entrusted with the authority to make such a declaration. It takes the form of an authoritative statement (which may be in the form of words or symbols) that is added to or inserted in the record attesting that the record is authentic.

Digital signatures are an example of an authentication technology that has been developed to address the need for secure electronic communication across open networks such as the Internet. Digital signatures, which identify the sender of a data object and verify that it has not been altered in transmission, can support the authentication of electronic records, but they are not sufficient to establish the identity and demonstrate the integrity of an electronic record over the long term. Further research is needed to determine the specific impact of digital signatures on the long-term preservation of authentic electronic records.\(^3\)

Research Design and Methodology

The theoretical-deductive approach
In the first stage of its research, the task force established the theoretical framework for the analysis of various types of electronic records and the identification of those elements that need to be preserved to ensure the records’ authenticity over time. The Template for Analysis

---

\(^1\) Oxford English Dictionary, 2nd ed., s.v. “authenticity.”
\(^2\) Merriam-Webster Online Collegiate Dictionary, s.v. “authentic.”
\(^3\) See "Technological development".
embodies this framework. The Template is a decomposition of an electronic record into its constituent elements. The decomposition defines each element, explains its purpose, and indicates whether, and to what extent, that element is instrumental in assessing the record's authenticity.

The theoretical perspective that shaped the development of the Template was contemporary archival diplomatics. Diplomatics emerged in the seventeenth century as an analytical technique for determining the authenticity of records issued by sovereign authorities in previous centuries. Its primary purpose was to ascertain "the reality of the rights or truthfulness of the facts" contained in such documents. The tenets and methods of diplomatics were laid out in 1681 in a treatise written by Jean Mabillon, a Benedictine monk. Mabillon examined, among other things, the language of the documents, their characteristic parts, their seals, and the systems of chronology used in dating them. On the basis of this examination, "Mabillon stated what, for a particular time and place, was the correct form for a genuine document, and presented the general principles of diplomatics." The original use of diplomatics was to determine a record's authenticity for legal purposes; that use continued into the eighteenth century, when many European faculties of law incorporated its concepts and principles into their curricula. By the end of the nineteenth century, however, under the influence of classical philology and the scientific school of historiography, diplomatics emerged as a tool for assessing the authority of medieval records as historical sources.

Over the past twenty years there have been numerous calls from within the archival community to revive and adapt diplomatics as an aid to understanding the record-keeping processes of contemporary bureaucracies. Delegates to the 1989 International Council on Archives' Second European Conference on Archives, for example, recommended "that the development of the discipline of modern diplomatics be promoted through research in the typology of contemporary records and in the records-creating procedures of contemporary institutions." In Europe, notable archival efforts to construct a modern diplomatics include the work undertaken by Dutch archivists to develop a typology of records created by organizations since the nineteenth century in the

---

4 See Appendix 1. The term elements is used differently in diplomatics to the way in which it is used in information systems design. In developing the initial research questions and the Template for Analysis, the task force used the diplomatic term elements to refer to both general and specific characteristics of a record that may be found in its documentary form, in annotations, or in one or more of its various contexts. As the research progressed, however, the task force found it necessary to narrow the scope of the concept. In the Requirements for Assessing and Maintaining the Authenticity of Electronic Records, therefore, the term record elements refers specifically to the intrinsic and extrinsic elements of a record's documentary form as these are identified in Template for Analysis. Such redefinition is illustrative of how diplomatics continues to evolve in response to the changing nature of the record.


Netherlands, and the adaptation of traditional diplomatic concepts and methods to the record-keeping environment of contemporary Italian administration undertaken by Paola Carucci.

In North America, the most comprehensive effort to adapt traditional diplomatics to contemporary record-keeping practices is embodied in the work of Luciana Duranti of the University of British Columbia. In a series of articles written between 1989 and 1992, Duranti examined the principles and concepts developed by diplomatic theorists for evaluating the authenticity of medieval documents to determine whether they could be adapted for application to the records generated by modern bureaucracies. Over the course of the six articles, she refined and interpreted the classical concepts, and introduced new ones to take into account the variety and complexity of bureaucratic record-keeping environments.

Duranti's series of articles resulted in a preliminary elaboration of contemporary archival diplomatics, an adaptation of traditional diplomatic concepts and methods to contemporary record-keeping environments, and an integration of these concepts and methods with those of archival science. It also laid the groundwork for a research project carried out between 1994 and 1997 at the University of British Columbia entitled The Preservation of the Integrity of Electronic Records (known as the UBC Project). The goal of that project was to identify and define conceptually the nature of an electronic record and the conditions necessary to ensure its integrity (i.e., its reliability and authenticity) during its active and semi-active life. The research resulted in a set of standards and rules for developing and implementing a trustworthy electronic record-keeping system.

The elements of an electronic record identified in the UBC Project provided the starting point for the identification of the InterPARES Template elements. Based on researcher input from a range of disciplinary perspectives, as well as data collected during the case studies, these original elements were revised and extended, and new elements were added as the research progressed. For example, the broader administrative and documentary contexts in which a record is created, handled, and maintained were more precisely articulated in the Template than they had been in the UBC Project, and a new category of context—that is, technological context—was identified and elaborated.

To help the researchers understand traditional diplomatic elements and their contemporary interpretation, student research assistants traced the lineage of the elements included in the Template back to their original elaboration in the work of traditional French, German, and Italian diplomats. The researchers reasoned that a sound understanding of the historical meaning of the elements would better equip them to assess the contemporary relevance of the elements.

---


12 The outcomes of the UBC Project were subsequently substantially incorporated into the Design Criteria Standard for Electronic Records Management Software Applications (DoD 5015.2-STD) promulgated by the U.S. Department of Defense.

13 The document showing the lineage of elements included in the Template for Analysis is available on the InterPARES Web site at <http://www.interpares.org/reports.htm>.
Research assistants also prepared a sample typology of papal chancery documents to facilitate the researchers’ understanding of how traditional diplomatists viewed the relationship between authenticity and documentary form, and, more specifically, how individual elements of documentary form supported the attestation of a record’s authenticity.14

Viewed from the perspective of contemporary archival diplomatics, an electronic record, like its traditional counterpart, is a complex of elements and their relationships. It possesses a number of identifiable characteristics, including a fixed documentary form,15 a stable content, an archival bond with other records either inside or outside the system, and an identifiable context. It participates in or supports an action, either procedurally or as part of the decision-making process (meaning its creation may be mandatory or discretionary), and at least three persons (author, writer, and addressee) are involved in its creation (although these three conceptual persons may in fact be only one physical or juridical person).

In a traditional record-keeping environment, these characteristics manifest themselves in explicit and implicit ways. For example, the name of the author may appear on the letterhead; and the archival bond may be expressed in a classification code or some other unique identifier that appears on the face of a record. The purpose served by these individual elements also depends on their specific form of expression. For example, the identification of the name of the author that appears in the letterhead serves the purpose of identifying aspects of the record’s provenancial context. When that same name appears as a signature at the bottom of the record, it serves the purpose of attesting to the validity of the record or its content. The working hypothesis of the Authenticity Task Force was that, even though they may manifest themselves in different ways, these same or similar elements are present explicitly or implicitly in electronic records. The Template for Analysis was created to test that hypothesis. The elements of an electronic record included in the Template fall into four main categories: documentary form, annotations, context, and medium:16

Documentary form is defined as the rules of representation according to which the content of a record, its immediate administrative and documentary context, and its authority are communicated. It possesses both intrinsic and extrinsic elements.

Intrinsic elements are the discursive parts of the record that communicate the action in which the record participates and the immediate context. They fall into three groups:

1) elements that convey aspects of the record’s juridical and administrative context (e.g., the name of the author, addressee, the date);
2) elements that communicate the action itself (e.g., the indication and description of the action or matter);
3) elements that convey aspects of the record’s documentary context and its means of validation (e.g., the name of the writer, the attestation, the corroboration).

Extrinsic elements refer to specific, perceivable features of the record that are instrumental in communicating and achieving the purpose for which it was created. For electronic records these include:

14 The sample typology is available on the InterPARES Web site at <http://www.interpares.org/reports.htm>.
15 According to the Authenticity Task Force’s Research Methodology Statement, a fixed form “means that (1) the binary content of the record, including indicators of its documentary form, are stored in a manner that ensures it remains complete and unaltered; and (2) technology has been maintained and procedures defined and enforced to ensure that the content is presented or rendered with the same documentary form it had when it was set aside.” The statement is available on the project Web site at <http://www.interpares.org/reports.htm>.
• overall presentation features (e.g., textual, graphic, image, sound, or some combination of these);
• specific presentation features (e.g., special layouts, hyperlinks, colours, sample rate of sound files);
• electronic signatures and electronic seals (e.g., digital signatures);
• digital time stamps;
• other special signs (e.g., digital watermarks, an organization’s crest or personal logo).

Annotations (additions made to a record after it has been created) constitute the next category of elements included in the Template for Analysis. They fall into three basic groups:

1. additions made to the record after its creation as part of its execution (e.g., the date and time of transmission added to an e-mail record at the moment it is sent, or the indication of attachments added before it is transmitted);
2. additions made to the record in the course of handling the business matter in which the record participates (e.g., comments noted on the face of the record, or embedded in it, and dates of transmission to other offices);
3. additions made to the record in the course of handling it for records management purposes (e.g., the classification code or file number assigned to the record, its draft and/or version number, cross references to other records, and an indication of scheduling actions).

Context shifts the analysis away from the record itself to the broader structural, procedural, and documentary framework in which the record is created and managed. The identified elements of context correspond to a hierarchy of frameworks ranging from the general to the specific. They include the record’s juridical-administrative context, its provenancial context, its procedural context, its documentary context, and its technological context. Knowledge of these elements is critical to an understanding of the business processes in the course of which electronic records are created, maintained, and used; the types of records generated from these processes; and the connection between those processes and the creator’s broader functions and mandate.17

Medium proved to be a problematic construct from the perspective of diplomatic analysis. In identifying and positioning the elements of the Template for Analysis, the Authenticity Task Force struggled with the question of whether to treat the medium—that is, the physical carrier on which a record is stored—as a part of the record itself or as part of its technological context. For diplomats examining medieval documents, the medium is an essential component of a record because the examination of the physical carrier on which the document is inscribed is one of the most obvious proofs of its authenticity.18 In the translation of traditional diplomatic concepts into modern, paper-based, record-keeping environments, the medium has continued to be treated as a part of the record itself, mainly because the medium and the message are inextricably linked.

The question was whether, in an electronic record-keeping environment, the medium should continue to be treated as an essential part of the record itself given that: (1) the medium and the message are no longer inextricably linked; (2) what is inscribed on or affixed to the medium is not a record as such (or words, or pictures), but a bitstream; and (3) the choice of a medium by those creating or maintaining the record is often arbitrary and carries no particular significance.

17 For a discussion of the embeddedness of electronic records within these contexts, see Anne J. Gilliland-Swatland and Philip Eppard, “Preserving the Authenticity of Contingent Digital Objects: The InterPARES Project,” Dlib Magazine 6 (July/August 2000), available at: <http://www.dlib.org/dlib/july00/eppard/07eppard.html>.
18 For example, a royal diploma of Childerbert I (King of Francs, sixth century) that is written on parchment instead of papyrus is considered false. The medium also provides evidence of the manner in which medieval documents were prepared. The documents from the German chancery have many erasures and corrections in comparison to the documents of the papal chancery, indicating a lesser degree of care and accuracy in the preparation of the final documents.
A record is assumed to be a representation of a fact or act that is memorialized on a physical carrier—that is, a medium—and preserved by a physical or juridical person in the course of carrying out its activities.\textsuperscript{19} It follows that a record cannot exist before its elements have been inscribed on or affixed to a medium. Similarly, in an electronic environment, the bitstream—that is, the source of the record—cannot endure for any length of time unless it is affixed to a medium. Storage of a bitstream on a disk or tape, however, while necessary for the bitstream to endure, is not sufficient to preserve a record as a record. As the Preservation Task Force observed early on in its deliberations, “strictly speaking, it is not possible to preserve an electronic record. It is only possible to preserve the ability to reproduce an electronic record. It is always necessary to retrieve from storage the binary digits that make up the record and process them through some software for delivery or presentation.”\textsuperscript{20} Moreover, although affixing a bitstream to a medium is a precondition to the existence of an electronic record, it is not necessarily a relevant factor in assessing that record’s authenticity. It is assumed to be neutral with respect to the record’s authenticity, at least from the perspective of the records creator and the records preserver. The Authenticity Task Force therefore concluded at the end of its research that the medium should be considered part of the record’s technological context, rather than an essential part of the record itself.

Initial development of the Template took place over a nine-month period from January to September 1999. During that time, the Template was revised numerous times by both the Authenticity Task Force and the InterPARES International Team. By June 1999, the Template was considered sufficiently developed for the start of the process of testing and refining it through case studies of real-life electronic systems.

**Empirical-Inductive Approach**

As discussed above, the Template for Analysis began as a model of an ideal record that, based upon prior archival knowledge of record types, delineated the possible known elements that a record may contain. However, where diplomatic typologies and analysis have in the past been developed retrospectively based upon what was known about existing records, one goal of InterPARES was to develop a predictive model that would assist archivists in identifying future record types and the necessary requirements for maintaining their authenticity over time.

In the first year of the project, InterPARES researchers determined that they could develop a richer picture of the complex nature of electronic records if they triangulated the theoretical, deductive diplomatics-based approach with an inductive, empirical approach that was based on an examination of actual electronic records and electronic record-keeping systems. This examination was conducted by means of purposively selected, interpretive case studies of electronic systems that contained, or were deemed likely to contain, electronic records. These case studies were directed towards understanding electronic records within their various contexts as well as the relationships of those contexts to each other.

The addition of this “bottom-up” approach extended InterPARES research activities considerably beyond those originally envisaged. It provided a rich dataset that informed the theoretical development by indicating the increasing role of procedural and technological context in ensuring and maintaining the authenticity of records. At the same time, the application of the Template of Analysis to existing records and record-keeping systems gave an indication of which necessary extrinsic and intrinsic elements of form were not present in systems as they were currently designed and operating, thus demonstrating potential weaknesses or deficiencies in the records or record-keeping systems examined.

\textsuperscript{19} Maria Guercio, “Principi, metodi e strumenti per la formazione, conservazione e utilizzo dei documenti archivistici in ambiente digitale,” *Archivi per la storia* XII, 1–2 (1999): 26.

Use and selection of case studies

The task force researchers adopted a grounded theory approach; four successive rounds of case studies of electronic systems that contained or potentially contained records were examined in order to identify and describe phenomena associated with the records and their contexts. Grounded theory is a method for discovering concepts and hypotheses and developing theory directly from data under observation.\textsuperscript{21} Cases are selected for study “according to their potential for helping to expand on or refine the concepts or theory that have already been developed. Data collection and analysis proceed together.”\textsuperscript{22}

Because of the grounded theory approach, researchers employed theoretical rather than statistical sampling in the selection of case studies. Glaser and Strauss describe the process of theoretical sampling as “a process of data collection for generating theory whereby the analyst jointly collects, codes, and analyzes his data and decides what data to collect next and where to find them, in order to develop his theory as it emerges.”\textsuperscript{23} In other words, task force researchers purposively identified the cases that seemed most likely to explain phenomena that the research was seeking to understand (e.g., what happens to active or inactive electronic records when they are subject to migration?). No attempt was made to draw a representative or statistically significant sample.

In the first two rounds of case studies, the focus was on electronic systems, although a considerable amount of contextual data was collected to elucidate the broader record-keeping environment. Following the International Team and Authenticity Task Force’s evaluation of these case studies and analysis of case study data, the criteria for selection were adjusted to support continued theory building.

The data gathered through these case studies were then used to test and extend the Template for Analysis. The translation of the case study data into a form that could be analyzed diplomatically by the Template was achieved by coding the data for interrelated themes and concepts using a Template Element Data Gathering Instrument (TEDGI).\textsuperscript{24} The data collected through the case studies were also made available to the Appraisal and Preservation Task Forces to assist them with modelling preservation processes and then with walking through their models.

First- and second-round case studies had to meet at least three of the following criteria:

1. Systems that contain, generate, or have the potential or possibility of generating records.\textsuperscript{25}
2. Systems that have gone through one or more migrations.
3. Systems where migration(s) was (were) from one electronic system to another.
4. Systems for which several aspects of technological context (storage media, system software, application software, data format, schema) were changed in the course of each migration.


\textsuperscript{23} Glaser and Strauss, \textit{Grounded Theory}, 45.

\textsuperscript{24} The TEDGI is available on the project Web site, at <http://www.interpares.org/reports.htm>.

\textsuperscript{25} As explained earlier, a record possesses a number of identifiable characteristics. These include a fixed documentary form, a stable content, an archival bond with other records either inside or outside the system, and an identifiable context. It participates in or supports an action, either procedurally or as part of the decision-making process, and at least three persons (author, writer, and addressee) are involved in its creation.
5. Systems for which the pre-migration and the post-migration versions were available and functional.

6. Systems for which detailed documentation (design, implementation, migration, metadata) exists.

7. Systems with a diversity of information configurations (e.g., contain both text and images).

An effort was also made to ensure diversity in content and type of records (i.e., case studies representing a variety of systems) among the candidate systems proposed by the same archival institution. Between institutions, an effort was made to identify and conduct case studies on record-keeping systems performing similar functions (e.g., student registration systems in different universities). The researchers believed that these two factors might enable them to see emergent patterns relating to the nature of organizational record-keeping and specific record-keeping functions.

A key issue encountered by the researchers, and indeed by any archivist or records managers working with electronic records, is the difficulty in identifying actual electronic records and their parameters. This issue stems from the nature of digital information systems, which are frequently multipurpose, highly networked database systems capable of containing a diversity of information elements that can be compiled and presented in a variety of ways (e.g., through hard-coded report formats, style sheets, and virtual “on-the-fly” views) and that can invoke a range of functionalities, according to the needs of different users. A single system may contain only raw data or information, one or more than one type of record, or a combination of record types and data or information. The diplomatic analysis of first- and second-round case studies indicated that few of the systems contained records within close range of the ideal promulgated in the Template (some systems proved to be information systems not containing records at all; some contained records that were able to achieve their purpose but were not intrinsically very good records). In line with the grounded theory approach, based upon what they had learned from the first two rounds of case studies, the researchers modified the case study selection criteria for the third and fourth round of case studies to define more precisely the types of cases in which they were now interested. Through this redefinition, researchers incorporated the following indicators of systems that are known to create records or have the potential to create records:

- if the action in which the system participates is juridically required;
- if there is a business procedure in place to carry out that action;
- if the system operates within the management or strategic decision-making levels of the organization.

For case study rounds two to four, the researchers decided to examine only live systems (i.e., systems still being actively used by the creator to carry out business activities), since the case studies of inactive electronic records indicated that too much contextual information had already disappeared for the task force to be able to analyze the records and record-keeping systems successfully. The researchers also eliminated criteria that related to systems and records that had undergone migration, since these had not proven to yield significant additional insights for either the Authenticity Task Force or the Preservation and Appraisal Task Forces. Additional desirable criteria identified for rounds three and four case studies were:

1. Systems that come from different hierarchical levels within an organization; and optimally, systems supporting management and strategic level activities.
2. Systems that contain supporting and narrative records.
3. Systems from the private sector.
4. Financial management systems.
5. Multimedia systems.
6. Computer-aided design (CAD) systems.

Case study data
Between spring 1999 and spring 2001, four rounds of case studies were conducted by institutional and student researchers in government, university, and corporate agencies in
Canada, the United States, Italy, the United Kingdom, the Netherlands, and China. The case studies included large-scale databases (such as patent and student registration systems), geographic information systems, and interactive Web-based applications as they existed at the time the case studies were conducted.

A drawback of any research that employs multiple selective case studies is the limited degree to which it is possible to compare across or generalize from individual case studies. Each case is highly sensitive to its own national, juridical, institutional, and technological contexts. Moreover, InterPARES case studies were conducted under a range of different conditions by different investigators. As a result, each case study had to be selected and analyzed on its own merits for how it might inform theory development by the researchers, and there was a need for caution about the extent to which one could look for patterns emerging across case studies located in similar institutional settings or performing similar functions in different settings. In an effort to control as much as possible the individual differences between case studies and within case study rounds, a Case Study Interview Protocol (CSIP) was developed by the Authenticity Task Force to standardize the interview process for the case studies as well as to provide data for populating the TEDGI. Several project investigators who would be conducting the case study interviews also participated in training sessions at UBC or the University of California, Los Angeles (UCLA) on how to conduct the case studies as well as how to complete the TEDGI and CSIP.26

The CSIP (essentially the interview script) was divided into five sections: context (juridical-administrative, provenancial, procedural, and documentary); intrinsic elements of form; extrinsic elements of form; annotations; and medium and technological context. A range of standardized questions was asked to elucidate each aspect. The same question was sometimes asked in different ways within the same section to check for consistency in responses. The same question was also sometimes asked in a different way in more than one section to identify any alternative perspectives of respondents with different backgrounds (for example, records managers and systems personnel). Interviewers, predominantly institutional archivists or archival science students participating in InterPARES, sought out respondents who were the records creators, records managers, and systems personnel primarily responsible for working with the electronic systems under study. Due to local requirements and practicalities, some interviews were with individuals, and some with groups of individuals. In some case studies, multiple interviews with different individuals were held. Interviewers also collected supporting documentation such as technical information, organization charts, and work-flow rules; and sometimes followed up with interviewees when further information was required. The interviewers were then responsible for translating the data they had gathered through the CSIP and supporting documentation into the TEDGI, and for transmitting copies of all the case study data to both UBC and UCLA for analysis. In the third round of case studies, researchers at UBC were responsible for compiling the TEDGI.

Version 2.1 of the CSIP and 1.0 of the TEDGI were used for the first round of case studies. Then the CSIP, TEDGI and the Template for Analysis were revised to eliminate, clarify, or expand aspects identified as problematic in the first round of case studies. In the second round of case studies, however, most interviewers still used version 1.0 (rather than 1.1) of the TEDGI, but version 3.0 of the CSIP. In the third round of case studies, researchers used version 1.1 of the TEDGI, and version 3.1 of CSIP. In the fourth round, researchers used version 1.1 of the TEDGI and version 3.2 of the CSIP. In total, data were analyzed for twenty-six completed case studies from the four rounds of case studies using two different versions of the TEDGI and four versions of the CSIP.

Multiple types of data were sought or created by the task force researchers in the course of conducting and analyzing each case study. These types included the CSIP and TEDGI, audio

26 As required by the different researchers’ individual institutions, the entire protocol for the case studies and all subsequent revisions to the protocol were submitted for review and approved by the institutional review boards/offices for the protection of human subjects. The CSIP and TEDGI are available on the InterPARES Web site at <http://www.interpares.org/reports.htm>. 
and videotapes of interviews, supporting procedural and technological documentation, and case study overviews. Not all data types exist for each case however, due to variations in how data were collected (e.g., interviewees could decline to be audio taped), or to lack of availability of specific supporting technological or procedural documentation or translations of that documentation into English. It is also important to note that in the majority of cases, although the case study focused on the electronic system, the actual record-keeping system comprised both paper and electronic components.

Case study data analysis

Each round of case studies was described and analyzed from the perspective of contemporary archival diplomacy—the primary emphasis of the work of the Authenticity Task Force—as well as through the application of analytical methods drawn from the social sciences. The rationale behind subjecting case study data to such a barrage of analyses was to render the most complete picture possible of the complexities of the modern electronic record, and to feed this emerging knowledge into the development of records theory, and archival diplomacy in particular.

Diplomatic analysis of case studies

The primary purpose of analyzing the case studies from the perspective of contemporary archival diplomacy was to consolidate information needed for (1) the drafting of the conceptual requirements for assessing the authenticity of electronic records, and (2) the development of a typology of electronic records based on those requirements. The analyses were undertaken by student researchers in their second year of the Master of Archival Studies Program at UBC. All the students were familiar with diplomatic analysis, having completed a course in diplomatics during their first year of the program.

The process of analysis took place in three phases: case studies from rounds one and two were analyzed in the first phase (fall 2000); those from round three, in the second phase (January 2001); and those from round four, in the third phase (May 2001). The process consisted of periodic team meetings of the research assistants with the Authenticity Task Force representative (Luciana Duranti) and the project coordinator (Tahra Fung) to discuss findings and brainstorm; and independent work by pairs of research assistants in the periods between team meetings. The work process differed from phase to phase as research assistants began to work more independently and as the responsibilities assigned to them grew. For example, responsibility for populating TEDGs, which was assigned to case study researchers in rounds one and two, was assigned to the research assistants for rounds three and four case studies. The process produced a considerable amount of documentation, including synopses of the systems, answers to assigned questions, inquiries directed to case study researchers, and speculative scenarios. The final product of the diplomatic analysis was constituted by a “final report” for each system analyzed. Final reports were written for twenty-two of the twenty-six completed case studies. Four case studies were excluded from the diplomatic analysis due to insufficient information or language difficulties with technical documentation.

The diplomatic analysis of the first two rounds of case studies commenced in September 2000. Each pair of research assistants was assigned a case study and asked to complete a diplomatic analysis of the electronic system on the basis of the Template for Analysis; the CSIP; case study supporting documentation (including organization charts, lists of employee responsibilities, printouts of screen views, interview tapes, legislation, diagrams of business procedures, glossaries used by the organization, etc.); and Duranti’s six-part exploration of diplomacy. After a preliminary examination revealed significant inconsistencies and differing interpretations on the

27 The account of the process of diplomatic analysis is based on a summary prepared by Ian McAndrew, with contributions from April Miller and Anna Gibson.

28 See fn. 10.
part of interviewers in the translation of data from the CSIP into the TEDGI, it was decided that only the CSIPs and not the TEDGIs completed by the case study researchers would be used in the diplomatic analysis in the first phase.

To complete the diplomatic analysis, the research assistants were assigned the following questions:

- How many records are in the system?
- What is the function of these records (i.e., dispositive, probative, supporting, narrative)?\(^{29}\)
- What is/are the action(s) associated with the system?
- What types of documentary forms are included?
- What is the status of transmission of each documentary form (i.e., original, draft, copy)?\(^{30}\)

Considering these questions was essentially a process of examination that led to four more questions:

- Does the system contain records?
- Should the system contain records?
- Is the system itself a record?
- With the nature and function of the system in mind, is there a presumption of authenticity? If yes, what is the basis for this presumption?

Answering both sets of questions proved considerably more challenging than had been anticipated. The fundamental problem the research assistants faced was that of identifying an electronic record in diplomatic terms. Although the research assistants had experience with the process of diplomatic analysis, they had only ever dealt with traditional paper records. To analyze the case studies in diplomatic terms, it was necessary first to penetrate the complexity of the electronic system and the surrounding record-keeping environment in order to establish whether records even resided within that system and, if so, to understand the specific ways in which they manifested themselves. To reach that understanding required a detailed knowledge of the electronic system and the record-keeping environment that was difficult to achieve. The difficulty stemmed in part from the fact that the knowledge had to be gleaned not on the basis of an examination of the system itself and the entities within it—which is the traditional diplomatic approach—but rather on the basis of the information found in the case study tools and related documentation.

Documentation from all sources was valuable in supporting the analysis of a given case study. The CSIP and its supporting documentation, however, did not provide enough information to enable the research assistants to gain a good understanding of the relationship between the electronic system and the business processes associated with it, and the relationship between the business processes and the types of records generated from them. Moreover, the supporting documentation was included only at the discretion of individual case study interviewers, who encountered issues not unfamiliar to institutional archivists responsible for appraising electronic records—up-to-date technological or procedural documentation may not exist, organizations or their units may be reluctant to provide copies of systems documentation for security or other reasons, and existing documentation may be intellectually inaccessible to the archivist for technical or language reasons. For some case studies, therefore, there was considerable supporting documentation; for others there was little or none. More supporting documentation did not always imply more and better information about the systems, however: some supporting documentation was difficult to understand and, in some cases, it was not clear why it was included at all. To fill in the gaps in their knowledge, the research assistants solicited the assistance of the interviewers who conducted the case studies. In some cases the interviewers were able to answer the questions put to them; in others, however, they were either unable to obtain the needed information, or unable to obtain it within the necessary time frame.

\(^{29}\) For definitions of these terms, see "Development of a Typology of Electronic Records" later in this section.

\(^{30}\) For definitions of these terms, see fn. 42.
As the analysis proceeded, it became increasingly clear that most of the systems under examination did not contain records, or at least, did not contain “good” records, when measured against the criteria established by contemporary archival diplomatics. In most cases this was because the entities identified within the electronic system did not appear to possess either a fixed documentary form or a stable content. To probe this situation further, UBC researchers decided to draft “scenarios” for certain cases. For those case study systems that had been found on first analysis to contain records, research assistants were instructed to answer the questions already devised. Reports for a number of the other case studies were drafted on the basis of two different scenarios: the first positing that “the system does not contain records, but if it did they could be analyzed diplomatically as follows . . .”; and the second positing that “the system does not contain records, but it should; it could be reconfigured such that it would contain records, as follows, and if this were done, the records could be analyzed diplomatically as follows . . .”

Given that the case studies so far had yielded very little information useful for the formulation of the requirements for assessing authenticity, the researchers also decided to incorporate into the analysis the procedural rules for creating and maintaining reliable and authentic electronic records that had been developed by the UBC Project. The procedural rules for creating and maintaining authentic records as laid out in the UBC Project were compared with data from each case study concerning the methods used by the creator to support its presumption of the authenticity of the records in the system under examination. On the basis of this comparison, the research assistants described the means currently in place that, from the creator’s perspective, supported a presumption of record authenticity; they also identified additional methods for supporting and strengthening such presumption, based on the procedural rules.

The development of hypothetical case studies and the comparative analysis of real-world data using the UBC procedural rules enabled the team to draft a preliminary set of conceptual requirements for presentation at the International Team workshop in October 2000. It was understood, however, that this was only a temporary solution and that the team would have to adjust the case study process to achieve better results from the diplomatic analysis in subsequent rounds. Accordingly, two changes were made to the process of conducting case studies.

The first change concerned the kinds of systems that would be targeted in subsequent rounds. Given that the majority of systems that the Authenticity Task Force had examined thus far had not contained records when viewed from the perspective of contemporary diplomatics, task force researchers were faced with two choices. On the one hand, they could revise the eligibility criteria for treating the entities within electronic systems or the electronic systems themselves as records to accommodate the various dynamic realities they were seeing; on the other, they could circumscribe the range of case studies to accommodate only systems that contained entities fitting the diplomatic construct of a record. The researchers opted for the latter route on the grounds that one of the reasons for choosing diplomatics as a means of analyzing electronic records was to evaluate its effectiveness. The researchers needed to examine a range of systems that fitted the construct in general terms before they could evaluate its effectiveness in more specific terms. Accordingly, the case study selection criteria were adjusted to ensure that only electronic systems containing, or having the potential to contain, records were selected in subsequent rounds.

The second change concerned the designation of responsibility for preparing the TEDGIs. It was decided that the UBC research assistants would prepare the TEDGIs because their knowledge of diplomatics made them the best equipped for mapping the answers to questions on the CSIP to the relevant archival-diplomatic element of the TEDGI. Once the research assistants had completed the TEDGIs, they were required to send them back to the case study interviewers for verification of the accuracy of the mapping before finalization. The TEDGIs subsequently became the basis for the preparation of draft versions of the diplomatic analyses of case studies, which

---

31 The procedural rules may be found on the Web site of the UBC Project at <http://www.slais.ubc.ca/people/faculty/duranti/duranti.htm >.

32 See above, fn. 25.
were also returned to the case study interviewers for approval prior to being finalized. It was decided also that the experiment of developing hypothetical scenarios would not be repeated in the next rounds. While the exercise had helped the research assistants to understand how authenticity requirements might manifest themselves in a given situation, the International Team found them to be confusing and overly prescriptive.

Apart from these changes, the process of diplomatic analysis in the second and third phases (rounds three and four) was similar to the process in the first phase (rounds one and two). The main difference was that case studies from rounds three and four contained more systems with records.

Twenty-two case studies were analyzed from an archival diplomatic perspective. Of these, twelve systems were found to contain records. In the systems containing records, many of the elements associated with uniquely identifying a record and placing it in its immediate context were either implicit or absent. For example, in most of the systems there was no explicit manifestation of the archival bond between and among the records participating in the same action. Moreover, while it was reasonably straightforward to identify the business processes supported by the electronic system in general terms, it was not always easy to determine how the records participated in or supported specific actions.

Authenticity Task Force researchers had hypothesized at the outset of the research that intrinsic and extrinsic elements of documentary form and annotations would play key roles in establishing the identity and demonstrating the integrity of electronic records. This hypothesis failed to be supported, however, by either the diplomatic analysis or the analysis of elements relating to the identity and integrity of records described later in this document. In the case studies analyzed, it was often difficult to determine the significance of the presence or absence of annotations or specific elements of documentary form. The determination of documentary forms in general and the establishment of required elements of form in particular appeared to be deeply embedded within specific institutional and procedural contexts and were resistant to any easy generalizations. As a result, the researchers were unable to draw any general conclusions about the relevance of specific intrinsic and extrinsic elements of documentary form or annotations to a consideration of an electronic record’s authenticity outside of the specific institutional and procedural context in which the record was created (this is discussed further in this section).

At the same time, however, it was possible to identify certain commonalities in the means used by creators to protect record authenticity from one institution to the next. The diplomatic analysis and the analysis of elements relating to identity and integrity revealed that record creators tend to rely on procedural means for protecting authenticity and to treat it as part of the management of the electronic system as a whole rather than as part of the management of individual records within the system. The commonest means identified were access privileges (including passwords, user IDs, and user profiles), followed by the use of audit trails and back-up procedures.

Development of a Typology of Electronic Records

The diplomatic analyses of case studies were undertaken to make it easier to identify general conceptual requirements for authenticity and to develop a typology of electronic records based on authenticity requirements for specific types of electronic records. The primary purpose of any typology is “to produce ordered and reproducible sets that can support the rapid identification of members of groups of sets in general and members of individual sets or subsets in particular.”

The design and implementation of a typology may be approached from the top down or the bottom up. As Seamus Ross explains:

In the former approach a researcher begins within the premise that a ‘group of entities’ . . . forms a bounded set. Then the researcher attempts to select and define characteristics shared by the material and to determine whether

objects/entities proposed as members of the group have the required attributes. In this approach the set becomes equivalent with the type. In the second approach the investigator starts with the objects and proceeds to describe the component elements. The elements are then grouped into attributes and the attributes subsequently grouped into restricted sets. These are shared component types that carry meaning.\textsuperscript{34}

The criterion for developing the typology of electronic records was the significance of specific extrinsic and intrinsic elements of documentary form and annotations for carrying out or attesting to the action or matter in which a record participated. Between October 2000 and April 2001, the task force explored numerous candidate types based on a range of criteria. A top-down approach was adopted for the identification of these types, mainly because there were insufficient data from the case studies to support a bottom-up approach.

The initial basic typology reflected the four categories of records identified by contemporary archival diplomatics, based on the relationship between a record and the action in which it participates. This categorization was chosen on the grounds that groups of records sharing the same function with respect to an action or matter form a bounded set. The categories are dispositive records (records whose written form is required by the juridical system as the essence and substance of an action), probative records (records whose written form is required by the juridical system as proof that an action has taken place prior to its documentation), supporting records (records whose written form is discretionary; they are created to provide support for, and are procedurally linked to, an action), and narrative records (records whose written form is also discretionary; they do not participate procedurally in the action but are created as part of the process of setting oneself to work). The task force extended these categories so that they referred to the smallest indivisible aggregation of records (e.g., the file unit) in each system rather than to individual records. This definitional extension of the record categories implied an extension of the authenticity requirements because the requirements for a given category of record aggregation (dispositive, probative, supporting, narrative) would apply to all the records within the aggregation, regardless of the different types of individual records contained within it.

The task force hypothesized that for dispositive and probative aggregations of records—that is, records whose written form is required—the elements of extrinsic and intrinsic form as well as annotations would be prescribed by the juridical system and therefore would have to be preserved in their entirety. For supporting and narrative aggregations of records—that is, records whose written form is not required—it was assumed that there would not be the same necessity to preserve all the elements and annotations. This hypothesis was not, however, supported by the case studies which suggested that: (1) the requirement of a written form does not necessarily translate into specific required elements of documentary form or annotations; and (2) the fact that a written form is not required does not necessarily translate into an absence or reduction of specific required elements of documentary form or annotations since there are cases of supporting and narrative records whose written form is highly regulated.

Next, the task force explored the possibility of a typology of electronic records based on the diplomatic categorization of procedures. These include: constitutive procedures (procedures that create, extinguish, or modify the exercise of power and that may be further subdivided into procedures of concession, of limitation, or of authorization); executive procedures (procedures that allow for the regular transaction of affairs according to rules established by an external authority); instrumental procedures (procedures connected to the expression of opinions and advice); and organizational procedures (procedures whose purpose is to establish organizational structure and internal procedures and to maintain, modify, or extinguish them). Because the categories of procedure imply different levels of documentary control, with constitutive procedures being the most controlled and instrumental procedures the least controlled, the task force hypothesized that records created in accordance with the more controlled procedures would have more required elements of documentary form and annotations than would those created in

\textsuperscript{34} Ross, 86.
accordance with less controlled procedures. This categorization was ultimately rejected, however, on the grounds that (1) records do not necessarily aggregate in accordance with these procedures; and (2) it is not possible to generalize, simply on the basis of the procedure, about the significance of elements of documentary form and annotations.

The task force experimented with a number of other candidate types, based on a range of criteria, such as whether the system contained records or was itself a record; whether the system contained one type of records, or many types; whether the records were digital or digitized, and so on. None of these types, however, resulted in a categorization of records on the basis of which specific requirements for authenticity could be formulated.

By April 2001, the task force had not yet succeeded in developing a typology that provided a meaningful differentiation and specification of requirements for authenticity according to types of records.35 Despite our efforts, we were simply unable to establish a correlation between authenticity and the presence of specific documentary elements or annotations. Since the deadline for submitting the final version of the requirements for authenticity was June 2001, the task force decided to suspend its efforts to develop a typology and to focus instead on refining the general conceptual requirements for assessing the authenticity of electronic records.

Additional analyses of case studies

In addition to the diplomatic analysis, and in order to support the Authenticity Task Force’s theory-building efforts, four other types of analysis were performed at UCLA and the University of Albany on some or all of the case studies.

Prior to these analyses, TEDGI and selected CSIP data were entered into a database and interview tapes, where available, were transcribed.

A preliminary analysis of completed TEDGIs and supporting documentation was undertaken in order to:

1. create a narrative overview for each case study;
2. generate tables of how each TEDGI element was completed across case studies;
3. verify how each TEDGI element was completed and attempt to reconcile any differences between the interviewer completing the TEDGI in the first two rounds and the research assistants analyzing the data;
4. identify which questions were used to support completion of which elements, and which questions were seldom, if ever used;
5. identify what supporting documentation was used to complete which elements;
6. identify elements that interviewers had difficulty completing, or where there appeared to be little consistency in how they were completed;
7. track the numbers of interviews and the position titles of interviewees necessary to complete case studies.

A detailed analysis of completed TEDGIs and supporting data was then undertaken to identify:

1. the elements that are most commonly present across case studies;

35 Researchers working on the D(igital) A(rchivering in V(laamse) I(nstellingen en) D(iensten) Project in Brussels reached a similar conclusion. The original aim of the DAVID Project “was to work out a typology from which a method for preserving the various types of digital archive documents over the long term would follow.” According to the researchers, “[t]his typology would stand or fall on its usefulness in formulating a preservation strategy, and was pursued with this goal in mind. The first attempt rested on the editorial form and function of the digitally preserved document, a method of description and classification borrowed from paper archiving. It was soon obvious, however, that this was no basis for managing digital archives and no basis for formulating a preservation strategy.” See Filip Boudrez, “The Digital Recordkeeping System: Inventory, Information Layers, and Decision-Making Model as Point of Departure” (Antwerp, June 2001), 4, at <http://www.antwerpen.be/david>. 
2. the elements that are most commonly absent, or cannot be discerned across case studies;
3. the business functions being supported by the electronic systems studied;
4. the activities and transactions performed by the electronic systems in support of the business functions;
5. at which level within the organization the electronic systems exist;
6. the relationships between paper and electronic components of record-keeping systems.

A narrative analysis of selected transcribed interviews was also undertaken to identify:

1. the ways that records creators, custodians, and systems personnel conceptualize the nature and role of the electronic records and/or record-keeping system being studied;
2. the variances in language used to describe records by records creators, custodians, and systems personnel;
3. the extent to which the findings of 1) and 2) should or could be factored into the design of a method to identify and ensure the preservation of authentic electronic records.

As outlined above, upon commencing the data analysis, researchers first created a brief narrative description of the case study, referred to as the case study overview, based upon the documentation submitted for analysis by the interviewers. The draft overview was then returned to the interviewers for them to review together with the respondents and make any necessary corrections that might be due to misinterpretation of the case study data.

The researchers then proceeded to undertake the following four analytical activities:

i) Analysis of how and to what degree the identity and integrity of electronic records is supported within and across case studies. In undertaking the diplomatic analysis of the case studies, the researchers had begun with an assumption that the diplomatic elements of electronic records would be the same (or at least the fundamental elements would be similar) as those of traditional records. However, researchers began to realize that these elements are less explicit in electronic records, and that more of the record’s identifying elements are found in its context, instead of on the face of the records, as was the case for traditional records. As a result, the diplomatic analysis often focused on what was wrong with the systems that were studied when held up against the ideal record represented by the Template, rather than effectively identifying alternative, new, or unanticipated ways in which authenticity requirements were being met in these systems. To help the task force’s efforts to develop a typology of authenticity requirements for electronic records, therefore, each case study was analyzed in order to determine which—if any—aspects of the systems examined corresponded to or supported elements establishing the identity and integrity of electronic records (the key concerns of authenticity) as delineated in the Template for Analysis. This analysis examined not only specific elements, but a variety of contexts, sources, and techniques through which elements might be manifested or their purposes achieved.

The case study data were coded to see whether any patterns were discernible, across all case studies, or across those that seem likely to contain similar types of records. The resulting analysis showed that both within individual case studies and across all case studies, authenticity is assured mainly through procedural means and treated as part of the management of the electronic system as a whole.

ii) Characteristics of case studies by type of information system. This analysis applied a model commonly used in business administration to identify types of information systems developed and used in an organization to support business processes and to fulfill the mission of the organization. The model provided one way to describe the nature of systems found in an organization, and, thereby, potentially a method to help discern systems that are likely to create records, and whether those records are likely to be dispositive, probative, supporting, or narrative.

In this model, an organization is divided into four levels:

1. **Operational-level systems**: information systems that monitor the elementary activities and transactions of the organization.
2. **Knowledge-level systems**: information systems that support knowledge and data workers in an organization.
3. **Management-level systems**: information systems that support the monitoring, controlling, decision-making, and administrative activities of middle managers.
4. **Strategic-level systems**: information systems that support the long-range planning activities of senior management.

Organizational functions are supported by six major types of systems:

1. **Transaction processing system (TPS)**: computerized system that performs and records the daily routine transactions necessary to conduct the business; these systems serve the operational level of the organization.
2. **Knowledge work system (KWS)**: information system that aids knowledge workers in the creation and integration of new knowledge in the organization.
3. **Office automation system (OAS)**: computer system—such as word processing, electronic mail system, and scheduling system—that is designed to increase the productivity of data workers in the office.
4. **Management information system (MIS)**: information system at the management level of an organization that serves the functions of planning, controlling, and decision making by providing routine summary and exception reports.
5. **Decision support system (DSS)**: information system at the management level of an organization that combines data and sophisticated analytical models to support semi-structured decision making.
6. **Executive support system (ESS)**: information system at the strategic level of an organization designed to address unstructured decision making through advanced graphics and communications.

Operational level systems such as transaction processing systems help operational managers keep track of the organization’s everyday activities. Knowledge level systems such as office automation systems and knowledge work systems help knowledge and data workers design products, distribute information, and manage paperwork. Management level systems such as management information systems and decision support systems help middle managers monitor and control business activities. Strategic level systems such as executive support systems help senior managers with long-term planning. The model also delineates the information inputs, processes, and outputs that serve as indicators of the type of system being examined.

This approach closely parallels certain traditional appraisal approaches that have targeted executive and administrative levels within an organizational hierarchy as being most likely to generate key records relating to policy, procedural, and organizational decision making. In the model used in this analysis, the types of information systems commonly associated with these levels would be management information systems (MIS), decision support systems (DSS), and executive support systems (ESS).

This analysis examined the organizational level and the information inputs, processes, and outputs associated with each case study in an attempt to identify the type and nature of each system and the likelihood that it generates, or should generate, records. In the analysis, in recognition of the “mixed” nature of most of the systems studied, the researchers also extended the model to identify more closely both electronic and paper outputs. Because stable content is considered to be an identifying characteristic of authentic records, researchers further categorized the status of system outputs in order to understand the degree to which they were stable:

---

• Fixed: Once output is created, it is immutable. If it needs to be changed, either an update must be appended or a new version must be created.
• Transient: Output is created for temporary use only (e.g., a screen display providing the results of an information query).
• Dynamic: Output is stored on the system but can be changed, updated, annotated, and overwritten.

The analysis indicated the complexity of the systems studied. Almost no system exists independent of a wider record-keeping system, and most relate to more than one organizational level and perform a range of functions rather than conforming to one of the discrete types contained in the business model. Equally, most of the systems studied have paper as well as electronic components. This “mixed” environment must be taken into account when understanding the nature of any potential record generated by the system. Many of the systems studied contained primarily transactional data, and most of them generated primarily transient or dynamic output.

The majority of the case studies focused on systems that function at the operational and knowledge levels within the organization, and less frequently at the management level. Comparing analyses of the same case studies, we see that those systems identified through the diplomatic analysis as ones that contain or should contain dispositive or probative records for the most part carry out at least some management as well as operational and knowledge-level functions. We could speculate, therefore, that systems addressing functions at the management level and above would be more likely to contain records and less transactional data.

iii) Functional analysis of case studies. As the research progressed, it became increasingly clear that understanding the nature and boundaries of electronic records required a detailed understanding of the business functions and activities of the record-keeping systems being studied. Researchers at UCLA selected the method delineated in the National Archives of Australia’s DIRKS (Designing and Implementing Recordkeeping Systems) Manual38 as one of the most robust and replicable extant approaches to functional analysis. The purpose of conducting the InterPARES analysis was to describe—unambiguously for non-archivists, particularly systems designers—the nature of the system's record-keeping function. The researchers attempted a narrative and graphical representation of the major functions of the systems being studied, and a breakdown of the actions and transactions that support those functions. They then received feedback from interviewers and respondents upon the draft breakdowns. The researchers concluded that it was not possible to render an accurate functional decomposition of each case study because the CSIP, developed from the diplomatic perspective of analyzing individual documents, had not been designed to capture the appropriate depth of functional detail about the record-keeping system as a whole.

iv) Narrative analysis of transcribed case study interview data. The researchers were concerned that their understanding of the nature of electronic records and the concept of authenticity, as well as the way that understanding was expressed through the terminology used in the CSIP and any other InterPARES products, would not match that of, or be understandable by, record keepers and systems personnel. Although the case study interviews were heavily scripted to ensure some level of consistency, some interviews were recorded and transcribed (where respondents gave their permission) and of these interviews, some contained additional discussion about the nature and functionality of the electronic record keeping in which respondents were engaged. Selected case study transcripts were examined to gain a closer understanding of respondent perspectives and terminology. A complete narrative analysis was conducted of one case study; it demonstrated that, even though the process of transcribing and analyzing interview data is laborious and time-consuming, such an approach would be valuable for future research.39


39 The narrative analysis was undertaken by Ciaran Trace. Her report, “Applying Content Analysis to Case Study Data: A Preliminary Report,” is available at <http://www.interpares.org/reports.htm>.
case studies, however, were not originally intended to be subjected to narrative analysis. Had this been the case, interviews or components of interviews would need to have been conducted in a more free-form or conversational manner that would have allowed respondents to expand their commentary and would have avoided providing respondents with InterPARES’ own terminology and rhetorical tropes.

Research Findings

Preamble
The purpose for developing the Template for Analysis and testing its effectiveness through four rounds of case studies was to lay the foundation for establishing conceptual requirements for assessing and maintaining the authenticity of electronic records over the long term. The requirements are described in detail in Requirements for Assessing and Maintaining the Authenticity of Electronic Records (Appendix 2) and embody the major conceptual findings of the Authenticity Task Force.

Conceptual findings: the requirements for authenticity

Terms of assessment of authenticity
To assess the authenticity of an electronic record, the preserver must be able to establish its identity and demonstrate its integrity.

The identity of a record refers to the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records. From an archival-diplomatic perspective, such attributes include: the names of the persons concurring in its formation (i.e., its author, addressee, writer, and originator); its date(s) of creation (i.e., the date it was made, received, and set aside) and its date(s) of transmission; an indication of the action or matter in which it participates; the expression of its archival bond, which links it to other records participating in the same action (e.g., a classification code or other unique identifier); as well as an indication of any attachment(s) since an attachment is considered an integral part of a record.

The integrity of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects. This does not mean that the record must be precisely the same as it was when first created for its integrity to exist and be demonstrated. Even in the paper world, with the passage of time, records are subject to deterioration, alteration and/or loss. In the electronic world, the fragility of the media, the obsolescence of technology, and the idiosyncrasies of systems likewise affect the integrity of records. When we refer to an electronic record, we consider it essentially complete and uncorrupted if the message that it is meant to communicate in order to achieve its purpose is unaltered. This implies that its physical integrity, such as the proper number of bit strings, may be compromised, provided that the articulation of the content and any required elements of form remain the same.

Assessment and maintenance of authenticity
The preserver must assess the authenticity of electronic records before electronic records are transferred to archival custody, and maintain it after transfer. The assessment is an integral part of the records’ appraisal while the maintenance is an integral part of their long-term preservation.

Before records are transferred to archival custody, the preserver must establish, as part of the process of appraisal, whether and to what extent the records have been maintained by the creator using technologies and administrative procedures that either guarantee their authenticity or at least minimize risks of change from the time the records were first set aside to the point at which they are subsequently accessed.
After the authenticity of the creator’s electronic records has been established in the appraisal process, and the records transferred from the creator to the preserver, the preserver needs to maintain the authenticity of the records. To do so, the preserver must maintain the electronic records in accordance with procedures that ensure their continuing authenticity and produce copies of those records in accordance with procedures that ensure that their authenticity is not compromised by the reproduction process. To support its attestation of the authenticity of copies of electronic records, the preserver must also produce and maintain documentation relating to the manner in which it has maintained the records over time as well as the manner in which it has reproduced them.

In light of the above, the Authenticity Task Force has developed two sets of requirements. The first set, termed “benchmark requirements,” includes requirements that support the presumption of the authenticity of the creator’s electronic records before they are transferred to the custody of the preserver. The second group, “baseline requirements,” includes requirements supporting the production of authentic copies of electronic records that have been transferred to the custody of the preserver.

Conceptual framework of the benchmark and baseline requirements

Both the benchmark and the baseline requirements are based on the notion of trust in record keeping and record preservation. The benchmark requirements draw specifically on the notion of a trusted record-keeping system, and the baseline requirements are predicated on the role of the preserver as a trusted custodian.

A trusted record-keeping system has been defined as “a type of system where rules govern which documents are eligible for inclusion in the record-keeping system, who may place records in the system and retrieve records from it, what may be done to and with a record, how long records remain in the system, and how records are removed from it.”

The role of the preserver as trusted custodian dates back to Roman antiquity, when citizens would deposit private records in the Tabularium for the express purpose of rendering them authentic. As the trusted custodian of records, ancient archival institutions sustained and lent credibility to contractual relationships between citizens. They also lent credibility to the implicit social contract between citizens and the state by preserving the records of the state’s past actions on the basis of which the state could be held to account. Today, the role of the preserver as a trusted custodian is also analogous to that of the trusted third-party record keeper in electronic contracting. A trusted third-party record keeper is a physical or juridical person entrusted with independently maintaining the records of electronic data interchange (EDI) partners. The reason for having a trusted third-party record keeper is to increase the probability that records of an EDI transaction will be accepted in court as evidence. To be considered a trusted record keeper, the person must demonstrate, among other things, that it has no reason to alter retained records itself; that it has no interest in allowing others to alter records; and that it is capable of implementing security procedures to a degree that meets the necessary standards of integrity and accuracy.

Similarly, to be considered a trusted custodian, the preserver must demonstrate that it has no reason to alter the preserved records, or to allow others to alter them, and that it is capable of implementing the baseline requirements.

---


Specific conceptual framework for the benchmark requirements for assessing the authenticity of the creator’s electronic records

The creator’s records belong to one of two categories. The first comprises those records that exist as created; they are considered authentic because they are the same as they were in their first instantiation. The second category comprises those records that have undergone some change and therefore cannot be said to exist as first created; they are considered authentic because the creator treats them as such by relying on them for action or reference in the regular conduct of business. However, the authenticity of electronic records is threatened whenever they are transmitted across space (i.e., when sent to an addressee or between systems or applications) or time (i.e., either when they are in storage, or when the hardware or software used to store, process, or communicate them is updated or replaced). Given that the acts of setting aside an electronic record for future action or reference and of retrieving it inevitably entail moving it across significant technological boundaries (from display to storage subsystems and vice versa), virtually all electronic records belong to the second category. Therefore, the preserver’s inference of the authenticity of electronic records must be further supported by evidence—provided in association with the records—that they have been maintained using technologies and administrative procedures that either guarantee their continuing identity and integrity or at least minimize risks of change from the time the records were first set aside to the point at which they are subsequently accessed. The requirements for assessing the authenticity of the creator’s electronic records concern this evidence.

The presumption of authenticity

A presumption of authenticity is an inference that is drawn from known facts about the manner in which a record has been created, handled, and maintained. The evidence supporting the presumption that the creator created and maintained its electronic records authentic is enumerated in the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records (Requirement Set A). A presumption of authenticity will be based upon the number of requirements that have been met and the degree to which each has been met. The requirements are, therefore, cumulative: the higher the number of satisfied requirements, and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity. This is why these requirements are termed “benchmark” requirements.

The verification of authenticity

In any given case, there may be an insufficient basis for a presumption of authenticity, or the presumption may be extremely weak. If this is so, further analysis may be necessary to verify the authenticity of the records. A verification of authenticity is the act or process of establishing a correspondence between known facts about the record and the various contexts in which it has been created and maintained, and the proposed fact of the record’s authenticity. In the verification process, the known facts about the record and its contexts provide the grounds for supporting or refuting the contention that the record is authentic. Unlike the presumption of authenticity, which is established on the basis of the benchmark requirements, this verification involves a detailed examination of the records themselves and reliable information available from other sources about the records and the various contexts in which they have been created and maintained. Methods of verification include, but are not limited to, a comparison of the records in question with copies that have been preserved elsewhere or with back-up tapes; comparison of the records in question with entries in a register of incoming and outgoing records; textual analysis of the record’s content; forensic analysis of aspects such as medium and script; a study of audit trails; and the testimony of a trusted third party.
Specific conceptual framework for the baseline requirements supporting the production of authentic copies of electronic records

After the records have been presumed or verified authentic in the appraisal process, and have been transferred from the creator to the preserver, their authenticity needs to be maintained. In order to do this, the preserver must carry forward the records in accordance with the baseline requirements that apply to the maintenance of records, producing copies according to procedures that also maintain authenticity. The production of authentic copies is regulated by the Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records (Requirement Set B). Unlike the benchmark requirements, all of the requirements included in the baseline requirements must be met before the preserver can attest to the authenticity of the electronic copies in its custody (hence the word “baseline”).

Satisfaction of these baseline requirements will enable the preserver to certify that copies of electronic records are authentic. Traditionally, the official preserver of the records has been the person entrusted with issuing authentic copies of such records. To fulfil that role, the preserver needed simply to attest that the copy conformed to the record being reproduced. With electronic records, the difficulties related to preservation make it prudent for the preserver to produce and maintain documentation relating to the manner in which it has maintained the records over time as well as the manner in which it has reproduced them to support its attestation of authenticity.

A copy is the result of a reproduction process. A copy can be made from an original or from a copy of either an original or another copy. 42 There are several types of copy. The most reliable copy is a copy in the form of an original, which is identical to the original although generated subsequently. An imitative copy is a copy that reproduces both the content and form of the record, but in such a way that it is always possible to tell the copy from the original. A simple copy is a copy that only reproduces the content of the original.

Any of these types of copy is authentic if attested to be so by the official preserver. By virtue of this attestation, the copy is deemed to conform to the record it reproduces until proof to the contrary is shown. Such attestation is supported by the preserver’s ability to demonstrate that it has satisfied the applicable baseline requirements for maintenance and all of the requirements for the production of authentic copies.

Methodological findings

Although the primary purpose of the work of the Authenticity Task Force has been to address authenticity requirements for electronic records, a significant by-product of its work, and indeed that of the entire InterPARES Project, has been an enhancement and extension of existing archival methodological knowledge and expertise. Drawing upon the multidisciplinary expertise of its researchers, InterPARES applied a diverse range of theoretical and applied approaches, including diplomatic analysis, modelling, and narrative analysis. This diversity of approaches was unprecedented in archival research to date, and throughout its work, the Authenticity Task Force

42 In common language, *copy* and *reproduction* are synonyms. For the purposes of this research, the term *reproduction* is used to refer to the process of generating a copy, while the term *copy* is used to refer to the result of such a process, that is, to any entity which resembles and is generated from the records of the creator. An original record is the first, complete record, which is capable of achieving its purposes (i.e., it is effective). A record may also take the form of a draft, which is a temporary compilation made for purposes of correction. For a discussion of the status of originals, drafts, and copies in an electronic environment see Luciana Duranti and Heather MacNeil, “The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project.” *Archivaria* 42 (Fall 1996): 56–57. For the definition and interpretation of an original in the context of international law and electronic commerce, see United Nations Commission on International Trade Law, *UNCITRAL Model Law on Electronic Commerce with Guide to Enactment* (New York: United Nations, 1997), esp. article 8 of the “Model Law” and para. 62–69 of the “Guide to Enactment.”
strived to assess and document what worked in the different methods that it used, what partially worked, what did not work— and why.

**Limitations of diplomatics as an analytical tool**

One reason for incorporating the perspective of contemporary archival diplomatics into the work of the Authenticity Task Force was to evaluate its effectiveness as an analytical tool. The Authenticity Task Force found it to be a useful means of assessing the strengths and weaknesses of current electronic systems. For example, it highlighted the extent to which electronic systems are still being designed to manage data rather than records. This appears to be the case even when the purpose for which the system is designed would appear to require the creation and maintenance of fixed records rather than fluid data. It also highlighted the significant extent to which elements relating to a record’s identity are implicit rather than explicit; and the consequent need to make certain identifying elements are explicit to ensure that knowledge of key indicators of identity is not lost when the records are removed from the system in which they have been created and actively used. Finally, the diplomatic analysis revealed a surprising level of indifference on the part of record creators to authenticity-related issues, an indifference attributable mainly to a (possibly misplaced) confidence in the capacity of generic technological and procedural controls over the electronic system to protect the authenticity of the records contained within it.

At the same time, contemporary archival diplomatics remains rooted in a very traditional conception of what a record is and is thus limited in its capacity to extend the range of archival understanding about the nature of different kinds of electronic systems and the variety of entities contained within them. While it is quite effective in analyzing electronic environments that are analogous to traditional record-keeping environments, it is considerably less helpful in analyzing electronic environments that are not so analogous. This finding points to the limits of the known as an aid to understanding the unknown. Increasing the utility of diplomatics as an aid to understanding diverse electronic systems will require the development of a more nuanced interpretation of the characteristics of electronic records and the manner in which they manifest themselves in a variety of electronic environments. The Authenticity Task Force began to move in this direction in the final two rounds of case studies, where it focused less on establishing whether the record was complete, stable, and unchangeable, and more on determining whether and to what extent the system was capable of tracking and preserving any changes; however, considerably more interpretive work is needed.

The limitations of the diplomatic model of a record as it is elaborated in the Template for Analysis are attributable mainly to the fact that the model was built on the premises of general diplomatics. **General diplomatics** seeks to decontextualize records, to eliminate their particularities, variations and anomalies in the interest of identifying the common, shared elements of records that cut across juridical, provenancial, and technological boundaries. Given the complexity and variety of electronic systems it might make more sense to adopt and adapt the approach of **special diplomatics**, which, traditionally, has focused on the records of individual chanceries and specific juridical systems. In such an approach, one would begin with an analysis of the various features of the systems themselves and the broader record-keeping environment in their own terms, with all their particularities, variations, and anomalies; and, on the basis of that analysis begin to build a more general framework.

Further refinement of the diplomatic approach is also needed to accommodate record aggregates. One significant difference between the diplomatic and the archival perspective is that diplomatics focuses mainly upon the individual record, while archival science tends to emphasize the record aggregate (e.g., files, series, and fonds). Although researchers attempted, during the development of the Template, to incorporate the aggregate approach, the Template remained predominantly focused on elements that are relevant only at the level of individual records. Many

---

43 The typology of papal chancery documents prepared in the course of the Authenticity Task Force’s work is an example of special diplomatics. See fn. 14.
of the systems examined through the case studies, however, contained heterogeneous aggregates of records. In fact, the archival extensions of the Template, such as the addition of the five categories of context (juridical-administrative, provenancial, procedural, technological, documentary), turned out to be the most relevant to an understanding of the record-keeping environment, and the grounds on which creators based their presumption of the records’ authenticity. These contexts were, however, the least well developed part of the Template. For example, in several case studies, audit trails were identified by the creator as a significant means of ensuring the authenticity of electronic records. Audit trails are part of system administration and therefore were considered an element within the record’s technological context. The element “system administration” was not decomposed sufficiently, however, to enable task force researchers to identify the various kinds of audit trails and the specific purposes they serve in a given environment. In the absence of that identification, it was difficult to assess the extent to which an audit trail supported the creator’s presumption of authenticity in particular cases.

To deal more effectively with such systems, therefore, the researchers believe that a contemporary archival diplomatic analysis should seek to identify and elaborate in a more comprehensive way the nature of archival aggregates and the elements that uniquely characterize them.

Limitations of case study design and instrumentation
One objective of the case studies was to make recommendations about the development of procedures, instrumentation, and analytical techniques to assist archivists and records managers in gathering and assessing the information they need in order to identify and preserve authentic records in electronic systems. Although the task force found that the case study method was a valuable approach to understanding the nature of the electronic record, the analysis of the successive rounds of case studies pointed up several areas where the design and instrumentation of the case studies were problematic or could effectively be refined. On the basis of a tandem evaluation of instrumentation and method after the analysis of each round of the case studies, the instrumentation and protocol were modified accordingly and tested through implementation in the succeeding round.

Issues that arose with the case studies included the following:

- The contact phase was often lengthy. It consisted of getting permission from the relevant institution and administrators to conduct the case study, and then identifying the appropriate respondents, prior to conducting the actual interview or interviews. Combined with the time it took to revise and get human subjects’ approval for the case study protocol; difficulties scheduling the interviews; the interviewers’ time to complete the TEDGI; and the time it took to analyze the resulting data, this contributed to difficulties in keeping to the tight schedules the task force had identified for each round of case studies. It is likely that a case study approach would also be time-consuming for practising archivists to implement when studying their institution’s records.

- The Case Study Interview Protocol (CSIP) was too long. It is a daunting instrument, both for interviewers and respondents. Notwithstanding this, a three-hour interview is generally insufficient to cover all the questions in detail and follow-up contacts are time- and labour-intensive for all parties concerned. Gathering sufficient information on electronic record keeping, whether for a research project or for institutional records management and archival purposes, is necessarily time-consuming and complex. Researchers, practitioners, and record keepers should not underestimate the resource-intensiveness of information gathering.

- Many questions from the CSIP were never used in populating the TEDGI and could potentially be eliminated from the CSIP, thus streamlining it further.

- The sometimes arcane terminology in the CSIP, drawing upon that of diplomacy, was not sufficiently oriented towards that used or understood by the people who are being interviewed. Moreover, some of the same terminology is used, but with different
meanings, by the systems community. While the establishment of a project glossary has sought to address these issues, they still potentially impeded the interview process with implications for the reliability of some of the answers obtained in the case studies.

- Making the translation between the CSIP and the Template Element Data Gathering Instrument (TEDGI) is difficult for anyone not trained in diplomatics even with the explanations provided with the Template for Analysis. Moreover, the correlation between the CSIP and TEDGI was not always clear. In many instances, the interviewers making the translation between the two instruments often got confused or simply made errors, thus potentially affecting the reliability of the data. Inconsistencies in how the TEDGI was completed became evident when completed TEDGIs were compared across case studies, requiring that the researchers analyzing the data go back to the CSIP, and sometimes also the interviewers, to verify how the TEDGI was completed. The translation process was also extremely time-consuming. In the final two rounds of the case studies these issues were addressed by having the TEDGI completed from the CSIP by researchers at UBC, rather than by individual interviewers.

- In situations where case study interviewers were also practising archivists, they brought to bear valuable experience and institutional knowledge, as well as archival expertise. Without such knowledge, some nuances of the records environment might be missed. The limitation of such situations, however, is that the interviewers have a considerable amount of implicit and unconscious knowledge, because of their familiarity with the institution and its records, that is not always captured overtly in case study data (especially the translation of the data into the TEDGI). The same would potentially be true if the interviewer came from a systems background. An alternative, although more labour-intensive approach might be to use pairs of interviewers with different backgrounds to conduct the interviews.

- It was often unclear what the focus of or unit of analysis for the case study was supposed to be—was it the entire record-keeping environment (i.e., the mandate, the business processes, the data input and output, whether electronic or paper) or the electronic system alone? Not all electronic records systems can be analyzed and managed at the same level of granularity. Sometimes it is possible to examine records document by document (e.g., with e-mail). In other cases, one has to approach an entire system (e.g., with databases). In yet other cases, a single record aggregate comprises both paper and electronic components. The best approach appears to be to commence with a thorough understanding of the record-keeping environment and then decide which types of records could be present and their intellectual and physical parameters. The CSIP, however, was weak in terms of collecting data that will allow for the analyses or understanding of specific record-keeping functions and events. It was therefore difficult both to achieve a functional analysis of the cases studied and to develop a typology based upon specific record-keeping acts or functions.

- Because it was first derived from diplomatics and what is known about traditional records, the CSIP unconsciously favoured record-keeping systems that look most like their paper counterparts. This made it difficult to understand whether the reason researchers were not finding specific diplomatic elements was because the elements were absent or because the researchers were not aware of how those elements might be manifested differently in electronic systems. It also meant difficulty in identifying whether elements were absent because the form or the record has changed or because the case studies were examining imperfect record-keeping systems that were not generating or maintaining good records.

Based upon its analysis of the quality and scope of the case study data, the task force makes the following recommendations for revisions to the design and instrumentation of record-keeping case studies:
1. Explicitly examine the entire record-keeping system, and not just its electronic components.
2. Reorient the CSIP to start with an analysis of business processes: proceeding from the general to the specific and delineating functions, activities, and then transactions. This will make it easier to identify actions in which records participate, and the nature of that participation.
3. Adjust terminology in the CSIP to reflect the language of records creators and systems managers more closely. The Glossary Committee may be able to provide some insight into how to map between the terminology used by interviewees and the terminology being used by the InterPARES Project. Additional analysis of the transcribed tape recordings of the case studies to date should also assist with this aspect.
4. Rewrite the CSIP questions so that they contain the definitions of the terms, rather than the actual glossary terms used in the Template of Analysis.
5. Eliminate questions that have failed to yield useful data through three or more rounds of case studies.
6. Interviewers’ comments from the CSIP and TEDGI suggest that a high level of technical expertise is needed to understand the systems being studied. In a best-case scenario, interviews should be conducted with both an archivist and an IT or computer professional.

Relationship Between Conceptual Requirements for Authenticity and Existing Standards

Preamble
In order to place its conclusions in context, the task force has conducted comparative analyses of the authenticity requirements against three prominent records management standards: the International Organization for Standardization’s (ISO) Draft International Standard on Records Management; the U.S. Department of Defense’s (DoD) 5015.2 Records Management Standard; and the European Commission’s (EC) Model Requirements Specification (MoReq). Each of the “mapping documents” produced in this exercise has been designed to reveal the extent of similarity between the Authenticity Requirements, on the one hand, and the particular standard under examination, on the other. The ISO and EC mapping documents identify provisions that can be considered as counterparts to the individual InterPARES benchmark requirements, while the DoD mapping locates provisions that function parallel to the stipulations contained in both the benchmark and the baseline requirements.

The mapping documents provide a basis for comparison from both microscopic and bird's-eye perspectives. With respect to the former, each mapping reproduces or summarizes individual provisions from the ISO, DoD, or EC standard alongside the particular authenticity requirement to which they relate. Thereby, the mapping documents allow for assessment of how InterPARES requirements are expressed differently from, and similarly to, pertinent provisions of the existing standards. At the same time, the documents can be used to make more general comparisons in that they reveal an overall portrait of the relationship between the InterPARES requirements and the three existing standards in question. For instance, the mapping documents demonstrate how many of the InterPARES requirements have counterparts in, respectively, the ISO, DoD, and EC standards.

A brief summary of the findings of each mapping exercise is presented below. Please note that making identifications between provisions of different standards involves recognizing degrees of similarity, and is rarely a simple yes/no question. This is a result of several factors, such as the fact that the specific wording used in any given standard tends to be unique, and the fact that an idea or concept treated in one single provision by, for instance, the authenticity requirements

44 The mapping documents are available at <http://www.interpares.org/reports.htm>.
might be scattered among several provisions in the ISO, DoD, or EC standard. Therefore, this
text generally makes statements to the effect of “DoD provision X is a parallel (or counterpart) to
InterPARES requirement Y.” Such statements are understood to mean that a general
resemblance exists between provisions X and Y, not that they correspond directly and completely
with one another. Conversely, by avoiding statements such as “Fulfilment of DoD provision X
satisfies InterPARES requirement Y in all respects,” the text attempts to avoid suggesting that
necessary and complete correspondence is entailed in identification of counterparts and parallels.

on Records Management

The ISO draft standard is designed to provide “guidance on managing records of originating
organizations, public or private, for internal or external clients” by making recommendations “to
ensure that adequate records are created, captured and managed.” In section 6, the ISO
standard indicates that organizations should “establish, document, maintain and promulgate
policies, procedures and practices for records management, the objective of which should be the
creation and management of authentic, reliable and useable records, capable of supporting
business functions and activities for as long as they are required.”

There are at least two noteworthy features of ISO/DIS 15489. First, while it provides a
considerable amount of technical detail in specifying required software functionalities, the
standard also addresses matters such as organizational policies and procedures. This
establishes an extent of similarity between the ISO standard and the InterPARES authenticity
requirements in that both sets of guidelines take into account the need for combining automated
and manual implementation methods. Second, a particular section of the ISO standard is devoted
to emphasizing the importance of record authenticity. This implies further common ground with
the InterPARES requirements, although, in accordance with the orientation of 15489 as a whole,
the pertinent sections treat this matter only as it relates to active records.

Of the eight benchmark requirements, only “A.6 Authentication of Records” and “A.7 Identification
of Authoritative Record” were found not to have counterparts within ISO/DIS 15489. On the other
hand, parallel provisions from the ISO standard have been identified for each of the remaining six
benchmark requirements. Note, though, that the counterpart for InterPARES Requirement A.1
concerning “Expression of Record Attributes and Linkage to Record” does not specify any
particular metadata fields for capture. Instead, this stipulation indicates that organizations should
determine what metadata is required according to their business needs and regulatory
circumstances: “To support the continuing conduct of business [and] comply with the regulatory
environment . . . organizations should institute and carry out a comprehensive records
management programme which includes . . . determining what metadata should be created with
the record and through records processes and how that metadata will be persistently linked and
managed.”

In consideration of these parallels, it can be said that full compliance with the 15489—or partial
compliance, if all provisions listed in the ISO mapping document were to be satisfied—would
result in satisfaction of Requirements A.1 through A.5, and A.8. However, note also that ISO
section 7.2.1 on “Authenticity” has been identified as parallel to Requirements A.1 and A.2 only,
suggesting that although certain ISO provisions satisfy Requirements A.3, A.4, A.5, and A.8,
these would support authenticity only in an implicit fashion.

---

45 International Organization for Standardization, Technical Committee ISO/TC 46 Information and
Documentation, Subcommittee 11, Archives/Records Management, International Organization for
Standardization Draft International Standard (ISO/DIS 15489) Information and Documentation—Records
document.

United States Department of Defense. Standard for Electronic Records Management Software Applications

The so-called Design Criteria Standard, better known as DoD 5015.2, “sets forth mandatory baseline functional requirements, and identifies non-mandatory features deemed desirable” for procurement of records management application (RMA) software by agencies of the United States government. 5015.2 has been implemented in this context for the purpose of “assur[ing] efficient and effective records management.” The scope of the standard is restricted to management of active records, and, as a procurement standard, its contents focus almost exclusively on required system functionalities. 47

These observations suggest some of the ways in which DoD 5015.2 differs from the authenticity requirements, and, for that matter, from ISO 15489: specifically, the Design Criteria devotes primary attention to software specifications over procedures and other implementation means, and to methods over principles. Furthermore, the focus of DoD 5015.2, like the ISO standard but unlike the InterPARES requirements, is solely on active records. Finally, this standard is distinct in that it does not overtly address authenticity, the record quality of principal concern to the task force, anywhere in its terms.

The DoD standard features provisions that can be understood as counterparts to six of the eight benchmark requirements. The exceptions are Requirement “A.6 Authentication of Records,” which has no parallel provision in 5015.2, and Requirement A.1 on “Expression of Record Attributes and Linkage to Record,” which is satisfied in several respects, although not entirely due to the fact that A.1 mandates capture of certain metadata fields not covered in the DoD standard. As for the InterPARES baseline requirements, counterpart provisions have been identified for “B.1 Controls over Records Transfer, Maintenance, and Reproduction,” and “B.2 Documentation of Reproduction Process and its Effects.” No parallel stipulation was located for “B.3 Archival Description.” 48

European Commission. Model Requirements for the Management of Electronic Records (Interchange of Data between Administrations)

The so-called Model Requirements Specification, or MoReq, “focuses mainly on the functional requirements for the management of electronic records by an Electronic Records Management System (ERMS),” and is designed for use by public and private sector organizations that are either introducing an ERMS, or assessing one already in place. MoReq has been designed to be

47 United States Department of Defense, Assistant Secretary of Defense for Command, Control, Communications, and Intelligence Design Criteria, Standard for Electronic Records Management Software Applications (DoD 5015.2-STD) June 2001, “C1.1. Purpose.” The characterization of the purpose of 5015.2 presented here is based on 44 U.S.C. § 2902, the passage of the United States Code cited in the “C.1.1. Purpose” section of the standard. In full, this law reads as follows:

It is the purpose of this chapter, and chapters 21, 31, and 33 of this title, to require the establishment of standards and procedures to assure efficient and effective records management. Such records management standards and procedures shall seek to implement the following goals: (1) Accurate and complete documentation of the policies and transactions of the Federal Government; (2) Control of the quantity and quality of records produced by the Federal Government; (3) Establishment and maintenance of mechanisms of control with respect to records creation in order to prevent the creation of unnecessary records and with respect to the effective and economical operations of an agency; (4) Simplification of the activities, systems, and processes of records creation and of records maintenance and use; (5) Judicious preservation and disposal of records; (6) Direction of continuing attention on records from their initial creation to their final disposition, with particular emphasis on the prevention of unnecessary Federal paperwork; (7) Establishment and maintenance of such other systems or techniques as the Administrator or the Archivist considers necessary to carry out the purposes of this chapter, and chapters 21, 31, and 33 of this title.

“pragmatic” and “usable,” and its purpose is to ensure that an ERMS will “manage electronic records with the desired levels of confidence and integrity.”

Like DoD 5015.2, MoReq is a software specification, and accordingly it differs from the InterPARES requirements in that it explicitly focuses on system functionality over procedures, and on implementation methods over records management principles. The European Commission standard also shares a point in common with ISO 15489 in that it addresses authenticity of records directly. However, MoReq defines “authenticity” in a manner that may or may not match the InterPARES definition. Note as well that MoReq features a greater extent of variability than any of the other standards considered here, including the InterPARES requirements. Having been designed to acknowledge that “different countries have their differing traditions, views and regulatory demands for managing records,” the EC standard presumes that, prior to use, it will be tailored to the business needs and the legal-regulatory requirements bearing upon an organization.

MoReq counterparts have been located for seven of the eight InterPARES benchmark requirements. In the remaining case, several provisions from the EC standard are listed as parallel to Requirement A.1 on “Expression of Record Attributes and Linkage to Record.” However, there are metadata fields mandated for capture in InterPARES Requirement A.1 that are not specified in MoReq.

Relationship of Findings to Other Research Initiatives

Issues that relate to the preservation and authenticity of digital information objects are being addressed from several perspectives by current research projects. These projects include:

- CAMiLEON (Creative Archiving at Michigan & Leeds: Emulating the Old on the New), which is investigating the viability of emulation as a preservation strategy that maintains the intellectual content, structure, and “look and feel” of software-dependent complex digital objects. Researchers are also assessing user preferences for different versions of emulators that vary considerably in how they reproduce those objects (for example, by analyzing how users define the authenticity of objects running in their native software environment, running under emulation, and delivered in migrated versions).

- Cornell University’s PRISM Project, which focuses on policy enforcement for ensuring information integrity in the areas of preservation, reliability, interoperability, security, and metadata. PRISM is investigating the long-term survivability of digital information, reliability of information resources and services, interoperability, and security (including the privacy rights of users of information and the intellectual property rights of content creators), and the metadata that make it possible to ensure information integrity in digital libraries. As part of this project, PRISM researchers carried out one-on-one interviews and discussion groups at Cornell to characterize the current environment and identify

---


50 See “1.5 Emphasis and Limitations of this Specification,” “4.5 Authenticity,” and “13.1 Glossary.” The Glossary defines “authenticity” as “the quality of being genuine,” but the quality of genuineness is not itself defined in the MoReq Glossary, or elsewhere. Note also that MoReq does distinguish between “mandatory” and “desirable” requirements, but that organizations implementing MoReq may nevertheless modify mandatory functionalities, and even omit individual requirements, when custom-designing the specification to suit their business needs.


digital preservation requirements. They found that few formal policies are in place for distributed resources and that as a result the level of trust about the preservation of content is low.  

- The San Diego Supercomputer Center’s (SDSC) Collection-Based Persistent Archives and Archivists’ Workbench projects, which are engaged in deriving XML information models from collections of software-dependent data objects and developing tools that can be used to ensure preservation and access to those objects over time. The Persistent Archives approach is built around the Open Archival Information System (OAIS) reference model. It supports archival processes from accessioning through preservation and use, and it recognizes the importance of collection-based management. It also exploits inherent hierarchical structures within records, predictable record forms, and dependencies between them. It is designed to be consistent, comprehensive, and independent of infrastructure.

- The Cedars Project (CURL exemplars in digital archives), which seeks to address strategic, methodological, and practical issues and provide guidance in best practices for digital preservation. Cedars is a United Kingdom collaboration of librarians, archivists, publishers, authors, and institutions (libraries, records offices, and universities). Working with digitized and born-digital materials, Cedars is using a two-track approach to evaluate different preservation strategies through demonstration projects at U.K. test sites; develop recommendations and guidelines; and develop practical, robust, and scaleable models for establishing distributed digital archives. Cedars is also examining other issues related to the management of digital information, including rights management and metadata.

Only one of these research initiatives, the Persistent Archives research at SDSC, focuses specifically on the preservation of electronic records. However, the task force believes that the delineation of the nature of electronic records and the conceptual requirements for authenticity provide a rigorous framework for approaching issues of preserving the integrity of complex digital objects in general, and electronic records in particular that could be applied in all these research initiatives.


Conclusion

Electronic records are very complex physical objects and intellectual constructs. Both the deductive and the inductive approaches employed by the Authenticity Task Force have constructed a detailed profile of the complexity of contemporary electronic records and identified the ways that they are embedded in their juridical-administrative, provenancial, procedural, documentary, and technological contexts.

In terms of what the Authenticity Task Force learned relating to issues of authenticity, we found that most contemporary records systems are a hybrid of electronic and paper records; that few explicit measures are employed to ensure the authenticity of electronic records; and that authenticity is generally assured through procedural means.

Although the task force developed a conceptual framework for establishing the requirements for preserving authentic electronic records, it failed to create a single, comprehensive typology of authenticity requirements for electronic records. It identified several possible perspectives from which a typology could be constructed, but none of these can be developed in such a way that it can be thorough, deep, and predictive. It seems likely that a typology based upon individual creators and the acts, procedures, and functions they perform would be the single most effective approach. Potentially such typologies could be generalized to other similar settings, but this generalizability would be limited because each creator interprets his or her own juridical context differently and implements it differently procedurally. The task force, however, has not at this point collected the necessary data to support such an hypothesis.

In terms of methodological outcomes, the task force found that because of the complexity of electronic records and record keeping, it is both difficult and problematic for those researching or managing electronic records to identify a single, appropriate unit of analysis. Diplomatics approaches the issue from the perspective of the individual record; archival science, from that of the record aggregate; and systems analysis, from that of the automated information or record-keeping system. Each of these perspectives contributes to both understanding the nature of the record and its long-term preservation. What is also required, however, is an overall systems approach that takes into account the total record-keeping environment, that is, the sum of all of the contexts identified through InterPARES.

Areas for Further Research

Several areas that need further research were identified in the work of the Authenticity Task Force. In some cases, these areas amounted to a more sophisticated formulation of questions that initially InterPARES sought to address but were found to be beyond the scope of a three-year research project. In other cases, new areas emerged in the course of the project. Research questions within these areas can be grouped under three rubrics: theory of the record, technological development, and record-keeping policy:

Theory of the record
1. Is it possible to develop an analytical framework integrating aspects of contemporary diplomatics and archival theory that addresses both the document and record aggregates and identifies and elucidates the role of the different contexts of the records in relation to individual records as well as record aggregates?

2. Can we provide a more detailed analysis of the various contexts in which records are created, maintained, and used, and the ways in which the archival bond might be expressed within those contexts? Can we develop more finely grained instruments that could extract specific aspects of different contexts and tie them more closely to the records?
3. Is it possible to develop meaningful typologies of records of specific creators or specific acts, procedures, and functions?

**Technological development**

1. Digital signature technologies have been implemented for the authentication of records across space, but what are the implications of their use for the management of authentic electronic records over the long term? Will their implementation impede the long-term management of authentic electronic records? Can the use of digital signatures be adapted and extended to support the long-term preservation of authentic electronic records. What specific adaptations and extensions would be necessary?

2. A related question concerns the infrastructure supporting digital signature technologies. The authority of a digital signature depends on the existence of a public key infrastructure (PKI), which is a hierarchical organization of certification authorities invested with the competence to authenticate the ownership and characteristics of a public key. The effectiveness of such infrastructure depends on the continuity of the chain of trust guaranteed by those certification authorities. As private sector organizations take on the role of certification authorities, what mechanisms are, or should be, in place to guarantee the continuity of the chain of trust in the event that an organization ceases to exist?

3. What specific technologies might support the implementation of the benchmark and baseline requirements in specific record-keeping and record-preservation environments?

**Record-keeping policy**

1. What are the juridical implications of developing a record-keeping system in which some requirements for authenticity are satisfied in an implicit rather than an explicit manner—for example, in a trusted record-keeping system where the integrity of the system as a whole, including the procedures used to maintain the system, creates a presumption of the authenticity of its component parts? Would such a record-keeping system have equal validity in common law and civil law jurisdictions?

2. To what extent can the models and principles developed by this project for administrative and bureaucratic records be applied to other kinds of digital objects such as records generated for cultural and creative purposes?
Appraisal Task Force Report

Task Force Members:

Terry Eastwood, University of British Columbia (Chair)
Barbara Craig, University of Toronto
Du Mei, Central Archives of China
Philip Eppard, University at Albany, SUNY
Gigliola Fioravanti, Central Direction of Archives, Italy
Normand Fortier, National Archives of Canada
Mark Giguere, National Archives and Records Administration
Ken Hannigan, National Archives of Ireland
Peter Horsman, Netherlands Institute of Archival Education and Research
Agnes Jonker, Netherlands Institute of Archival Education and Research
Leon Stout, Pennsylvania State University
Su-Shing Chen, University of Missouri, Columbia
Introduction

The Appraisal Task Force set out to determine whether the theory and methodology of appraisal for electronic records differs from that for traditional records, and what role the activities of appraisal play in the long-term preservation of electronic records. In doing so it relied heavily on and coordinated its work with the other two InterPARES Project task forces. The original InterPARES research plan asked the Appraisal Task Force to answer seven questions in its domain of investigation (Domain 2):

- What is the influence of digital technology on appraisal?
- What is the influence of retrievability, intelligibility, functionality, and research needs on appraisal?
- What are the influences of the medium and the physical form of the record on appraisal?
- When in the course of their existence should electronic records be appraised?
- Should electronic records be appraised more than once in the course of their existence, and, if so, when?
- Who should be responsible for appraising electronic records?
- What are the appraisal criteria and methods for authentic electronic records?

Although these questions animated the research, many other questions—often related but sometimes new—were to arise. The research moved through three phases. It began with a review of the literature on the appraisal of electronic records. It then examined the available documentation from archival institutions on their appraisal policies and procedures, as well as reports on actual appraisals of electronic records. The final phase, and the most important, involved developing a function model of the selection of electronic records in order to gain a detailed understanding of the nature of the activities in appraising electronic records to be preserved in authentic form. Selection encompasses both appraising and carrying out the disposition of electronic records. It was mainly during the modelling that new questions arose and were answered.

It is worthwhile emphasizing that the task force aimed primarily to identify the specific junctures in the selection process when issues of authenticity come into play. It was for this reason that it has relied heavily on the conceptual requirements for assessing authenticity developed by the Authenticity Task Force and on the preservation requirements developed by the Preservation Task Force. The activities of appraisal determine which records need long-term preservation and the various terms and conditions applying to their continuing preservation beyond the time of their active life in relation to the affairs that created them. Conceptually, this determination involves passing the responsibility for preservation from the creator to some entity assigned responsibility for long-term preservation. The most important aspect of this passing of responsibility is to ensure, as much as is possible, that the identity and integrity (as defined by the Authenticity Task Force) of the records can be established and preserved over time. The conceptual requirements for assessing authenticity are therefore an important guide in the process of identifying electronic records and establishing their integrity during appraisal. The requirements for preservation come into play in a vital way when we assess the feasibility of preserving authentic electronic records. The work of the Appraisal Task Force must therefore be viewed in the light of the work and results of the two other task forces. The main connections and consequences of these relationships are outlined in this report.

The primary outcome of the work of the Appraisal Task Force is the function model of the selection process. It and the thinking behind it represent the main contribution of the task force to the problem of long-term preservation of authentic electronic records, as this report will explain at length.

---

Basic Assumptions of the Research

Definition and situation of appraisal

In common usage, *appraisal* is defined as “the act of estimating the nature, quality, importance etc.” of something. Archival dictionaries and glossaries speak of appraisal as being “a basic archival function” aimed at determining the disposition or disposal of records, that is, usually either their continuing preservation or their destruction. In some cases, records may be alienated from their creator. Appraisal involves making a judgement or estimation of the worthiness of continued preservation of records. The Appraisal Task Force treats the term *appraisal* in this sense of coming to a determination of the disposition of records. However, the task force considered that the function at issue is broader than the matter of determining the disposition of records. It did so because appraisal is usually part of the act of acquisition of records by the entity responsible for their long-term preservation, the preserver. Most commonly, the preserver appraises records, that is, determines the worthiness of their continuing preservation, and carries out their disposition. Carrying out the disposition of records usually (but not always) involves a transfer from the creator to the preserver of the custody of those records determined to be worthy of continued preservation and/or destruction of records not deemed worthy of long-term preservation. In situations where the preserver has responsibility for appraisal, it often comes to decisions about the fate of records in consultation with the creator. Carrying out disposition is also often a shared responsibility between the creator and the preserver. Responsibility to destroy records may fall to the creator or to the preserver, or be shared by them.

The preserver is the juridical person whose primary responsibility is the long-term preservation of authentic records. The preserver may be an archival institution, such as a national, state, or provincial archives given responsibility for the long-term preservation of the records of a governmental organization. It may be an office of an organization, as is often the case with the archives division or department of such organizations as churches, businesses, and universities. It may even be an office (or officer) within the entity creating the records; this would be the case, for instance, where an agency’s archives within an organization such as a government were given responsibility for long-term preservation of electronic records. In short, there must be some entity that is assigned responsibility for preservation. Therefore, the task force proceeded from the perspective of this entity and with long-term preservation in mind.

Whatever the division of responsibilities may actually be, it is necessary to conduct appraisal to identify records worthy of continued existence, then carry out the disposition of records determined to be of long-term value, and finally set in motion arrangements for their preservation. It is this selection function, rather than the differences in the way responsibility for them is actually assigned, that interests the task force. It is assumed that the activities of appraisal and carrying out disposition, once they are understood in sufficient detail, can be conducted in numerous administrative contexts.

The other premises of the research are found in concepts, such as “electronic record” and “authenticity,” of the project at large. Where it is necessary in this report, these other concepts are discussed.

---


3 The creator is the physical or juridical person in whose archival fonds the record exists. The fonds is the whole of the records created (meaning made or received and set aside for action or reference) by a physical or juridical person in the course of carrying out its activities.
Research Design and Methodology

Literature review
In the first stage of its research, the task force conducted a review and analysis of the existing literature on appraisal of electronic records to confirm or reject the various research questions with which it began. For the most part, it confirmed the wisdom of these initial questions. The full review is reproduced in Appendix 3; however, some of the main findings are worth repeating here.

The review determined that a consensus had developed that electronic records must be appraised from the same theoretical and methodological standpoint as traditional records. In other words, important as the influence of technology is on certain aspects of methodology, writers still employed the same general concepts in their writing about electronic records, particularly in assessing the full context of electronic records and their continuing value. In the view of many writers, the main influence of the technology was in fact a negative one. Few creators made adequate provision for electronic record keeping, with the result that it was difficult to determine what an electronic record was in many cases and therefore to conduct appraisal of such records. In particular, archivists wrestled with the problem of what to do with dynamic databases, many of which were implicated in record keeping but did not actually produce records.

Many writers discussed the question of the timing of appraisal. Almost all of them concluded that appraisal had to be conducted early in the life of systems producing the records. There were accounts of both appraisal and preservation activities being hindered because records had been removed from active systems without adequate documentation of their context of creation. To avoid these problems, most writers advised early archival involvement with creators to determine records of continuing value and to develop procedures for their disposition.

The review implicitly reveals that writers rarely addressed the application of the concept of authenticity to appraisal. For the most part, they assumed that preserving electronic records in authentic form is a matter for the preservation function. Appraisal determines which records are to be preserved; it is then up to the preserver to ensure their continuing identity and integrity. Although that is true, the work of the task force explicitly shows that the concept of authenticity has important application during appraisal, as will be explained.

Review of policies and procedures
The review of policies and procedures of archival institutions and programs amplified the findings of the literature review, as might be expected. It also revealed that only a small number of institutions and programs had anything like extensive experience appraising electronic records. Those that did appraised electronic records in conjunction with and using similar methodology to that for traditional records.

The main documentation of value proved to be actual reports of appraisal of electronic records. These reports revealed that archivists expended much time and energy to appreciate the various contexts of the records, including of course the technological context; applied criteria familiar in the appraisal of traditional records; and spelled out the terms and conditions of disposition of records deemed worthy of continuing preservation. Together the reviews of the literature and policies and procedures provided a body of empirical knowledge used in the task force’s main exercise to develop a function model of selection.

Modelling the selection function
A function model represents the various activities of a functional process in a series of structured diagrams. The task force used IDEF(0) or Integration Definition for Function Modelling, which is derived from Structured Analysis and Design Technique. IDEF(0) is a U.S. Federal Information
Processing Standard, which is detailed in Publication 183 of the National Institute of Standards of Technology.\(^4\) The model and the related definitions of terms are reproduced in Appendix 4.

The purpose of the model is to characterize the relationships of the activities involved in selection of authentic electronic records for long-term preservation. The model is produced from the viewpoint of the entity responsible for the long-term preservation of electronic records of an organization. The assumption is that the same activities occur in any context where selection is performed. Arrows pointing into a box representing an activity indicate inputs. Arrows pointing outwards from boxes indicate outputs. Arrows pointing down from the top of boxes indicate constraints on the activity, and arrows pointing upwards the mechanisms necessary to accomplish the activity.

The following discussion can be read independently of the model. However, diagram numbers are given to aid those readers interested in finding the part of the model relevant at any point in the explanation of the process.

The task force employed modelling methodology in order to isolate and characterize the various activities of the selection function. The methodology requires consistency and careful definition of the concepts and terms used in the model. However, the model itself, because it depicts a highly intellectual and complex process, needs considerable explanation, which the next section of this report provides.

**Research Findings**

**Preamble**

The work on the model began in early 2000. At this time, the Authenticity Task Force had already developed the *Template for Analysis* of electronic records to be used by the project. The *Template* and, subsequently, the *Requirements for Assessing and Maintaining the Authenticity of Electronic Records* provided an important conceptual framework for the thinking behind the modelling. In several meetings of the task force over the next year and a half, the model was refined and coordinated with the work of the other two task forces.

**The Scope of the Main Activities in Selection (Diagram A0)**

From the perspective of the preserver, appraisal is obviously a vital first step in the process of preservation. Selecting electronic records involves appraising them *and* carrying out their disposition. Carrying out disposition acts as a bridge between the activities of appraisal and those of preservation. Information about electronic records amassed during their appraisal is vital to the actions taken to determine and carry out their disposition and then, later on, to the actions taken to preserve them. Nevertheless, it is important to note that responsibility for the actions of carrying out disposition will probably be shared between the creator and the preserver in most instances. Clearly the organization’s policies and procedures will have to sort out the responsibilities that fall to the creator and those that fall to the preserver as part of the disposition rules guiding transfer of records.

The point about disposition rules makes it clear that one of the activities of selection is to establish, implement, and maintain a framework for the selection function. Managing the selection function also sets the rules and conventions for the preserver that govern appraisal.

Many problems that occur in the archival treatment of electronic records stem from changes in the records’ context over time. Monitoring these changes is a distinct activity; it ensures that up-to-date information about records is compiled and that appraisal decisions are updated accordingly or, where there is a need, revisited. To a large extent, monitoring electronic records selected for preservation is our answer to the research question, “When in the course of their

\(^4\) See <http://www.idef.com/idef0.html> for more information.
existence should electronic records be appraised?" The answer is frequently dictated by the circumstances of change in the context of the records. In cases where the appraisal decision is built into design of electronic systems, such as by records scheduling, or where it is conducted sometime after a system has been in operation, monitoring records selected for preservation and making adjustments as needed is part of the process of selection. By contrast, appraising electronic records long removed from the active system in which they were generated is usually made more difficult because the relevant information about their technological and other contexts is no longer available or difficult to obtain.

Selection, therefore, encompasses four main activities: (1) managing the selection function, (2) appraising electronic records, (3) monitoring electronic records selected for preservation, and (4) carrying out the disposition of electronic records.

The Broad Picture of Selection (A0)

Selecting electronic records for long-term preservation—like selecting records in general—responds, broadly speaking, to societal needs and to the creator’s needs for continuing reference to the records. It also responds, explicitly or implicitly, to certain legal requirements, that is, to the concepts, principles, and specific statements in law relevant to the selection of records. All the activities of selection are conducted with an understanding of the theory, methodology, and practice of archival science, including the requirements for ensuring authenticity of records. Societal needs, creator’s needs, legal requirements, and archival science and authenticity requirements all condition or influence the process of selection. How they influence actions and decisions from juridical system to juridical system or for any one preserver is a question that is beyond the scope of our research. Nevertheless, it seems obvious that managing the selection function is largely a matter of taking these conditioning factors into account when developing policies, strategies, procedures, and standards.

It hardly needs saying that the successful selection of electronic records requires knowledgeable persons, certain facilities, and computer equipment and software. These are the necessary instrumentalities of selection. As mechanisms employed in carrying out selection, they are needed for all of the activities described in this report.

Broadly speaking, selecting electronic records means identifying records for transfer to the preserver for continuing preservation. From among the electronic records produced by an organization, some will be selected and transferred to the preserver and some will not. The outcome in any given case will either be a transfer of electronic records selected for preservation from the creator to the preserver or a designation of electronic records not selected for continuing preservation. It is a matter of organizational policy whether or not the preserver plays a role in the disposition of electronic records not selected for preservation. In any event, the outcome or result of selection is that electronic records both destined and not destined for continuing preservation are identified.

The work of the task force has confirmed something that is implicit, but not spelled out clearly, in the literature on appraisal of electronic records. In large measure, selection of electronic records depends upon a gathering and assessment of information concerning the context of a given body of records or deriving from the records themselves. Relevant information gathered during the process must then be associated with the records so that they can be managed effectively by the preserver and easily understood by future users. Obviously, a great deal of information about the context of electronic records exists while they are in active use, because it is needed for the continuing management of the records. This information often disappears or is difficult to assemble once records are removed from the active system in which they were generated. This situation provides a strong argument for beginning the appraisal process while records are still “live” in a system, and monitoring each phase of their existence to keep appraisal decisions relevant and disposition plans practicable.

In particular, information about the technological context of electronic records comes into play at two vital stages of selection. It is needed when assessing records’ authenticity, and when determining the feasibility of preserving authentic electronic records, as later parts of this report
spell out in greater detail. The other contextual information (juridical-administrative, provenancial, procedural, and documentary) tends to be relevant when assessing the continuing value of records, that is, judging their capacity to serve the continuing interests of society and their creator. For the most part, appraisers draw inferences about the continuing value of records from an understanding of the records and their various contexts.

Two kinds of information result from the appraisal process. There is information about the appraisal decision itself, and information about the electronic records selected for preservation and “packaged” with them as part of a transfer from the creator to the preserver. The latter is the information about electronic records necessary to maintain them continuously in authentic form, and includes the terms and conditions of transfer, to which the preserver may have to refer from time to time, such as when determining that a transfer contains the actual records designated to be transferred in a particular case.

Managing the selection function (A0, A1)

Constraints on the process
The preserver needs to establish, implement, and maintain a framework of policies and procedures guiding the selection function. The purpose of management is to make sure that the preserver’s requirements for selection of authentic electronic records are met effectively and efficiently. Managing the selection function means taking responsibility for the whole process, both the quality of its outcome and the efficiency of the process. The main responsibility is to ensure that those records of continuing value are identified and capable of being maintained according to the appropriate authenticity requirements.

Typically, the activity of managing a process transforms external requirements into internal directions. Furthermore, it generates feedback from internal processes and reacts to these feedback signals by modifying directions. Like any higher-level management process, it is not too structured, and this is also true for the knowledge and information used and processed in the function. The main process could be described as follows: collecting and evaluating requirements for, constraints to, and opportunities for potentially appropriate operational appraisal strategies. Indeed, like any management process, it is focused on matching external needs, requirements, and constraints with the possibilities of the system or processes to be managed. One other characteristic of managing is its relative autonomy in interpreting external requirements and gathering relevant information. This means that it is not possible to establish policies, strategies, procedures, criteria and standards that will fit all circumstances. Instead, we can only indicate the general considerations that go into building the managerial framework.

The external conditioning factors on the selection function are assessed during analysis of the creator’s needs for effective disposition of records, the broader societal needs for reference to records, the necessity to observe and meet authenticity requirements, the imperatives of any legal requirements bearing on the records, and, of course, the need to observe the concepts and principles of archival science. The notion of creator’s needs and societal needs is a familiar one in archival science. Schellenberg recognized creator’s needs in his concept of primary value, and

5 The Template for Analysis has defined the various contexts of an electronic record and therefore, by implication, of the various aggregations of electronic records such as are examined during appraisal. Technological context is “the hardware and software environment in which the record exists or was created.” Juridical-administrative context is “the legal and organizational system in which the creating body exists.” Provenancial context is “the creating body, its mandate, structure, and functions.” Procedural context is “the business procedure in the course of which the record is generated.” Documentary context is “the fonds to which a record belongs, and its internal structure.” Internal structure refers to the relationships among the records in a fonds.

6 The task force has defined terms and conditions of transfer as “a document that identifies in archival and technological terms electronic records to be transferred, together with relevant documentation, and that identifies the medium and format of transfer, its timing, and the parties to the transfer.”
societal needs in his concept of secondary value.\textsuperscript{7} These needs include requirements for rendering accountability in the political, administrative, fiscal, legal, and broader societal senses. All these constraining factors are considered when developing the policies, strategies, procedures, criteria, and standards guiding the selection function.

Obviously, external factors such as creator’s and societal needs and legal requirements will vary from situation to situation. They are factors that will, through policies and so on, influence the value judgements made during appraisal. The requirements dictated by the concepts and principles of archival science are another matter. In fact, the concept of authenticity and the conceptual requirements for assessing authenticity are matters of archival science. Because they are the most important concepts bearing on long-term preservation of \textit{authentic} electronic records, and to ensure that they are recognized as an important conditioning factor of the process, authenticity requirements are highlighted as a separate constraint on managing the selection function. They explain the concepts guiding practice. Even though the application of authenticity requirements in any given case may be a matter of judgement, it is not a value judgement that is at issue, as is the case in judging continuing value to the creator and society. The manner in which authenticity requirements guide the conduct of appraisal will be looked at in some considerable detail as we move through the specific activities in which they come into play.

\textbf{Information needed for management}

Essentially, four kinds of information are needed to support the development of policies, strategies, and procedures guiding selection: information about the records’ context, information about appraisal decisions, information about updated appraisal decisions, and information about disposition.

In any given case, the preserver aiming to develop a framework for selection must gather information about the context in which records for which it has responsibility are created. For instance, if the preserver is a government archives, what range of agencies is it responsible for, under what administrative arrangements, performing which functions, and so on. Information of this kind feeds directly into the process and comes out of it in statements (such as on the scope of records to which the policy applies) in the framework.

Information about appraisal decisions already made provides valuable intelligence about the success or lack of success of the process, and as such is an important input to establishing, implementing, and maintaining an effective framework. Where the process of monitoring electronic records selected for preservation results in updated appraisal decisions, information about these updated decisions and the reasons for them also provides valuable intelligence to feed into the process of managing the framework.

Much the same is the case with information about the disposition of electronic records. Experience of actual dispositions over time will reveal information useful for managing the framework, such as whether appraisal decisions are properly implemented.

\textbf{Results of managing selection}

There are two aspects of the framework. One is a set of rules and conventions governing the conduct of appraisal, which for convenience we call the \textit{appraisal strategies}. The appraisal strategies operate as controls on the processes of appraisal of electronic records and monitoring of electronic records selected for preservation. The appraisal strategies encompass:

- criteria for appraisal
- guidelines on how to apply authenticity requirements
- procedures for carrying out appraisal
- guidelines for reporting the results of appraisals

procedures for reporting on appraisal activities

The second aspect is a set of rules and procedures governing the conduct of disposition of electronic records. The rules and procedures act as a control on the activity of carrying out the disposition of records. These rules and procedures include:

- procedures for carrying out disposition (e.g., roles and responsibilities of the creator and the preserver)
- rules for disposition (e.g., acceptable formats for transfer, means of transmission of records, etc.)
- procedures for reporting about disposition activities (e.g., character and volume of records acquired and/or destroyed)

**Appraising Electronic Records (A2)**

When applied to any given body of records, selection can be broken down into two main processes. First, a decision is made with respect to the records’ disposition; then that decision is implemented, that is, the records are transferred or otherwise disposed of. Furthermore, the creation of information “packages” to document the appraisal decision and the records to be preserved is crucial to allow for the performance of other archival functions, such as preservation and description. Within the larger context of the selection of records, therefore, appraisal is the activity during which relevant information is gathered and compiled, and a disposition decision is made.

We have viewed appraisal as being made up of four distinct activities: compiling information about the records and their contexts; assessing the value of the records; determining the feasibility of preserving them; and finally, making the appraisal decision. This breakdown is based on a decomposition of appraisal into its logical component activities or functions, and it is not meant to specify a precise workflow. It does not make any assumptions about the organizational setting in which the activities take place. It does, however, assume that there is a strong chance the continuing value of records differs from the records' operational value to their creator, and that the continuing value should be determined according to a different set of criteria from that for operational value.

Appraising a body of electronic records is to decide on their disposition. If they are deemed to serve some enduring need of their creator or society, the records will be preserved. One common way of doing this is by transferring them to an entity, such as an archival institution or program responsible for the records' continuing preservation. It is also possible that the creator will preserve them indefinitely, possibly by an archival unit, possibly under the supervision of an outside archives authority. If the records are not deemed valuable, they will be destroyed or, perhaps, be alienated to the care of some other entity.

**Compiling Information**

In order to conduct an appraisal, the person or persons conducting the appraisal (the appraiser) needs information drawn from reading the form and content of the records, and information about the records' various contexts (juridical-administrative, provenancial, procedural, documentary, technological). The appraiser gathers, organizes, and records this information as a vital part of the process of determining disposition of the records. Information may come from publicly available sources, as well as be obtained from the creator's personnel or documentation, and from the records themselves. The precise nature and scope of the information required depends on the particular appraisal methodology and criteria that the preserver has implemented.

It should be stressed that during this vital activity of appraisal, inferences are accumulated about the continuing value of the records and about the grounds for presuming them to be authentic. Referring to this activity as "compiling information" may mislead. Appraisal must rest on a foundation of solid research, which will assist in performing several of the activities we have identified, particularly assessing the value and the authenticity of the records, and identifying the digital components that have to be preserved.
Assessing Value (A22)
The archivist uses the information gathered and compiled to determine the capacity of the records to serve the continuing interests of both their creator and society. The archivist answers the question: “How valuable are these records? How important is it to preserve them?” The output of that activity is an assessment of the continuing value and authenticity of the records, as well as information about the criteria that were used to make that assessment and how they were applied. This assessment may be further decomposed into three activities: assessing the continuing value of electronic records, assessing their authenticity, and determining their value.

Assessing continuing value
This first activity results in a statement of the reasons why the records should or should not be preserved, according to the criteria decided upon by the preserving institution. Because it involves values and judgement, appraisal may be performed differently according to different national or intellectual traditions, juridical systems (including legislation), value systems, and theoretical choices. Archivists engage in heated debates about appraisal criteria and methodologies, and for good reason. As an example, one could study the records themselves and determine which elements are likely to give them continuing value, for example, their usefulness for legal purposes, their value as evidence of the functioning and organization of their creator, or their potential for research. Another approach—particularly useful when there are vast amounts of records, created during complex, intertwined processes—is to start by appraising not the records themselves but the functions performed by the records creator, to determine which ones should be documented for posterity, and then to find out which records better reflect the accomplishment of these functions and their impact on society. Since our goal here is to come up with a model of the appraisal activity that applies in a number of different contexts, we deliberately omitted specifying which criteria or values, strategies, and methodologies should be employed.

Assessing authenticity (A222)
A second component in the assessment of value is an assessment of the records’ authenticity. The appraiser must establish the grounds for the presumption of authenticity. He or she must ensure that the records’ identity (e.g., parties involved, date, subject matter, and archival bond) is preserved, and must ascertain the degree to which the records’ creator has guaranteed their integrity, by making sure that the records remain intact and uncorrupted. The questions to be asked of the records at this stage correspond to the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records defined by the Authenticity Task Force of InterPARES in the Requirements for Assessing and Maintaining the Authenticity of Electronic Records. Answering the questions requires an in-depth knowledge of the records, the electronic systems they reside in or were created in, and the wider context of their creation and use. For example, such an important element of identity as the author of the record may in fact be found in the provenancial context (the owner of the system) rather than on the face of the record. Integrity may be maintained through safeguards built into an integrated electronic record-keeping system (the technological context of the records), or it could be ensured through policies, procedures, and practices in the environment of the electronic system. Examples would be physical restrictions on access, policies on access privileges, procedures for data entry and validation, as well as procedures for back-up and storage in different locations. The benchmark requirements give full details of the various factors to be analyzed when assessing authenticity.

Therefore, the first step for the archivist is to compile the evidence supporting the presumption of authenticity. That evidence must then be measured against the benchmark requirements. If such an evaluation does not produce a high presumption of authenticity, the archivist must try to verify authenticity by other means, such as comparing different versions or copies of the records, examining system audit trails, or interviewing personnel involved in the creation, use, and

---

8 See Appendix 2.
preservation of the records. The resulting assessment may affect the determination of the records' value. The information compiled at this stage is also crucial to understanding and using the records once they have been transferred to the preserver. Future users of the records must know how well founded the presumption of authenticity of the records is, and what information that presumption is based on in order to make their own assessment, long after the fact and when accumulating relevant information is likely to be difficult if not impossible.

**Determining Value**

The appraiser's assessment of the records' value reflects continuing value and authenticity. However, the impact of authenticity on the archival value of records is not straightforward, and requires some explanation.

For example, let us suppose that an objective of appraisal is to identify records documenting a process or function performed by the records' creator and deemed worthy of long-term preservation. To do so, an appraiser identifies records that will allow the preserver to maintain an accurate reflection of that process or function. If the records creator were performing a function that had a very high impact on society, and its record-keeping practices were very poor, the archivist would most probably still want to preserve evidence of the function by acquiring the appropriate records. Furthermore, if evidence of poor record-keeping practices and of possible willful or fraudulent tampering with the records comes to light during appraisal, that might make it more important to preserve them, in order to attest to that.

Assessment of how authenticity affects the value of electronic records is largely a matter of gathering and evaluating evidence of what has happened to them during the course of their existence prior to the time of appraisal. Of course, there is a prima facie case for presuming records to be authentic if their creator relies on them in the usual and ordinary course of business. Nevertheless, in cases where the records no longer reside in their original environment through, for instance, conversion or migration, it is necessary to determine whether what is being appraised is what originally existed and whether changes to the records have seriously impaired their ability to act as evidence of the activity that generated them. In cases where the chain of custody and preservation has been broken or where migration has resulted in missing records, missing parts of records, or inadequate or inaccurate documentation of changes, there may be good reason to suspect the value of the records. If the appraiser has good reason to suspect that the records no longer reflect what they were at the time of their creation and primary use, he or she may decide not to preserve them. Another case in which authenticity is important is when the value of the records resides in the accuracy of the information they contain, such as with survey or scientific data, rather than in how well they represent the process during which they were created.

Thus, the archivist must assess both the continuing value of the records to their creator and society and the authenticity of the records in order to determine their overall value, and decide how important it is to preserve them or not.

**Determining Feasibility of Preservation (A23)**

Assessing the value of the records is not enough, however. The appraiser must also determine the feasibility of preserving them as authentic records. More precisely, the appraiser (acting as an agent of the preserver in effect) must decide whether the digital components embodying the essential elements that confer identity and ensure the integrity of the records can be preserved, given current and anticipated capabilities. This determination is based on the same type of information from the records and about the records that is used to assess their value, but it also requires knowledge of the preserver's current and anticipated capability to preserve electronic records. This would include the state of preservation knowledge; hardware and software capabilities; staff expertise; and financial resources. That information is actually provided by the preservation function. The result of this determination is information about the resources and technical capability required for continued preservation of the records.
The activity of determining the feasibility of preserving authentic electronic records breaks down into three phases.

1. The appraiser determines both the record elements containing informational content and those elements that need to be preserved according to requirements for authenticity, as formulated by the Authenticity Task Force’s benchmark requirements.

2. The appraiser identifies where these crucial record elements are manifested in digital components of the electronic record that must be preserved.

3. The appraiser reconciles these preservation requirements with the preservation capabilities of the institution that is responsible for the continuing preservation of the body of records being appraised.

This feasibility determination gathers and records technical information that is necessary to accomplish preservation of the individual elements conveying both the intellectual content and the authenticity of electronic records being appraised. This information also includes the projected cost of preservation and an indication of whether or not the preserver has the capability to preserve the records in question.

The first activity in determining the feasibility of preserving a body of records being appraised is to determine which record elements are to be preserved to ensure the authenticity of the records. This activity identifies the extrinsic and intrinsic elements of form, as outlined in the benchmark requirements, and the elements of content of electronic records that need to be preserved in order to maintain their authenticity over time. Depending on the design characteristics of the system that produced the records, these elements of form and content may be observable on the face of the record or in metadata associated with the record, or implied in contextual information associated with the records’ creation. This contextual information relates to the legal and organizational system in which the creating body belongs; the mandate, structure, and functions of the creating body; the business procedure in the course of which the record is created; the fonds to which the records belongs, and the fonds’ internal structure. This internal structure comprises the relationships that link each record incrementally to the previous and subsequent ones, and that convey meaning to the records.

There is other relevant information, in addition to that concerning form and content, that aids in determining the feasibility of preserving authentic electronic records. In particular, the appraiser needs information about the technological context of records in order to understand how they were generated. This would involve gathering and analyzing information about the electronic system itself, the hardware, software, operating system, and the type of files created—for example, word-processing files, image files, and so on. This information is normally gathered as part of the process of compiling information to support the activities of selection, but it is important to note its special relevance in the essentially technical exercise of identifying records elements and digital components.

An archival institution’s rules and conventions for appraisal—that is, its appraisal strategies—affect this determination. Indeed, appraisal strategies are taken into consideration at all three stages in determining the feasibility of preservation.

The activity of determining the record elements to be preserved can be illustrated by using one of the InterPARES case studies. The Canadian Intellectual Property Office (CIPO) has a system called TECHSOURCE, which contains a variety of records produced during the patent-granting process. One very important record in the TECHSOURCE system is the patent application, a legal document that constitutes the first step, or act, in the process.

In the act of determining the feasibility of preserving authentic patent applications (or other legal records) maintained in this system, the appraiser would be particularly concerned about the rigour with which the creator met authenticity requirements in their creation. This is typically accomplished through controls embodied in the TECHSOURCE system, and in external procedural controls that were specified during the system's design. These system design requirements constitute the measures CIPO felt were needed to preserve the identity and integrity of the electronic records in the system.
These requirements are often expressed in the intrinsic and extrinsic elements of the records’ form. Preservation of these elements of form maintains the records’ authenticity over time and across technologies. In the case of a patent application, certain extrinsic and intrinsic elements—such as the application’s standard format, and the chronological date and time of receipt by CIPO—must be protected from tampering. In fact, such a requirement is stipulated in the Canadian Patent Act. Intrinsic elements relating to the identity of the record, such as the names of the persons involved and expression of the archival bond in the application number, would also be included.

The appraiser would first consider the records from a system such as TECHSOURCE and their various contexts. Then the appraiser would determine that a dispositive record (in which the record represents the act) such as a patent application would contain many of these elements that need to be preserved in order to maintain its authenticity, and therefore its trustworthiness. This is a critical component of appraisal and the result of this analysis would be a list of intrinsic and extrinsic record elements that must be preserved in order to ensure authenticity.

Once the appraiser has identified both the diplomatic elements of the record that confer authenticity and the content elements that need to be preserved, the next activity is to identify how these elements are manifested electronically as digital components. In the analog realm, the extrinsic and intrinsic elements are typically united on the medium; however, this is not so with electronic records. The identified elements in the electronic realm may be manifested in various ways, in what the project calls digital components. As defined by the InterPARES Preservation Task Force, a digital component is “a digital object that contains all or part of the content of an electronic record, and/or metadata necessary to order, structure, or manifest the content, and that requires specific methods for preservation of one or more electronic records, and that has specific methods of preservation and reproduction.” The concept of digital object has its roots in the object-oriented paradigm, whereby the characteristic of discernment of such an object is that it has one or more particular methods associated with it, such as presentation software. This identification of digital components is made using previously ascertained information about the record elements to be preserved, along with information already gathered about the record’s technological context.

For instance, in CIPO’s TECHSOURCE system, standard correspondence generated in the course of the patent-granting procedure is produced by combining standard templates containing formulaic language with attribute information from various tables in the relational database management system. Each of these, the templates and the attribute information, are separate digital objects, or components. In the case of the former, word-processing software is necessary to invoke the template object and in the case of the latter, database software is necessary to understanding the table and relationships represented by a particular instance of an attribute. The template digital objects contain extrinsic elements of form conferring authenticity (e.g., English and French versions of the formulaic correspondence language). The digital objects representing the table-derived attribute information contain security privilege and work-flow information that guarantee intrinsic elements of form by ensuring that only someone with proper authority can issue a particular type of correspondence. In order for the archives to preserve authentic electronic records over time, the appraiser must be aware of what these components are, what records elements are contained therein, and the means by which the elements can be united to reproduce the record in a comprehensible form.

The final stage of determining the feasibility of preserving authentic electronic records is a reconciliation of the record components’ preservation requirements with the archives’ preservation capabilities. The question is: “Can the components that manifest the informational and authenticity elements be preserved, in light of current and/or anticipated future capabilities of the archives?” Simply put: “Can the preserver preserve these digital components?” This is answered by knowing the preserver’s current and anticipated capability to preserve electronic records. This information includes the state of preservation knowledge and the institution’s hardware/software capabilities, as well as practical matters of staff expertise and financial resources available for preservation services.
The attempt to reconcile preservation requirements with preservation capabilities produces two bodies of information that inform the appraisal decision. The first body of information concerns the digital components to be preserved; it includes information that would explain where records elements that are vital for maintaining authenticity are manifested in the (potentially various) components of the electronic records, as well as the technical information (e.g., invocation methods) about those components that would be required for subsequent preservation activities. To use the TECHSOURCE example, the first type of information would include the identification of specific tables within the records and database management system (RDBMS) that correspond to specific elements of form conferring both content and authenticity. The second body of information would indicate, for example, what type of viewer software would be needed to view the system’s scanned images or what information in the image file headers could be exploited for retrieval purposes. The feasibility of preserving a given body of authentic electronic records would be based on current or anticipated finances and technical capabilities. Equipped with this information as well as the valuation information articulated in the value assessment activity, the ultimate appraisal decision and documentation supporting it is then made in light of the preserver’s appraisal strategies.

Making the Appraisal Decision
If the assessment of value determines that records are not worthy of long-term preservation, the appraisal decision becomes easy. In all other cases, however, the determinations of value and feasibility come together in determining the appraisal decision. We could describe this as balancing what the appraiser would like to preserve against what the preserver is capable of preserving and can afford to preserve. However, that would be simplistic. The balance between value and feasibility rests on an exercise of judgement, on a case-by-case basis. For example, an appraiser could be confronted with a situation where preserving records would be either extremely difficult for technical reasons, or would entail considerable costs. But this does not necessarily tip the decision against preserving the records. If the records were of extraordinary importance or their preservation were mandated by law, the archivist might look for either alternative sources of funding or another preserver, or come to an arrangement by which the creator would preserve the records—at least for a certain period of time. Nevertheless, preservation capabilities do come into play, because resources are not infinite, and choosing to preserve any given body of records often affects decisions made about other records.

The outcome of this decision making is of course an appraisal decision, which sets out the disposition of the records. The decision is made up of two parts. First, it must list what must be transferred to the preserver, or disposed of in other ways (destroyed, transferred to an entity other than the preserver, etc.). The list is laid out at a level of detail appropriate for someone to be able to carry out the disposition. Depending on the type of electronic records, and the precision of the records management system, this could mean a high-level description of records (e.g., based on their functional context or their classification), a list of record elements, or a detailed list of digital components. Ultimately, however, persons effectively carrying out disposition need precise instructions and a list of digital components.

In addition to the list of records and digital components, persons responsible for carrying out the disposition of records must be provided with information specifying how and when disposition must be effected. That includes the responsibilities of each party, and interim measures such as a monitoring schedule. The terms and conditions of transfer specify the conditions of the disposition the records, as well as more general clauses that apply to all records (such as rules about the frequency of monitoring). These general clauses are established as part of the management function in the form of disposition rules and procedures.

If all or some of the records appraised must be preserved, the content of the appraisal decision, as well as any further relevant information about the records’ technological environment, must be included in an information package for people responsible for continuing preservation.

Finally, the appraisal process must produce documentation explaining and justifying the appraisal decision. It characterizes the various contexts of the records that were relevant to the decision,
explains the methodology and criteria used, details the research method, presents the assessments of value and of feasibility, and outlines the rationale for the decision. It should make clear which records were preserved, and which were not, out of the universe of records created. This documentation is vital for accountability purposes, on the one hand, and so that future users of the records understand the records, on the other. In fact, it constitutes permanent records of the preserver that must be accessible to anyone wanting information about the disposition of records. Information about appraisal decisions is also a crucial feedback mechanism for those managing the selection function (especially in devising appraisal strategies and methodologies), and for other archivists engaged in appraisal.

Recommendations on Appraisal

1. Appraisal is a knowledge- and research-intensive activity. Appraisers must be provided with the proper training, tools, information, support, and resources to conduct the necessary research.

2. Accurate and thorough documentation of the appraisal process in its various phases and outcomes is essential. Information about the appraisal decision, as well as about the appraised records themselves, should be considered as an outcome of appraisal in its own right, as much as the appraisal decision itself. That information is required for further archival functions—such as preservation, arrangement, and description—to be performed adequately.

3. The preserver should develop an interview protocol (along the lines of the InterPARES Case Study Interview Protocol or CSIP) to ensure that the relevant information is compiled to determine the records elements that need to be preserved.9

4. The preserver should use the Requirements for Assessing and Maintaining the Authenticity of Electronic Records as the conceptual basis for its assessment of the grounds for presuming records to be authentic and for its identification of records elements that need to be preserved to ensure authenticity.

Monitoring Appraised Electronic Records (A0, A3)

Relationship to Other Activities

Monitoring appraised electronic records and the activities associated with doing so are necessary to ensure the continuing preservation of the appraised authentic electronic records. This activity occurs conceptually after an appraisal decision is made and before disposition is undertaken. This placement recognizes that any decision is fixed in time, place, and circumstance. Appraisal decisions need to be revised as required, to ensure that the information about the appraised electronic records is still valid, that changes to the records and their context have not adversely affected their identity or integrity, and that the details of the process of carrying out disposition are still workable and applicable to the records.

Logically, the appraisal decision should be monitored to ensure that time and its changes are attended to when disposition actually takes place. Disposition may be immediate upon reaching an appraisal decision, but it might not take place for some time. The appraiser, acting as a monitor of electronic records earmarked for continued preservation, fulfils two important functions. The first is to see that the appraisal decision, the detailed information about the appraised electronic records, and the terms and conditions for transfer required by the preserver reflect contemporary realities. Many changes to the records and their context will require relatively minor revisions to appraisal documentation and to the terms and conditions of transfer. However, in cases where the business processes and related computer systems are significantly revamped or rebuilt, it will obviously be necessary to consider initiating a disposition under the terms of the original appraisal and—for the two will likely go together—redoing the appraisal to take into account the radically altered situation appraisal. This kind of “redoing” of appraisal should be distinguished sharply from reappraisal in the sense of second-guessing the valuation of the

---

9 The CSIP is available on the InterPARES Web site at <http://www.interpares.org/reports.htm>.
original appraisal. It is redoing in the sense that one has to begin afresh to appraise what is in fact a new situation of records of a particular creator, the former manifestation of which has in fact ended. Describing this redoing as "a new appraisal" would not be inaccurate. However, the need to do so is very likely to go unnoticed, with perhaps disastrous consequences, unless monitoring is part of the selection process.

The Framework for Monitoring
Monitoring takes place within a framework established by specific appraisal strategies and acts upon the appraisal decision in the light of the circumstances of the records and their contexts. The appraiser is more or less constrained by the degree to which the appraisal decision is, for working purposes, embodied in statements of terms and conditions and other information about the appraised electronic records. Among other things, this information should identify the record(s) that is(are) selected for preservation; provide technical information about the electronic system and the digital and record components in that system; specify a schedule for copying, transfer, or other type of process that allow an authorized disposition to take place; and confirm these actions by an appropriate attestation from the authority with the competence to dispose of records officially.

Monitoring Tasks
One of the tasks of monitoring is to see that scheduled dispositions are carried out. The appraiser doing the monitoring keeps up-to-date information about the appraisal decision, the appraised records, and the terms and conditions of transfer so that, when it comes time to make a disposition, there are no unforeseen problems or difficulties. It is especially important to keep up-to-date information on how records are manifested in the system; how to destroy those records that do not have to be preserved; and how to acquire, copy, format, and otherwise prepare and package records for continuing preservation.

Effective monitoring ensures that the appraisal decisions and the information about the appraised electronic records meet the needs of carrying out disposition and preservation. Adjustments or minor change to the electronic records—either at the level of the record-keeping system, or in the broader contexts of document, provenance, or technology—may have a direct bearing on the implementation of the initial disposition of electronic records. Such alterations or adjustments in the course of the ordinary business of the creator also may have implications for subsequent dispositions. Monitoring, therefore, regularly confirms that the decision and its related terms and conditions can be and are implemented. Updates to the appraisal decision and associated information about the appraised electronic records will constantly be available to the persons carrying out disposition and preservation activities.

Major alterations to the records, or significant changes in the system, its platform, and/or the context of its records might alter the circumstances of the records sufficiently that the original assessment of value and determination of feasibility are no longer sound grounds on which to continue the selection of electronic records for continuing preservation. Another way of looking at it is that monitoring will result in a recommendation to initiate a transfer and redo the appraisal only when major changes take place that cannot be addressed by adjustments to the terms and conditions in the original or updated appraisal decision.

Relationship to Real Situations
Although the previous discussion has been explicit about the conceptual logic connecting activities of the monitor to appraisal and the carrying out of disposition, the model also implicitly suggests related practical steps for the real tasks of selecting electronic records for continuing preservation. The model does not prescribe specific policies, procedures, strategies, rules, conventions, or criteria, nor does it describe the detailed contents of any of these. However, the relationship of parts, the categories of control, the nature of facilities and resources, and the types of inputs and outputs required for action suggest what specific information the selector needs to achieve the continuing preservation of electronic records for reference with their qualities of authenticity identified, maintained, and preserved intact.
Monitoring is essentially a matter of the entity responsible for preservation keeping in close contact with the creator to keep track of changes to its records system. The realization of duties, roles, and responsibilities is a practical matter for organizations to decide. A large national archives, for example, with major responsibilities for the preservation of records for cultural ends, and perhaps with defined roles in the overall management of recorded information resources, will have a very different arrangement of responsibilities compared with that of a multinational private corporation that views continuing preservation as an aspect of its business needs alone. Whatever the goal of continuing preservation, monitoring decisions to ensure that they are implemented is a key activity for any organization.

Those who monitor electronic records must have access to all the components of the appraisal decision. These include information about the electronic records appraised for preservation, other related information of a contextual nature, and all details of the preserver’s system, platform, and capabilities. In addition, documentation must be revised to ensure that changes made to the initial terms and conditions as well as recommendations to redo the appraisal are understandable, accessible, and preserved. A fully developed monitoring activity will integrate the continuing and natural changes to records and the systems in which they are kept with the preserver’s need to know of changes. Monitoring ensures that the operation of the selection function meets the needs of both the creator and the preserver over the long term.

**Recommendations on Monitoring**

1. The preserver should set guidelines for the roles and responsibilities of monitoring appraised electronic records and develop work flows to ensure smooth operation of this activity.

2. The appraisal strategy and disposition rules should take account of the needs of monitoring appraised records.

**Carrying out Disposition of Electronic Records (A4)**

To carry out the disposition of electronic records is to effect disposition of the electronic records according to the appraisal decision. This activity is made up of three distinct activities: prepare electronic records for disposition; prepare electronic records for transfer to the responsibility of the preserver; and transmit electronic records to the preserver. The appraisal decision, and the terms and conditions under which there is this transfer of responsibility for preservation from the creator to the preserver, set out what the disposition is to be, as well as who will do what, and when. The disposition rules and procedures developed in the management of the selection function govern the process of disposition of records selected for preservation and records not selected for preservation. Carrying out disposition is a three-step process.

**Preparing Electronic Records for Disposition**

The first step, preparing electronic records for disposition, includes copying and, if necessary, formatting those selected for preservation so as to prepare them physically for transfer; and/or, if this falls to the responsibility of the preserver, taking those records not selected for preservation and preparing them for destruction, alienation to another entity, or such other disposition as has been determined in the appraisal decision. The inputs to this activity are the electronic records themselves and updated information about the appraised electronic records. This includes the information necessary for disposition and continuing preservation of electronic records, including the terms and conditions of transfer. Updated information will result from the monitoring activity that keeps track of the changes to electronic records in the period since the appraisal decision was first made. During this step in the process, records eligible for disposition are identified and prepared for transfer or destruction.

Therefore, two of the outputs of this activity are electronic records selected for preservation and electronic records not selected for preservation. Electronic records not selected for preservation are identified for destruction or disposition to an entity other than the one responsible for continuing preservation. Those electronic records selected for preservation are copied and formatted for transfer to the entity responsible for continuing preservation. A third output of this
activity is information about disposition. This is information about the quantity and quality of records selected for preservation and records not selected for preservation, and about the cost of disposition of electronic records, utilized for management purposes. As an output of the appraisal process, this information accompanies the decision made: transfer of electronic records, or destruction (or otherwise).

Preparing Electronic Records for Transfer
The next step, one that either the creator or preserver may take or that they may take together, is to associate records selected for preservation with the necessary information for their continuing preservation, including the terms and conditions of transfer, identification of the digital components to be preserved, and associated archival and technical documentation needed for their treatment. The relevant information should have already been compiled and recorded during the various stages of appraisal and monitoring. The task at this stage is to extract the information necessary for continuing preservation of the records from the mass of appraisal documentation, and associate it with the records.

There are two outputs from this activity: the electronic records themselves, prepared for transfer, and information about the electronic records prepared for transfer. Electronic records are copied and, if necessary, formatted for transfer, and associated with the information necessary for transmittal and continuing preservation. Information about the electronic records prepared for transfer spells out the terms and conditions of transfer of electronic records, and identifies the digital components to be preserved together with the archival and technical specifications necessary to guide continuing preservation.

Transmitting electronic records
The third step, the final activity of carrying out disposition, is to transmit the records selected for preservation—along with the accompanying information necessary for continuing preservation—to the office responsible for the preservation function.

The outputs of this activity include information about transferred electronic records and the transfer of electronic records selected for preservation. Information about transferred electronic records constitutes the record or records providing the information about electronic records necessary to maintain them continuously in authentic form, including the terms and conditions of transfer. The second outcome of the transmit function is the actual transfer of electronic records selected for preservation. These electronic records are copied and, if necessary, formatted for transfer and sent to the office responsible for the preservation function.

Recommendations on disposition
1. As part of its disposition rules, the preserver should work out a standard protocol setting out the roles and responsibilities of the creator and the preserver in carrying out disposition of electronic records.

2. The preserver should develop a standard format for recording the information necessary for continuing preservation that is associated with transfers of electronic records.
Relationship of Findings to Other Research

Although the work to develop a picture of the process of selection of authentic electronic records builds on the general literature on archival appraisal, and on the specific literature on appraisal of electronic records as reflected in the literature review, no other research we know of has delved deeply into the questions we set out to answer. We believe that the picture of appraisal of electronic records afforded by the model we have developed provides the most extensive and detailed account of the process of selection currently available.

The most important relationships of the work of the Appraisal Task Force is of course to the work of the Authenticity Task Force and Preservation Task Force of InterPARES, the main lines of which we have indicated. In this regard, it is worth reiterating that our work depends to a very great extent on the work of the Appraisal Task Force, in particular on its template for analyzing electronic records and its conceptual requirements for assessing authenticity. Readers of this report are encouraged to follow the explanation of the template and conceptual requirements in the “Authenticity Task Force Report.”

Conclusion

The task force set out to determine whether the theory and methodology of appraisal for electronic records differs from that for traditional records, and what role the activities of appraisal play in the long-term preservation of electronic records. To summarize the ways in which appraisal and disposition differ for electronic records and traditional records, it will be instructive to look again at the original research questions in light of the findings, particularly in light of the knowledge encapsulated in the model of the selection function.

1. What is the influence of digital technology on appraisal?

Some of the influences of the digital environment on appraisal simply heighten tendencies already evident in the traditional environment, while others are new.

The need to monitor or keep track of changes in the record-keeping environment is not unknown in the traditional environment. The functions and activities of creators, and their internal organization and procedures—including documentary procedures—all change over time, with the result that appraisal decisions must be revisited and amended to take account of these changes. By contrast, in the digital environment, changes in the system generating the records can present at least three scenarios.

In the first scenario, relatively minor changes to a system may lead to a relatively inconsequential revision to an appraisal and information about appraised electronic records. That is, one can live with the main lines of the original appraisal and determination of disposition. In the second, significant changes in the technological context may require one to adjust the appraisal to take account, for instance, of new work processes and their automation or technological advances. In the third, drastic changes, such as introduction of a completely new system, may lead the appraiser to initiate a disposition under terms of the existing appraisal and then, of course, a “redo” of the appraisal of records in the new system when it is determined one should be made. Because much data in the digital environment is dependent for its meaning on an understanding of that environment, deciding the disposition of records in systems about to become outmoded is likely to be an important tactic. In the traditional environment, records committed to paper did not so easily lose important aspects of the original context of creation, even if they migrated into a new record-keeping environment. Nevertheless, it is probably fair to say that monitoring change and determining its effects on selection decisions is nothing new. The need for it is just more pressing in the digital environment.
Something along similar lines can be said about the need to appraise records early in their life, when the appraiser can see a fully operational live system. In fact, modern records schedules, which in effect constitute a series of disposition decisions class by class, are often created before records are created. The difficulty in the digital environment, one discussed widely in the literature reviewed by the task force (see Appendix 3), is that designers of digital systems, particularly in the early years of office automation, paid little or no attention to questions of the disposition of records. It was this fact, rather than any inherent characteristic of the digital environment, that pushed archivists to suppose that appraisal capability had to be built into the design of systems.

The need to appraise early in the digital environment is, by contrast, vital for quite another reason. Information about the technological context, much of it now contained in the systems themselves, cannot be found or reconstructed, we know from sad experience, even a short time after systems have reached the end of their life. It is exceedingly difficult to assess the authenticity of such records, determine the feasibility of preserving them, and understand them in the future, without this information about the technological context. Once again, archivists are familiar with the difficulties of having to construe the context of the records with little else to speak of it but the records themselves. This is hardly an argument for expecting the acuity of Jean Mabillon (the Benedictine Monk who laid out the concepts and tenets of diplomatics in the seventeenth century) in all future users of electronic records where information about their technological context is concerned. Rather, considerable information about the technological context of the records needs to accompany them through time in order for the records to be intelligible in anything like an acceptable fashion in years and centuries to come. It is a principal task of appraisal to gather this information so that it can be associated with the records.

These somewhat shaded and not entirely novel influences of the digital environment are quite different in kind from the influences on two aspects of appraisal of electronic records: assessing their authenticity and determining the feasibility of their preservation. Archivists have rarely assessed authenticity overtly and as a matter of gathering evidence to support a presumption of authenticity, such as we recommend be done during appraisal using the benchmark requirements. Those requirements spell out evidence derived in large measure from analysis of the technological context (of the kind spoken about in the previous paragraph) but it is with the end of gauging the play of authenticity in the overall determining of value of electronic records. It is precisely because the digital environment is so frail that this needs to be done. Even less often have archivists gone the extra length, during appraisal, to verify the authenticity of records. Both assessment and verification along the lines recommended in the Requirements for Assessing and Maintaining the Authenticity of Electronic Records are likely to become, and should become, the rule rather than the exception.

Still, the pièce de résistance of this particular recounting of influences may be found in the activity of determining feasibility. It is here, particularly in determining how record elements are manifested in the digital environment and in identifying digital components to be preserved that the appraiser must be immersed in the technical details of the digital environment. In some cases, it may be surmised that, for reasons of the character of the digital environment, it will be determined that records cannot be preserved or not in authentic form. Could there be a greater influence?

2. What is the influence of retrievability, intelligibility, functionality, and research needs on appraisal?

The task force did not investigate these questions directly. It is implicitly clear that part of the exercise of determining the feasibility of preserving authentic electronic records is to ensure that the preserver has the capability to read and retrieve or present them in a form that does not compromise their identity or integrity. We have little to say beyond what is implicit in the final report of the Authenticity Task Force about the question of functionality. Some researchers have suggested that proper preservation of electronic records means perpetuating the functionality of the system creating the records. We have not worked upon the assumption that this is necessary if the message the record was meant to communicate is preserved and its identity and integrity
are evident. However, the means by which records are presented to researchers is really dependent on preservation capabilities, not on appraisal as such. In any given case, though, should the capability exist to replicate aspects of the functionality of the originating system, the appraiser would naturally take that into account.

3. **What are the influences of the medium and the physical form of the record on appraisal?**

As the Authenticity Task Force has determined that the medium is in fact part of the technological context and that not all aspects of physical form necessarily need to be reproduced in order to have authentic electronic records, these questions are no longer apposite. They have been proven to be the wrong questions to ask.

4. **When in the course of their existence should electronic records be appraised? Should electronic records be appraised more than once in the course of their existence, and, if so, when?**

These two questions are addressed together because they both concern the timing of appraisal. In fact, it is not possible to answer the first with a single answer. For good reason, the task force has assumed that records must exist before they can be appraised. It is indeed possible to build records retentions scheduling into the design of electronic record-keeping systems, but until records are actually created in the system and can be examined, questions around their authenticity and the feasibility of preserving them cannot adequately be answered. Of course, it might be the case that scheduling is regarded as the first step in the appraisal process, when continuing value alone is judged. This initial step would then be followed by assessment of authenticity and determination of feasibility, most likely at the time that a transfer of records to the preserver is anticipated.

The ideal scenario as we see it is that an initial appraisal is made, preferably when records can be seen “live” in the system that generated them, and the applicability of that appraisal is regularly monitored to take into account changes in the records and their contexts, with the last monitoring being at or near the time of transfer (disposition). So, yes, electronic records must be appraised more than once in the sense that the dynamic nature of the digital environment means the assumptions and judgements of the appraisal as it exists at any point in time must be validated before disposition action is taken. In short, the idea of monitoring is the answer to questions about the timing of appraisal of electronic records that have been raised in the literature. For a full discussion of opinion on this question, see the literature review in Appendix 3.

5. **Who should be responsible for appraising electronic records?**

As already discussed, one of the basic assumptions of the research is that appraisal is part of the primary responsibilities of the preserver, although obviously there are nuances, as already discussed, to the way in which responsibility may actually play out in a given administrative setting. We believe that our work buttresses this assumption in several ways.

The preserver has the long term and the interests of entities other than the current creating body in mind when appraising records. The assessment of authenticity and the documentation generated and preserved during that assessment are actions associated with the preservation function. This assessment and its documentation are preserved for the benefit of future users (whether inside or outside the creating body) wishing to establish the grounds for the presumption of authenticity of the records.

Another argument follows the logic of determining the feasibility of preserving electronic records. If appraisal is not undertaken by the preserver and with the current and expected future capabilities of the preserver in mind, there is the chance that appraisal and preservation will simply be disconnected. It is hard to imagine that the preserver should accept decisions about
what it must preserve without having responsibility for making the decision. The methodology of appraisal implied in the model leaves ample room for the interests of the organization creating the records to be taken into account.

Finally, we hope that the complexity of appraising electronic records such as we have indicated makes it abundantly clear that appraisal requires considerable professional expertise to perform. Because of this, it seems unreasonable to expect that anyone other than persons devoted to the primary task of preservation should be saddled with the responsibility to appraise electronic records.

6. What are the appraisal criteria and methods for authentic electronic records?

The methodology explicit and implicit in the model is our answer to the aspect of this question about methods. The only criteria that, in our view, can be established to cover all situations are, first, the requirements for assessing authenticity as part of assessing the value of electronic records; and, second, the concepts that have been developed for determining the record elements to be preserved and for identifying the digital components to be preserved as part of determining the feasibility of preservation. As we have explained, we did not regard it as part of our charge to establish criteria governing assessment of continuing value, because assessing continuing value is so sensitive to the entire context in which appraisals are made.

We are also of the view that we cannot go beyond the conceptual requirements developed by the Authenticity Task Force for assessing authenticity. They in fact provide sufficient criteria for assessing authenticity and for determining the records elements that are vital for the identity of electronic records. Although we did not take it as part of our charge to develop criteria guiding the determination of continuing value, we do recognize that the appraisal strategies of preservers should include criteria to apply in assessing continuing value.

Nevertheless, the model of the activities of selection shows that appraisal is a vital first step in long-term preservation of authentic electronic records in innumerable ways. It gathers and synthesizes essential information and evidence to ensure the authenticity of electronic records and to set in motion their disposition and long-term preservation.

Summary of Recommendations

1. Appraisal is a knowledge- and research-intensive activity. Appraisers must be provided with the proper training, tools, information, support, and resources to conduct the necessary research.

2. Accurate and thorough documentation of the appraisal process in its various phases and outcomes is essential. Information about the appraisal decision, as well as about the appraised records themselves, should be considered as an outcome of appraisal in its own right, as much as the appraisal decision itself. That information is required for further archival functions—such as preservation, arrangement, and description—to be performed adequately.

3. The preserver should develop an interview protocol (along the lines of the InterPARES Case Study Interview Protocol or CSIP) to ensure that the relevant information is compiled to determine the records elements that need to be preserved.

4. The preserver should use the Requirements for Assessing and Maintaining the Authenticity of Electronic Records as the conceptual basis for its assessment of the grounds for presuming records to be authentic and for its identification of records elements that need to be preserved to ensure authenticity.

5. The preserver should move set guidelines for the roles and responsibilities of monitoring appraised electronic records and develop work flows to ensure smooth operation of this activity.

6. The appraisal strategy and disposition rules should take account of the needs of monitoring appraised records.
7. As part of its disposition rules, the preserver should work out a standard protocol setting out the roles and responsibilities of the creator and the preserver in carrying out disposition of electronic records.

8. The preserver should develop a standard format for recording the information necessary for continuing preservation that is packaged with transfers of electronic records.
Preservation Task Force Report

Task Force Members

Kenneth Thibodeau, U.S. National Archives and Records Administration (Chair)
Richard Blake, Public Records Office, U.K.
Paola Caruci, National Archives of Italy
Robert Chadduck, U.S. National Archives and Records Administration
Michele Cloonan, University of California, Los Angeles
Babak Hamidzadeh, University of British Columbia
P.C. Hariharan, Johns Hopkins University
Ross Harvey, Curtin University of Technology
Hans Hofman, National Archives of The Netherlands
Torbjörn Hörnfeldt, National Archives of Sweden
Richard Lysakowski, CENSA
Reagan Moore, San Diego Supercomputer Center
Christine Petillat, National Archives of France
William Rhind, CENSA
William Underwood, Georgia Tech Research Institute
Bruce Walton, National Archives of Canada
Introduction

InterPARES Domain 3, Methodologies for Preserving Authentic Electronic Records, focused on the topic of preservation, and the InterPARES Project established the Preservation Task Force to address questions in this area of investigation. The original InterPARES research plan stated, “The goal of the research in this domain is to identify and develop the procedures and resources required for the implementation of the conceptual requirements and the criteria identified in the first two domains.”

The research plan articulated the following research questions for Domain 3:

- What methods, procedures, and rules of long-term preservation are in use or being developed?
  - a) Which of these meet the conceptual requirements for authenticity identified in Domain 1?
  - b) Which methods of long-term preservation need to be developed?
  - c) Which of these methods are required or subject to standards, regulations, and guidelines in specific industry or institutional settings?

- What are the procedural methods of authentication for preserved electronic records?
  - a) In what way can archival description be a method of authentication for electronic records?
  - b) In what way can appraisal and acquisition/accession reports be constructed to allow for the authentication of electronic records?
  - c) What are the procedures for certifying electronic records when they cross technical boundaries (e.g., refreshing, copying, migrating) to preserve their authenticity?

- What are the technical methods of authentication for preserved electronic records?

- What are the principles and criteria for media and storage management that are required for the preservation of authentic electronic records?

- What are the responsibilities for the long-term preservation of authentic electronic records?

Although the Preservation Task Force was charged with addressing these questions, several of the questions presume knowledge that would only be developed by other InterPARES groups. All five questions relate to authenticity and authentication, and their answers depend specifically on the articulation of these concepts and related requirements by the Authenticity Task Force. Part (b) of the second question depends in part on the work of the Appraisal Task Force. However, the Preservation Task Force could not delay starting its work until the products of the other task forces were finished, or it would not have been able to address most of these questions within the time frame of the project.

The task force therefore proceeded to address the issues of concern in the original questions, rather than literally to answer the questions as originally formulated. Pending results of the Authenticity and Appraisal Task Forces, the Preservation Task Force proceeded along lines that are essentially independent of those products. On the one hand it gathered empirical data about existing programs, plans, and technologies for preserving electronic records. On the other, it undertook a structured analysis of the process of preserving electronic records. In the empirical domain, the task force conducted a survey of programs that are preserving, are planning to preserve, or are conducting research related to the preservation of electronic records, and developed a white paper on media for storage of digital information. It also collected information on methods of authentication being used in organizations that are participating in InterPARES and preserving electronic records; however, in the absence of criteria for authentication, these data could not be analyzed.

---

In the analytic realm, the task force addressed the preservation of authentic electronic records by recasting the research questions. The key question addressed was:

- What activities are necessary to preserve electronic records?
  - a) What are the inputs to this process?
  - b) What controls govern the process?
  - c) What does the process produce; that is, what are its outputs?
  - d) What resources or mechanisms are necessary to carry out the process of preserving electronic records?

In addressing the above question and its sub-questions, the task force explored related issues:

- How do requirements for authenticity impact the preservation process?
  - a) How is compliance with requirements for authenticity demonstrated?
  - b) How can technological methods be evaluated in light of requirements for authenticity?

- How does the appraisal of electronic records impact on the preservation process?
  - a) How does preservation impact appraisal?

The survey also addressed the original Domain 3 opening questions and asked three additional questions as a result of initial responses:

- What is the meaning of preservation?
  - a) Does the meaning change when it is applied to electronic rather than paper-based records?

- Will current strategies for preserving electronic records ensure longevity and authenticity?

- How are costs for the preservation of electronic records derived? Have effective cost models been developed?

Although the Authenticity Task Force articulated requirements, they were not received by the Preservation Task Force in time to incorporate them fully in its products. Nonetheless, this task force has produced, in the formal model it developed of the function of preserving electronic records, a framework in which the requirements for authenticity and authentication can be applied. In essence, this model is neutral with respect to these requirements; that is, the model includes "place-holders" where the requirements could be introduced. In an initial review of the requirements for authenticity articulated by the Authenticity Task Force, the Preservation Task Force determined that no substantial revision of the model is needed to accommodate these requirements.

The Preservation Task Force review of the model of the process of selecting electronic records for preservation, produced by the Appraisal Task Force, showed that the two models are not incompatible, but some adjustments and clarifications are needed to align them so that they can readily be used together.

Research Design and Methodology

Survey Design

A survey, rather than a case study, method was selected because it was too early in the development of long-term retention strategies of most of the target respondents to study individual programs in depth. Given a target group of respondents consisting of fifteen sites known to be developing one or more of the techniques for digital preservation, the survey did not warrant a quantitative research design. Rather, the survey adopted a purposive sampling strategy—one that would show different perspectives on the problems we wanted to address—of archives, projects, and programs in the United States, Canada, and Europe. The investigators
developed a questionnaire with feedback from other members of the task force. The broad, sometimes open-ended questions further justified a qualitative design.

The target institutions were sent a consent letter explaining that, if they volunteered to participate in this study, they agreed to read the survey instrument, which was attached to the letter, and participate in an interview based on this instrument. Representatives from thirteen of the selected fifteen sites were ultimately interviewed. Some interviews were conducted in person, others by telephone.

**Modelling Method**

The Preservation Task Force developed a functional model of the process of preserving authentic electronic records following the Integrated Definition (IDEF) method prescribed by the InterPARES International Team. Specifically, it used IDEF(0) to describe processes or functions involved in preserving electronic records. In IDEF(0), “A function model is a structured representation of the functions, activities or processes within the modeled system or subject area.” An IDEF(0) model includes activities and entities. An activity is depicted as a box whose name indicates the nature of the activity. An entity either goes into or comes out of a process (activity). Three types of entities go into a process: inputs (I) that are transformed or consumed in the process, controls (C) that govern its execution, and the mechanisms (M) needed to carry it out. Only one type of entity comes out of a process: the outputs (O) that are produced by acting on the inputs under conditions and constraints imposed by the controls. In IDEF(0) diagrams, the four types of entities are always depicted as arrows in the following arrangement: Inputs enter a process box at the left side. Controls enter at the top. Outputs exit from the right, and mechanisms enter at the bottom. Given this invariant order, the entity arrows are collectively referred to as ICOMs.

In IDEF(0) diagrams, there are two basic icons: boxes are used to represent activities or processes and arrows represent ICOMs. In IDEF(0), a process may be decomposed into its subprocesses. This is depicted by creating a new, child diagram in which the parent process box becomes the outer boundary of the diagram and the subprocesses are depicted as boxes within that diagram. All ICOMs connected to a box at a higher level are shown entering or exiting at the corresponding edge of the decomposition diagram. Successive decompositions can be delineated to achieve whatever level of precision or clarity is desired. Such successive decompositions constitute a decomposition hierarchy. All IDEF(0) models start at the highest level, labelled “A0,” showing only one process box, which is the function being described taken as a whole, and the ICOMs that enter the function from the outside and that are output from the function. This simple notation provides a systematic and highly coherent method for describing a process to whatever degree of granularity is needed.

The boundaries of the preservation model derive from the viewpoint according to which the model is constructed. The IDEF(0) models “functions (actions, processes, operations), functional relationships, and the data and objects,” according to the National Institute of Standards and Technology. The relationships between functions are logical, and not necessarily chronological. IDEF(0) does not explicitly model temporal sequences. Moreover, in IDEF(0),

The viewpoint determines what can be “seen” within the model context, and from what perspective or “slant”. Depending on the audience, different statements of viewpoint may be adopted that emphasize different aspects of the subject. Things that are

---

2 National Institute of Standards and Technology (NIST), INTEGRATION DEFINITION FOR FUNCTION MODELING (IDEF0), Draft Federal Information Processing Standards Publication 183, 21 December 1993.

important in one viewpoint may not even appear in a model presented from another viewpoint of the same subject.\(^4\)

The horizon for the viewpoint of the preservation model is determined by the scope of the InterPARES Project as whole. The project is concerned with the preservation of electronic records that have been selected for preservation after they are no longer needed for the practical purposes for which they were originally created. Therefore, the process described in the “Preserve Electronic Records” model begins with the transfer of the records from their creator, or from an agent acting for the creator, to a person whose primary responsibility is that of preserving authentic records—that is, the preserver. However, the preserver, as defined by the InterPARES project, has responsibilities that are broader than the preservation process itself. For example, the preserver is presumed to be responsible for selecting the records that are to be preserved. In the “Preserve Electronic Records” process model, the viewpoint is literally and strictly that of “the person responsible for preservation.” The model’s viewpoint includes only those entities and processes that someone, or some organization, carrying out the role of preserving the records would carry out. The same person or organization may have other roles or other, coincidental responsibilities, such as appraisal or reference, but coincidental responsibilities are excluded from the “Preserve Electronic Records” model. The role of preserving records includes all and only those activities necessary to ensure the transmission of authentic electronic records over time.

The “Preserve Electronic Records” model is intentionally generic. It identifies and describes the processes necessary to preserve electronic records, articulates the inputs needed by each process, the controls under which it operates, the mechanisms necessary to accomplish the process, and the output(s) produced by each process. The model defines the relationships among these entities and processes. While the model is systematic, it does not prescribe an implementation. Rather than defining a preservation system, the “Preserve Electronic Records” model provides a comprehensive, precise, and coherent road map that institutions and persons concerned with the preservation of electronic records can use in designing, developing, and evaluating systems which address their specific requirements, objectives, and constraints.

**Data**

The multidisciplinary, expert knowledge of the members of the Preservation Task Force was the principal basis for elaborating the “Preserve Electronic Records” model. In addition, task force members from institutions that are actively engaged in the preservation of electronic records supplied information about their institutions’ practices, and plans. Two walkthroughs were conducted to test and refine the model using empirical data. The first walkthrough was of a hypothetical system defined by combining elements from several actual systems supporting related business functions, including two legacy systems used in different organizations, and a third system that is being deployed in one of these organizations. In order to provide a rigorous test of the model, the hypothetical system was made more complex by the addition of several data types. This walkthrough was used to test version 4 of the model and refine it to produce version 5. The second walkthrough was of a single empirical case. It was applied to version 5 and used to derive version 6. This walkthrough is described in an appendix.\(^5\)

**Findings**

**Survey Findings**

Responses to the survey of digital preservation practices, plans, and research indicate three broad themes. First, the perception of what preservation is goes beyond archival and library

---

\(^4\) NIST, 1993.

\(^5\) Appendix 7.
practice to the media being preserved. Traditional definitions of preservation may not apply in the digital arena, and a shift is already apparent. Second, the rush to develop the technological processes necessary to preserve authentic electronic records appears to be at the expense of directly addressing cost and policy issues at the start of projects. The problems posed by permanent preservation of authentic electronic records require the development of a unique cost model. Last, the paucity of preservation policies in place represents a distinct gap in the research design of many of the projects and possibly reflects a lack of commitment among the stakeholders in institutions. It appears that meeting the technological challenges of preserving electronic records is more of a priority within these institutions than developing policy. Such prioritization courts the risk that overall progress in this new arena will be more uneven than is necessary.

Results of Analytic Modelling

The process model "Preserve Electronic Records," developed by the Preservation Task Force, took an approach contrary to that discovered by the survey in many projects. Rather than giving priority to the technological challenges of preserving electronic records, the task force developed a model in which alternative technical solutions can be evaluated and adopted as appropriate. The model is motivated by the perception that, while preserving electronic records requires technological solutions, it is impossible to determine whether any given technology constitutes a solution on technological grounds. The criteria for evaluating technologies derive from the archival and institutional requirements that determine the goals and objectives of preservation and act as controls on the preservation process. Technology plays a role in selecting solutions only in that any solution must be feasible: the technology to implement the solution must exist and be applicable. Feasibility also includes affordability. While the "Preserve Electronic Records" model does not include or assume a cost model, it provides for the application of a cost model in developing preservation strategies and plans and in evaluating their execution.

The most fundamental finding that emerged from the structured analysis of the process of preserving electronic records was a paradigmatic shift in the concept of preservation of electronic records. This shift had both archival and technological dimensions. While the phrase "preserve an electronic record" is convenient and undoubtedly will continue to be used, in many variations it is a shorthand expression that belies reality. Empirically, it is not possible to preserve an electronic record; it is only possible to preserve the ability to reproduce the record. That is because it is not possible to store an electronic record in the documentary form in which it is capable of serving as a record. There is inevitably a substantial difference between the digital representation of the record in storage and the form in which it is presented for use. It is always necessary to use some software to translate the stored digital bits into the documentary form of the record. This entails an inevitable risk that, regardless of how well the digital data were protected in storage, the record may be inappropriately altered when the stored bits are retrieved and presented for use as a record. Thus, in contrast to prevailing notions about the preservation of records in hard copy, the process of preserving an electronic record goes well beyond keeping it safely in storage. The process of digital preservation begins with the initial act of storage and extends through reproduction of the record. To reflect the empirical situation, the task force constructed the concept of a "digital component of an electronic record." An electronic record is stored as one or more digital components. Digital components have no necessary relation to the elements of documentary form recognized in diplomatics analysis of records. Rather, they are determined technologically by the way the bits are stored and by the methods (software) that must be applied

---

6 With continuing progress in digital information technology we can expect to reach the point where computers can input information recorded in human-readable form. Nonetheless, this assertion will remain valid: while the digital display of a record—for example, narrative text recorded on paper—may preserve the "look and feel" of the paper version, the digital version will be inscribed on a different physical medium and the process of producing the display version from the stored version may result in alteration of the record.
to reproduce the record. Reproducing an electronic record entails (1) reconstituting it, that is, reassembling its digital components if it has more than one, or extracting any digital component stored in a physical file that contains more than one such component; and (2) presenting it in proper form.

The Process of Preserving Electronic Records

At the highest level, three factors control the process of preserving electronic records. First, in order to preserve records, and especially to preserve them as authentic, we need to know what the requirements are for doing so. These requirements are archival in nature: they derive from archival science and principles and from related standards and best practices for managing records. Second, preserving electronic records entails using digital information technology; the possibilities for such preservation are limited by the state of the art of information technology, which constitutes the second type of control on the preservation process. Third, the exercise of the preservation function will also be governed by requirements of the institution in which this function is carried out.

Three mechanisms are necessary to perform the preservation process: an information and communications technology infrastructure; facilities where the electronic records will be stored and processed; and persons competent to carry out the process. While the state of the art of technology determines what is possible and impossible to do, the technology infrastructure comprises the hardware, general-purpose software, and physical media used to store and process the digital components of electronic records. These mechanisms are used in all preservation activities.

There are two primary inputs to the process of preserving electronic records. The first, and most obvious, are transfers of electronic records selected for preservation. In simple terms, the records are what the process is all about. Records are preserved because they have been determined to have enduring value. That value is realized in use. So the second primary input consists of requests for the records, or for information about them. The preservation process also needs a third input, information about the records that have been selected for preservation. This information is necessary to determine what preservation methods, information technology, facilities and staff will be needed to preserve the records and to organize the process to guarantee that the records can be preserved as authentic.

The preservation process produces two primary outputs: reproduced electronic records and reproducible electronic records. A preservation system outputs a reproduced electronic record when the record is reconstituted and presented within the system itself. However, in many cases those who want to access a preserved record will want to do so on systems outside of the preservation system, such as in web browsers on their own computers. In such cases, the preservation system can output only the digital component(s) of the record along with instructions on how to reconstitute and present the record; that is, it outputs a reproducible electronic record. There are two other, derived, outputs of the preservation process: a certificate of authenticity, which attests to the authenticity of a reproduced record; and information about preservation, which attests to the integrity and reliability of the preservation process overall. A certificate of authenticity is produced when a requester demands tangible evidence that a reproduced electronic record is authentic. Information about the overall system and processes of preservation is produced either as required by higher authorities or in response to a challenge to their adequacy or appropriateness for preserving authentic records.

The process of preserving electronic records includes four principal sub-processes: a management process and three processes that carry out or execute preservation. The management process governs the other three. It establishes a comprehensive approach, which is executed in the three other processes, and it evaluates these processes to ensure that the goals and objectives of preservation are achieved. To do this, the management process interprets the external archival and institutional controls into a coherent synthesis of requirements, or preservation framework, which controls other management sub-processes, as well as the execution processes.
In each case where the appraisal process identifies a body of records as worthy of preservation, the preservation management process determines whether it is possible to preserve the records—given the technical characteristics of the records and the state of the art of information technology—and, if so, how they will be preserved. This determination feeds back into the appraisal process, enabling a two-part decision that the body of records both has enduring value and can be preserved. For each body of records thus selected for preservation, managing the preservation process requires articulating a preservation strategy.

The preservation strategy encompasses a set of rules or procedures for processing the records in each of the execution activities; criteria for determining whether each process defined in the strategy is executed properly and achieves its desired outcome; and specific technological methods that will be used to preserve the records, up to and including their reproduction. The preservation strategy acts as a control on the execution processes. A preservation strategy will entail requirements for specific information technology infrastructure needed in order to implement the strategy. Preservation management thus includes sub-processes to identify and acquire the technological infrastructure and the technological preservation methods that will be used in preserving the records. The technological preservation methods used to preserve the records are specific to classes of digital components and control the processing and maintenance of those components over time and the reproduction of records from the components; the technological infrastructure enables these methods to be executed. For example, a preservation strategy might prescribe that, in the case of textual records whose visual appearance is critical for authenticity, the records will be preserved as bitmapped images. Preservation methods used to implement such a strategy would include software to convert textual records from other formats, such as word-processing files, to bitmaps and software to render such bitmaps for viewing. The technological infrastructure needed to execute the software would include appropriate processors, storage devices, display devices, and drivers. Basically, preservation methods directly support the preservation and reproduction of electronic records from their digital components, and the preservation infrastructure supports the execution of preservation methods.

A preservation strategy also defines specific actions that should be taken with respect to the body of records, either at specified times, such as when the records are first brought into the preservation system, or under specific conditions, such as when the media on which the digital components of the records are stored need to be replaced. For example, a preservation strategy of preserving textual records as bitmapped images would entail converting textual records transferred in different formats to bitmaps. The same strategy should also specify what to do if the software used to display the records stored in bitmaps becomes obsolete. The preservation strategy remains constant, unless there is a management decision to change it.

Each execution process must produce information about itself and about the results achieved in its execution that is appropriate and adequate to enable management to determine whether a preservation strategy is successful and, if not, what corrective action is needed.

The reports of the Authenticity and Appraisal Task Forces give rise to different types of findings, specifically highlighting the need to review the work of the three task forces with the objective of synthesizing results where appropriate and of identifying where additional analysis is needed to align the products.

One such undertaking would be to develop a third IDEF(0) model that links those of the Appraisal and Preservation Task Forces. The Appraisal Task Force’s model may be described as constructed from the viewpoint of the preserver exercising the role of selecting electronic records for preservation, while the Preservation Task Force’s model may be described as constructed from the viewpoint of the preserver exercising the role of preserving electronic records. The third model would take the viewpoint of the preserver given its responsibility for coordinating both processes. This effort would not be merely academic. A substantial result that should be expected from constructing the third model would be articulation and clarification of the feedback loop between selection and preservation. Currently, when appraisal identifies a body of electronic records as having enduring value, information is needed about the feasibility of preserving the records. In the case where the preservation system has the capability and capacity to preserve.
the records, confirmation of this fact may be all that is needed to reach a selection decision. However, most cases will require more extensive communication between the preservation and selection processes. For the records to be preserved successfully, the two processes must reach complete agreement on terms and conditions for transfer of the records from the active system to the preservation system or, alternatively, from the state of active or open records to the state of closed, inactive, preserved records. These terms and conditions prescribe the initial steps in the preservation strategy. Where the preservation system does not have the capability or capacity of preserving the records, there should be additional communication between the two processes concerning requirements, alternatives, costs, and other related factors. Furthermore, to develop an adequate preservation strategy for a body of records, preservation management will need information about the appraiser’s benchmark assessment of authenticity as soon as it is available.

It does not appear that the “Preserve Electronic Records” model needs to be modified in any substantial way to accommodate the benchmark and baseline requirements produced by the Authenticity Task Force. The Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records do not apply to the records themselves nor to the preserver. Rather, they are criteria that the appraiser should use to assess authenticity when selecting records for preservation. The result of applying the benchmark requirements is information articulating a presumption of authenticity. The "Preserve Electronic Records" model provides for receipt and preservation of this information as part of the chain of preservation. This model also provides an opportunity for updating the assessment when records are examined as part of the process of bringing them into the preservation system. The Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records do apply to the records and to the preserver, but they are largely contextual in character. The "Preserve Electronic Records" model can satisfy these requirements as it stands, although it would probably be beneficial to make this more explicit.

Nonetheless, there are points that should be explored simultaneously from both authenticity and preservation perspectives. For example, the first baseline requirement requires that “the content of the record remains unchanged after reproduction.” Given that this requirement applies to “transfer, maintenance, and reproduction,” clearly the operative meaning of “unchanged” is with respect to the state of the content as delivered by the records creator. However, a variety of factors, such as the fragility of digital storage media, may result in some partial loss or corruption of content. The requirement should be enhanced by specifying that any such loss or corruption should be documented and, perhaps, by indicating when such problems would be critical.

The benchmark requirements include provision for documenting whether the creator established the documentary forms of records. The baseline requirements require documentation of “the impact of the reproduction process on their form.” If the creator did not articulate documentary form, there is no obvious basis for determining the impact of reproduction. There is a large gap between these processes that needs to be addressed.

Products

The Preservation Task Force produced a detailed IDEF(0) model of the process of preserving electronic records, a report explaining basic concepts of the model and providing simplified views of the model, a case study illustrating application of the model, a report on the results of its survey of current digital preservation practices, and a report on digital storage media:

- IDEF(0) Model, "Preserve Electronic Records"
- How to Preserve Authentic Electronic Records
- Walkthrough Applying the "Preserve Electronic Records" Model.
• P.C. Hariharan, "Media." <http://is.gseis.ucla.edu/us-interpares/Mediareport.pdf>

The first three of these are included as appendices.7

**Relationship to Existing Standards**

**The Open Archival Information System Reference Model**

The basis for the content of the preservation process model is the Open Archival Information System (OAIS) Reference Model, a new ISO standard that was being developed while the Preservation Task Force articulated the "Preserve Electronic Records" model.8 “An OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community.” The "Preserve Electronic Records" model is built on the basic assumptions of the OAIS that the records are produced outside of the archival system, that they are to be available to a user community which is also outside of the archival system, and that the archival system is thus a mediator which takes information from producers and delivers it to users over long periods of time. Thus the OAIS model has a much broader scope than the "Preserve Electronic Records" model. The reference model is intended to apply to any type of information, not just records. For example, the information preserved in an OAIS might be scientific data, or it might be information about physical objects in a museum. At a high level, it may be said that the "Preserve Electronic Records" model is a specification of an OAIS for the specific classes of information objects comprising electronic records and archival aggregates of such records.

---

7 Appendix 5, Appendix 6, and Appendix 7.
It is necessary to distinguish between the function described by the "Preserve Electronic Records" model and a system that would implement the model. The preservation function might be carried out by a system that provides only the functionality described in the model. But it might equally well be implemented in a system that features additional functionality, including the appraisal of records, the management of current and temporary records, and reference and dissemination functions.

This reveals another aspect in which the "Preserve Electronic Records" model is narrower than the OAIS: the preservation model does not include all activities related to making records available—only those that are inextricable from the preservation function. The preservation function extends to the production of copies of records, because that is necessary to guarantee their authenticity, but it does not include order agreements as described in the OAIS model or any "value-added" dissemination or access services. Similarly, the preservation model does not include processes, which inform potential users what records are being preserved or what conditions govern access to the records.

While the "Preserve Electronic Records" model is narrower than the OAIS model, the InterPARES model has substantial more depth on the topic of preservation in general and, obviously, the preservation of authentic electronic records in particular. The Preservation Task Force communicated its work to the committee responsible for the OAIS standard and worked with that committee to enhance the standard in light of our findings.
Information Technology Standards

The “Preserve Electronic Records” model does not explicitly adopt or implicitly entail specific information technology standards, such as those concerning various digital data formats, storage media, information interchange, etc. Instead, the model provides a context, specifically in the preservation management process, for identifying, evaluating and adopting such standards as appropriate. One of the principal controls on the entire preservation process, the state of the art of information technology, includes current standards. Other things being equal, the preserver should develop preservation strategies that adopt those standards that best support archival objectives; however, other things often will not be equal. In archives of corporations and universities, for example, the information technology infrastructure may be largely determined by corporate information technology architecture, leaving the preserver no option but to develop strategies that can be implemented on that infrastructure. Other major factors that will constrain the preserver’s adoption of standards include the costs of doing so and the availability of products that implement the standards as well as support for those products. For example, other things being equal, to achieve the archival goal of permanent preservation, the preserver would tend to select storage media that are subject to standards and are durable. However, in an environment of continuous change in information technology, the preserver needs to anticipate that the longer any given type of digital medium is kept, the more expensive and difficult it will be to maintain.

Conclusion

Much attention to the preservation of electronic records has focused on the twin problems of the relatively short life expectancy of digital media and the rapid obsolescence of hardware and software. The InterPARES project started with recognition of these problems and cast the preservation issue in terms of evaluating practical methods for solving them. The research plan called for the Preservation Task Force “to identify and develop the procedures and resources required for the implementation of the conceptual requirements [for preserving authentic electronic records] and criteria [for appraising electronic records] identified in the first two domains.” This formulation of the problem of preserving electronic records clearly situates it not in technology, but in the interface between the goal of preserving electronic records and the technology on which they depend. Technology itself is not a problem. If we did not need to preserve records beyond the life expectancies of hardware, software, and digital media, we would not have any preservation problem. Similarly, technology cannot determine the solution. It is archival and records management requirements, that define the problem. It must be archival and records management criteria that determine the appropriateness and adequacy of any technical “solution.” The question “What is the best technological method for preserving electronic records?” is as meaningless as the question “What is the best medicine for making people healthy?” Neither can be answered without specifying the conditions they are meant to address. The InterPARES project defined these conditions as the archival requirements for authenticity and the archival criteria for selecting records to be preserved.

As previously stated, because the InterPARES task forces on authenticity, appraisal and preservation worked in parallel, the Preservation Task Force could not formulate solutions based on specific conceptual requirements and criteria. Nonetheless, through communications and cross-fertilization among the task forces in the entire course of the research, the Preservation Task Force has been able to produce a model of the process of preserving electronic records that does in fact identify the procedures and resources needed to implement the requirements and criteria. The procedures are the processes defined in the “Preserve Electronic Records” model, and the resources include the mechanisms needed to carry out these processes as well as the information about both the processes and the records that needs to flow across processes. This model does not describe a computer system, and it does not itself reach conclusions about what technological systems, tools, or methods are best suited for preserving electronic records. Rather, it provides an extensive, detailed and highly coherent framework for identifying and analyzing the specific challenges faced in implementing appraisal decisions that select specific bodies of electronic records to be preserved. This framework guides the evaluation of
technological options and the articulation of specific preservation strategies addressing both the archival and technological characteristics of the records to ensure the continuing availability of authentic copies of the records across time and generations of technology.

Thus the "Preserve Electronic Records" model can be a guide to implementation, but it does not prescribe an implementation. There is greater value in this model than there would be in one that described how to design a particular preservation system. It would be simplistic, and erroneous, to assume that a single technical solution would be optimal in all circumstances. The "Preserve Electronic Records" model can be used to develop solutions that address varying circumstances, including not only diversity in the characteristics of the records to be preserved, but also variety in the external requirements imposed on the preserver, and in the goals and objectives to be achieved in preserving the records.

**Recommendation 1.** The primary recommendation that comes out of this work, then, is for analysts and institutions to use the "Preserve Electronic Records" model as a framework for developing solutions to the challenges of preserving electronic records.

**Recommendation 2.** Use of the "Preserve Electronic Records" model should be based on understanding of the particular characteristics of electronic records and what those characteristics entail for preserving these records, as summarized in the foundation concepts:

- a) Digital Components of Electronic Records,
- b) Preservation Control,
- c) Archival Requirements for Preservation,
- d) "Original" Electronic Records,
- e) The Need to Reproduce Electronic Records, and
- f) The Chain of Preservation.

These concepts are set out in Appendix 6, "How to Preserve Authentic Electronic Records." Key to all of these concepts is the recognition that the chain of preservation for electronic records must extend over their entire life and that the process of preserving electronic records extends to and includes reproducing them.

**Recommendation 3.** Solutions to the preservation of specific bodies of electronic records should be inherently dynamic. The solutions need to be dynamic for two different reasons. First, most archives and other preservers will accumulate electronic records over time. Over time, the specific properties of the records being brought into the archives will change. The preservation system must be capable of being expanded, adapted, or modified to accommodate new and different types of electronic records, and new ways of organizing, accessing, and presenting such records. Second, the goal of preserving electronic records is not to keep them, in archives or elsewhere, but to make them available to persons who have a need for, or an interest in, them. While the preserver has a fundamental responsibility for providing access to authentic records, their availability will be impacted by the continuing evolution of information technology. Preservers should assume that future users would want to use the best available technology for access to the records. The design of preservation systems should take into consideration the need to be able to interface with evolving technologies for information discovery, retrieval, communication and presentation.

**Recommendation 4.** The InterPARES Project has been so fruitful that it has gone far beyond providing valuable products in response to the research questions that it originally posed. It has also raised the threshold of research by articulating issues that are entailed by the original questions, but not explicit in them, by identifying new questions and by opening up lines of research that should provide grounds for valuable results for years to come. For example, the project has moved beyond its foundation in the science of diplomatics to recognize that, in the digital environment, many of the concepts and methods that traditionally were applied to individual documents need to be applied to sets of records. This insight needs to be explored more fully. The work of the Preservation Task Force has focused on defining a comprehensive framework for preserving authentic electronic records. This work should not stop when the current
project ends. The archival profession, our collaborators, and our stakeholders, have an interest and responsibility to see that further progress is made.

Much more work is needed to analyze the data and information requirements for executing the processes defined in the preservation model.

The model should also be applied to additional test cases both to validate and enrich it. The model should also be extended to address the application of specific technologies for overcoming technological obsolescence.

Methods must be developed for analysis and categorization of the documentary forms of electronic records and criteria for determining which elements or aspects of documentary form must be preserved to ensure the integrity of the record.

While the Authenticity Task Force found that it was not possible to construct a typology of electronic records from which requirements for authenticity could be derived, the concept of authenticity elucidated by that task force entails preserving documentary form. Benchmark Requirement 5, for assessing the authenticity of records, requires evidence that “the creator has established the documentary forms of records associated with each procedure either according to the requirements of the juridical system or those of the creator.” Similarly, Baseline Requirement 2, for reproducing authentic copies of electronic records, entails documenting “the impact of the reproduction process on their form, content, accessibility and use.”

There is a significant opportunity for the InterPARES project to contribute to the enrichment of the Open Archival Information System.

While the scope of the OAIS model extends far beyond the domain of records, that model could be informed by archival understanding of authenticity. Regardless of the nature of the information objects being preserved, those responsible for preserving them should be able to attest to and explain the authenticity of the products they deliver to their customers. Such a need is signalled by the concern in many disciplines of natural science with “data lineage” or “data parentage.”

The accomplishments of the InterPARES Project should be applied to related areas of concern, such as the process of archival description.
Strategy Task Force Report

Task Force Members

Luciana Duranti, University of British Columbia (Project Director, Chair)
Heather MacNeil, University of British Columbia
Terry Eastwood, University of British Columbia
Ken Thibodeau, U.S. National Archives and Records Administration
Sharon Farb, University of California at Los Angeles
Jason Baron, U.S. National Archives and Records Administration
Introduction

The goal of the Strategy Task Force was to define the principles that should guide the development of international strategies and standards for the long-term preservation of authentic electronic records, and the criteria for developing from them national and organizational policies and strategies.¹ Chaired by the project director, its members are the chairs of the three task forces on authenticity, appraisal, and preservation; an expert on legal and policy issues (Jason Baron); and an expert on standards, especially those related to copyright (Sharon Farb).²

The intellectual framework developed by the Strategy Task Force derives from the distillation and synthesis of the findings of the Authenticity, Appraisal, and Preservation Task Forces. The definitions of the key concepts expressed in the framework can be found in the international, interdisciplinary glossary produced by the Glossary Committee.

The original InterPARES research plan identified three questions that were to be addressed within Domain 4, which is the area of inquiry under the responsibility of the Strategy Task Force:

- What principles should guide the formulation of international policies, strategies, and standards related to the long-term preservation of authentic electronic records?
- What should be the criteria for developing national policies, strategies, and standards?
- What should be the criteria for developing organizational policies, strategies, and standards?

The distinction between international, national, and organizational policies, strategies, and standards derives from the recognition that each cultural, juridical, and organizational environment has its own needs, which must be articulated in a manner suitable for that environment. In each case the articulation of policies, strategies, and standards will reflect the synthesis of archival requirements for authentic electronic records with other requirements and constraints operative in that environment, as well as with possibilities and mechanisms for acquiring and committing resources and obtaining the support and commitment of stakeholders.

This document intends to provide a framework for the articulation of policies, strategies and standards, ensuring that they are well-grounded and consistent. It has therefore been necessary to take into account applicable laws and regulations; general policies concerning archives, records management, information technology, and client service; and realistic assessments of resource availability and stakeholder commitment. This document represents the summary of the principles and criteria that have emerged from the research project.

---

¹ A policy is a formal statement of direction or guidance as to how an organization will carry out its mandate, functions or activities, motivated by determined interests or programs. A strategy is the complex of practical means formally articulated by an entity for reaching a specific purpose, that is a plan or a road map for implementing policies. A standard is the complex of established norms aiming to make the characteristic of a product, process, or service uniform within or across a sector, a country, or a system. In technical or industrial use, a standard is a model or a type of a given product.

² Copyright and intellectual property issues may arise from the fact that we cannot preserve electronic records but only copies of electronic records.
The Development of the Intellectual Framework

Frameworks are shared maps for advancing understanding and implementing concepts, ideas, and issues. Similar to other maps, intellectual frameworks are used to model concepts by capturing important dimensions. The main purpose of an intellectual framework as a model is to "serve as an instrument of understanding." The value of an intellectual framework is the "real world ability to generate insights about issues."

The Strategy Task Force developed the intellectual framework through analysis and synthesis of results of work in the first three InterPARES domains. Taking those results as its data, the task force distilled principles and criteria for the formulation of policies, strategies, and standards. This framework reflects a multidisciplinary, international, and iterative collaboration among the task forces and international and national research teams.

Building on the work done by Lal Verman, related to developing a logical model of standardization, and Yesha Y. Sivan, related to the development of a five dimensional framework for development of standards, the Strategy Task Force has designed an intellectual framework that can be used by international and national bodies, and individual organizations to understand and implement the research findings and results of the InterPARES Project. An intellectual framework is not intended as a tool to develop theory; rather, it can be used as a tool to promote and communicate a shared understanding of the key concepts, issues and proposed solutions related to the long-term preservation of authentic electronic records. This framework for the development of policies, strategies, and standards related to the long-term preservation of authentic electronic records has three dimensions: purpose, principles, and criteria.

The Intellectual Framework

1) Purpose
The purpose of the intellectual framework is to support the development of policies, strategies, and standards.

2) Principles and Criteria
Principles are statements that have general validity in a given sector or field. In applied sciences, they are conceptual statements on which a science, an argument, or a reasoning is based, derived from the observation of individual facts. Criteria are the norms on which distinctions are based, judgements are made, and different lines of action or conduct are decided.

The extensive and in-depth investigations of the InterPARES project over three years have been distilled into a set of fourteen principles and corresponding criteria for the development of policies, strategies, and standards.

6 Knowledge Age Standards. See also Y.Y. Sivan, Nine Keys to a Knowledge Infrastructure: A Proposed Analytic Framework for Organizational Knowledge Management (Cambridge, MA: Center for Information Policy and K2K Inc., 2001).
Any records preservation policy, strategy, or standard should:

<table>
<thead>
<tr>
<th>Principle</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. address records specifically rather than digital objects generally; that is, it should address documents made or received and set aside in the course of practical activity.</td>
<td>A record is distinguished from other digital objects by virtue of the fact that it possesses a fixed documentary form, a stable content, an archival bond with other records, and an identifiable context. It participates in or supports an action and at least three persons are involved in its creation (i.e., an author, a writer, and an addressee).</td>
</tr>
<tr>
<td>2. focus on authentic electronic records.</td>
<td>An authentic electronic record is one that is what it claims to be and that is free from tampering or corruption. Accordingly, proving the authenticity of an electronic record involves establishing its identity and demonstrating its integrity on the basis of the benchmark and baseline requirements for authenticity. The identity of a record refers to the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records. The integrity of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects. This does not mean that the record must be precisely the same as it was when first created for its integrity to exist and be demonstrated. When we refer to an electronic record, we consider it essentially complete and uncorrupted if the message that it is meant to communicate in order to achieve its purpose is unaltered.</td>
</tr>
<tr>
<td>3. recognize and provide for the fact that authenticity is most at risk when records are transmitted across space (i.e., when sent between persons, systems, or applications) or time (i.e., either when they are stored offline, or when the hardware or software used to process, communicate, or maintain them is upgraded or replaced).</td>
<td>Assertions that electronic records are more susceptible to tampering and corruption than traditional, hard-copy records need to be placed in context. While threats to the integrity of electronic records undoubtedly exist, digital information technology offers possibilities for very strong protection of their integrity. These possibilities are strongest within the confines of a specific system. For example, it is possible to track every access to a records system and every action on any record in the system. A system can be designed so that, once filed, a record is never out of file: users get access only to copies of the record. System design can also preclude any alteration or destruction of records except by authorized persons. Simple procedures such as redundant storage and regular back-up can also make it easy to recover from any inappropriate alteration or deletion. However, such controls are only effective within the confines of a system. When a record is taken out of a system, or when the system itself is modified, systematic control is at risk.</td>
</tr>
</tbody>
</table>
4. recognize that preservation of authentic electronic records is a continuous process that begins with the process of records creation and whose purpose is to transmit authentic records across time and space. This process is defined as “chain of preservation,” that is, a system of controls that extends over the entire life cycle of records and ensures their identity and integrity in any action that affects the way the records are represented in storage or presented for use. The benchmark requirements for assessing the authenticity of the creator’s electronic records define evidence that demonstrates how the records creator established and maintained the chain of preservation while the records remained in its custody. The baseline requirements supporting the production of authentic copies of electronic records articulate what the preserver must do to ensure that the chain remains unbroken from the moment the records are transferred to the archives. 

5. be based on the concept of trust in records keeping and record preservation and specifically on the concepts of a trusted record-keeping system and the role of the preserver as a trusted custodian. Records should be made and maintained in a trusted record-keeping system and preserved by a trusted custodian. A trusted record-keeping system comprises the whole of the rules that control the creation, maintenance, and use of the records of the creator and that provide a circumstantial probability of the authenticity of the records within the system. To be considered a trusted custodian, the preserver must demonstrate that it has no reason to alter the preserved records or allow others to alter them, and is capable of implementing all of the baseline requirements.

6. be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object: it is only possible to preserve the ability to reproduce the record. Reproducing an electronic record means to be able to render it with the content and any required elements of documentary form and annotations that such record possessed before reproduction.

7. recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form. A digital component is distinguished from an element of documentary form on the basis of the fact that a digital component is a digital object that contains all or part of the content of an electronic record, and/or data or metadata necessary to order, structure, or manifest the content, and that requires specific methods for preservation. In contrast, extrinsic and intrinsic elements of form are those characteristics of a record that constitute its external appearance and convey the action in which it participates and the immediate context in which it was created.
<table>
<thead>
<tr>
<th>8.</th>
<th>specify the requirements a copy of a record should satisfy to be considered equivalent to an original.</th>
</tr>
</thead>
</table>
| In principle, an original electronic record is the first complete and effective record. However, in an electronic environment, no original survives. Every faithful copy of such a record's content and of its documentary form is to be considered a copy in the form of the original, which is equivalent to the original as to its consequences. Any kind of copy that is declared authentic by an officer entrusted with such a responsibility is also equivalent to the original. 
| 9. integrate records appraisal in the continuous process of preservation. |
| Records should be selected for long-term preservation on the basis of their continuing value, assessment of their authenticity, and the feasibility of their preservation. |
| 10. integrate archival description in the continuous process of preservation. |
| Archival description should serve as a collective attestation of the authenticity of the records and their relationships in the context of the fonds to which the records belong in conformance with the baseline requirements. |
| 11. explicitly state that the entire process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term. |
| To support the assertion of the authenticity of preserved electronic records, the preserver should document, at a minimum: the records creator's practices to support a presumption of authenticity, in accordance with the benchmark requirements for authenticity; and the processes of bringing the records into the archives and maintaining them over time, and the reproduction of records, in accordance with the baseline requirements for the production of authentic copies of records. |
| 12. explicitly recognize that the traditional principle that all records relied upon in the usual and ordinary course of business can be presumed to be authentic needs to be supplemented in the case of electronic records by evidence that the records have not been inappropriately altered. |
| In addition to the evidence that they were created and used in the usual and ordinary course of business, records should be presumed authentic on the basis of the criteria listed as benchmark requirements in the "Authenticity Task Force Report" or verified authentic by the preserver. |
| 13. recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records. The assessment of the authenticity of electronic records takes place before records are transferred to the custody of the preserver as part of the process of appraisal, while the maintenance of the authenticity of copies of electronic records takes place once they have been transferred to the preserver's custody as part of the process of long-term preservation. |
| The assessment of the authenticity of electronic records should be based on the benchmark requirements, while the maintenance of the authenticity of copies of electronic records should be based on the baseline requirements. |
| 14. draw a clear distinction between the preservation of the authenticity of records and the authentication of a record. |
| Authentication is a declaration of a record's authenticity at a specific moment in time by a juridical person entrusted with the authority to make such declaration. It takes the form of an authoritative statement, which may be in the form of words or symbols, that is added to or inserted in the record attesting that the record is authentic. |
Digital signatures—which identify the sender of a data object and verify that it has not been altered in transmission—can support the authentication of electronic records, but they are not sufficient to establish the identity and demonstrate the integrity of an electronic record over the long term.\textsuperscript{iv}

Conclusion

According to Lewis Branscomb and Brian Kahin, the development of a theoretical framework provides for more consistent policy development, contributes to the formulation of models of best practices, and establishes a process for addressing interoperability, harmonization, and improvement of cost efficiencies.\textsuperscript{7} Branscomb argues that, in the absence of a coherent intellectual framework, conflicts arise, because traditionally distinct domains and multiple approaches result in inconsistency, lack of interoperability, permanent loss of the cultural, historical, and financial record and the inability to implement long-term strategies and goals.\textsuperscript{8}

The InterPARES Project has been carried out by an international, multidisciplinary group conducting investigations in the context of research units working separately. Consistency of result was ensured by the regular meeting of representatives of the research units in the context of the International Team workshops, and by the work of the Glossary Committee. However, the development of this intellectual framework has been invaluable to test the actual coherence of the three final reports on authenticity, appraisal, and preservation. The concepts on which the research carried out by the task forces that authored those reports was based had themselves been developed in the context of an intellectual framework for the InterPARES Project generated by assembling several theoretical and methodological hypotheses. This common foundation of the research of the three task forces did yield consistent findings that could be integrated into an intellectual framework for strategies, policies, and standards development. Some refinements of the three reports occurred in the course of the articulation of the intellectual framework, but only for the purpose of making explicit their existing intrinsic coherence. The entire process showed us the vital importance of an intellectual framework for guaranteeing that the long-term preservation methods adopted in contexts that are administratively, legally, and culturally diverse will be coherent, interoperable, and cost-effective. Given the interconnection between policies, strategies, and standards, this intellectual framework provides a comprehensive basis for ensuring consistency also across policies, strategies, and standards for the preservation of authentic electronic records.

---


\textsuperscript{iii} See Appendix 2; Appendix 6, p 7.

\textsuperscript{iv} See Appendix 2.


\textsuperscript{vi} See Appendix 2, p 7; Appendix 6, p 8.

\textsuperscript{vii} See Appendix 6, p 5.

\textsuperscript{viii} See Appendix 6, p 8; "Authenticity Task Force Report", p 23.


\textsuperscript{8} Ibid.

x See Appendix 2, p 8.


xiii See Appendix 2.

Australian Research Team Report

Prepared by Livia Iacovino and Sue McKemmish, Australian Research Team
Introduction

Australia’s archival system reflects the country’s federal structure; each state and territory, and the Commonwealth at the national level, has legal and administrative arrangements in place for the management and preservation of records. Legislative reform in recent years has resulted in new laws that reflect changing directions in the role of archival authorities, establishing them as standard-setting bodies concerned with the totality of record activity—with both the long term preservation of archival records and the accountable management of public records in the interests of good governance.¹

The Australian records and archives industry and its network of national and state archival institutions have developed cooperative arrangements that ensure consistency of record-keeping practice. Vigorous debate on both theory and practice has taken place. The records continuum model has been adopted as the framework in which diverse developments have emerged—from the Australian Records Management Standard, which provided the basis of the International Standard, to record-keeping metadata standards.

Legislation, policy, standards, codes of practice, and guidelines are some of the tools that archival authorities use to establish good record-keeping principles and practice. Usually standards are issued nationally or by a lead state, supported by guidelines and codes of best practice. Other jurisdictions then adopt and adapt these benchmark standards and codes of practice. For example, the Australian Records Management Standard was issued as a voluntary national standard, but was mandated by most archival authorities in their jurisdictions. It has also been adopted by private sector organizations, particularly those seeking accreditation under the quality suite of standards (ISO 9000 suite). The development of legislation, standards, codes of practice, and guidelines takes into account international standards and benchmark practice, and the outcomes of research projects nationally and internationally. For example, the Australian Records Management Standard and related standards issued by national and state archival authorities have drawn on the findings of both the Pittsburgh Project, especially the “Functional Requirements for Electronic Records,” and the first UBC Project; while the record-keeping metadata standards issued by the National Archives of Australia and the State Records Authority of New South Wales were developed within the framework provided by the Monash Recordkeeping Metadata Schema.² This schema in turn drew on analyses of existing national standards.

---

¹ No distinction is made between archives and records in most Australian archival acts. Since 1997 the title of all acts adopts the term record rather than archives. The change reflects the archival authority’s role over the totality of records through record-keeping standards. Victoria dispensed with the legislative distinction between records and archives as early as 1973 in the Public Records Act. The Commonwealth Archives Act 1983; State Records Act 1997, South Australia; and State Records Act 2000, Western Australia do not have the distinction. However, the New South Wales Act (1960) established three categories of record (public records, public archives, and state archives). In the State Records Act (1998), the distinction between “State record” and “State archive” is retained, the latter term now defined to mean records (regardless of location) over which the new State Records Authority (SRA) has assumed control. From Chris Hurley, "From Dust Bins to Disk-Drives and Now to Dispersal: the State Records Act (New South Wales) 1998," Archives and Manuscripts 26, 2 (November 1998), fn. 1, 407.


Apart from the Australian Capital Territory and the Northern Territory, all Australian administrations have records legislation or bills. Both the Australian Capital Territory and the Northern Territory have exposure drafts and a bill respectively; Queensland has a Public Records Bill 1999.

and international archival descriptive standards, and the metadata requirements explicit or implicit in other records and archives standards and statements of best practice.

The outcomes of the Authenticity, Preservation, and Appraisal Task Forces of InterPARES will have an impact in Australia as existing legislation standards, policies, codes of best practice, and guidelines are revised and new tools developed. Processes for taking into account such research outcomes are already a routine part of such revisions and developments in Australia.

Many recent developments in Australia in the areas just described already incorporate the kinds of principles and criteria articulated in the "Strategy Task Force Report." In addition, more generic Australian legislation supports the Strategy Task Force principles, including evidence acts, electronic transaction legislation, and, to a lesser extent, freedom-of-information acts and copyright acts. Privacy legislation on the whole is less supportive, in particular as personal data that identify the parties to a transaction—which is essential to the reliability of the record—are required to be de-identified once their "immediate" use has ceased.

Below are comments from an Australian legislative perspective, together with some policy references, on the records preservation policies, strategies, and standards recommended by the "Strategy Task Force Report."

---


<table>
<thead>
<tr>
<th>Principle</th>
<th>Australian Comments</th>
</tr>
</thead>
</table>
| address records specifically rather than digital objects generally; that is, it should address documents made or received and set aside in the course of practical activity. | Recent Australian archival legislation and standards support a provenancial definition of a record; see for example the ACT Records Bill and the AS4390 definition: “Records ... created or received and maintained by an organisation or person in the transaction of business or the conduct of affairs and kept as evidence of such activity.” Older legislation is more object-based, for example the Commonwealth Archives Act 1983–Sect 3. National Archives policy specifies capture of electronic information that is used for practical activity, even if it does not have all the criteria as identified in the Authenticity Task Force (ATF) report.

The Australian Evidence Act 1995, Dictionary Section 3, Part 1 defines a document as “any record of information.” This is far less specific than records legislation. |
| focus on authentic electronic records. | All agencies as defined in modern Australian archival legislation must make and keep “full and accurate records,” regardless of who has custody. The standard on “full and accurate records” was first developed by SRA NSW (see NSW Records Act 1998 Sect 12), and is based on the Pittsburgh Requirements. Some classes of records have higher degrees of authenticity requirements than others, e.g., see Electronic Transactions Act 1999 Sect 4—exceptions for records of high risk of fraud, including visa and citizenship records.

The Electronic Transactions Act, Sect 11, (3) relates to integrity of information. It states that the integrity of information contained in a document is maintained if, and only if, the information has remained complete and unaltered, apart from:

a) the addition of any endorsement; or
b) any immaterial change; which arises in the normal course of communication, storage or display.

The Copyright Amendment (Digital Agenda) Act 2000, Sect 16B on removal or alteration of electronic rights management information supports the integrity of a work, i.e., that it not be altered; and the author should be linked permanently to it. |
recognize and provide for the fact that authenticity is most at risk when records are transmitted across space (i.e., when sent between persons, systems, or applications) or time (i.e., either when they are stored offline, or when the hardware or software used to process, communicate, or maintain them is upgraded or replaced).  

Provisions that address this principle already exist in archival authorities standards and guidelines, e.g., see the National Archives of Australia "Preservation and migration of electronic records" <http://www.naa.gov.au/recordkeeping/er/summary.html#information>.

recognize that preservation of authentic electronic records is a continuous process that begins with the process of records creation and whose purpose is to transmit authentic records across time and space.


be based on the concept of trust in record keeping and record preservation and specifically on the concepts of a trusted record-keeping system and the role of the preserver as a trusted custodian.

The Australian Standard addresses this principle in specifying record-keeping responsibilities and accountabilities. This principle has also been incorporated in recent law, e.g., NSW 1998 Sect 10; WA Act 2000, Sect 10 and Sect 29 (which establishes the archival authority as auditor with powers to enter premises, or an independent body).

The National Archives policy on "The custody of archival electronic records" also supports this principle.

be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object: it is only possible to preserve the ability to reproduce the record.

The National Archives policy on "The custody of archival electronic records" includes requirements based on this principle—e.g., relating to maintaining records and their associated metadata in an accessible form through successive migrations.

recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form.

See above.
| Specify the requirements a copy of a record should satisfy to be considered equivalent to an original. | The Commonwealth *Evidence Act* Sect 47(2) allows for "a document that is not an exact copy of the document in question but that is identical to the document in question in all relevant respects."

The NSW *Electronic Transactions Act* Sect 11 states: "If ... a person is required to retain ... a document that is in the form of paper ... that requirement is taken to have been met if the person retains ... an electronic form of the document."

The Commonwealth *Electronic Transactions Act 1999* Sect 2 says: "If, under a law of the Commonwealth, a person is permitted to produce a document that is in the form of paper, an article or other material, then, instead of producing the document in that form, the person may produce, by means of an electronic communication, an electronic form of the document." |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrate records appraisal in the continuous process of preservation.</td>
<td>This principle is also embedded in Australian records continuum theory and practice, e.g., the appraisal procedure in <em>DIRKS</em> and the Monash Recordkeeping Metadata Schema.</td>
</tr>
<tr>
<td>Explicitly state that the entire process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term.</td>
<td>This principle is supported by the National Archives policy on preservation (see above).</td>
</tr>
<tr>
<td>Explicitly recognize that the traditional principle that all records relied upon in the usual and ordinary course of business can be presumed to be authentic needs to be supplemented in the case of electronic records by evidence that the records have not been inappropriately altered.</td>
<td>Legal advice on the Commonwealth <em>Evidence Act</em> 1995 states: 'While the 'original document rule' has been abolished, it is still necessary for parties to authenticate evidence of the contents of documents given by one of these alternate ways. For example, in relation to a document in writing that is signed, it remains necessary to lead evidence (if the point is contested) that the signature appearing on the document is the signature of the person who has purported to sign it. In the case of computer records, it is necessary to give evidence that the computer output is what it purports to be.&quot; For more on the <em>Evidence Act</em> provisions, see <em>Records in Evidence</em><a href="http://www.naa.gov.au/recordkeeping/overview/evidence/contents_page.html">http://www.naa.gov.au/recordkeeping/overview/evidence/contents_page.html</a>.</td>
</tr>
</tbody>
</table>
recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records. The assessment of the authenticity of electronic records takes place before records are transferred to the custody of the preserver as part of the process of appraisal, while the maintenance of the authenticity of copies of electronic records takes place once they have been transferred to the preserver's custody as part of the process of long-term preservation.

In Australia, this would be regarded as an implementation strategy issue rather than a principle.

| Draw a clear distinction between the preservation of the authenticity of records and the authentication of a record. | This is not enunciated in Australia, in these terms. However recordkeeping metadata standards require preservation of "in time" person metadata, as well as other metadata to indicate that any changes have taken place. In Australia the current authentication regulatory framework is inadequate because it rejected initial recommendations for a statutorily based central root registration authority. See the National Office for the Information Economy, Establishment of a National Authentication Authority, A Discussion Paper, 19 August 1998 <http://www.noie.gov.au/>. |

---

i S3. Interpretation, "record means a document (including any written or printed material) or object (including a sound recording, coded storage device, magnetic tape or disc, microform, photograph, film, map, plan or model or a painting or other pictorial or graphic work) that is, or has been, kept by reason of any information or matter that it contains or can be obtained from it or by reason of its connection with any event, person, circumstance or thing."

ii "The digital data that make up these email messages, database systems, websites and other information systems have significance beyond the immediate business needs of the Commonwealth agency that creates and maintains them. As this data is created, used, and communicated in the course of an agency's business, it also provides useful evidence—a record—of that agency's past activities." <http://www.naa.gov.au/recordkeeping/er/summary.html#information>

iii S4 "Certain other laws not affected. This section does not affect the operation of any other law of the Commonwealth that makes provision for or in relation to requiring or permitting electronic forms of documents to be produced, in accordance with particular information technology requirements:

(a)on a particular kind of data storage device; or
(b)by means of a particular kind of electronic communication.

Exemption—migration and citizenship documents."

iv In 2000, National Archives of Australia began to develop the technological infrastructure to accept archival electronic records into custody and to provide for their ongoing access over time. NAA acts at this point as the preserver. Agencies act as creators and preservers as they remain responsible for managing electronic records of archival value until they are transferred to the archives custody. This involves maintaining the records and their associated metadata in an accessible form through successive migrations of hardware and software. See <http://www.naa.gov.au/recordkeeping/er/summary.html#information>.

v The role of root certification authority is to support the certification of subordinate intermediate certification authorities and to hold root cryptographic information. The authority of a digital signature depends on the existence of a public key infrastructure (PKI), which is a hierarchical organization of certification authorities
invested with the competence to authenticate the ownership and characteristics of a public key. The effectiveness of such infrastructure depends on the continuity of the chain of trust guaranteed by those certification authorities. As private sector organizations take on the role of certification authorities, there are currently no mechanisms in place to guarantee the continuity of the chain of trust in the event that the organization ceases to exist.
Canadian Research Team Report

Prepared by Bruce Walton, Yvette Hackett and Normand Fortier,
National Archives of Canada
**Introduction**

As a final part of the work of the InterPARES Project on the preservation of authentic electronic records, the general findings relating to authenticity, appraisal, and preservation must be assessed against the specific national and multinational contexts in which they will be applied. The purpose of these investigations is to determine whether there is anything in a given context that will affect, either positively or negatively, the implementation of the project’s findings.

This report looks at the situation of the federal government in Canada. The discussion of the various contexts relates to the government as a whole; the analysis of the individual InterPARES principles focuses primarily on national archives; and the review of the processes of appraisal and preservation addresses the practices of the National Archives of Canada (NAC) specifically.

The overall conclusion is that there is nothing in the InterPARES intellectual framework that is at odds with the various contextual aspects affecting record keeping and archiving in the Government of Canada (GOC). There is, however, a significant gap between the government’s legal and policy framework (which are largely consistent with InterPARES requirements) and the current state of the matter on the ground, meaning that implementation of InterPARES requirements will be possible only once remedial measures that began in the government in 2002 take effect.

**Assessment of the InterPARES principles against GOC contexts**

In this section, the InterPARES principles as defined in the report of the Strategy Task Force are first considered as a set, and assessed against the various contexts that define the record-keeping (creation, management, and preservation) regime of the Government of Canada. Where appropriate, individual principles are then assessed against those contexts as well.

**Principles overall**

**Juridical/administrative context**

This is the key context defining the possibilities for successful implementation of the InterPARES findings. The juridical and administrative elements must be dealt with separately, as their effects are not uniform.

- **Juridical** - Canadian information legislation (*National Archives of Canada Act, National Library of Canada Act, Access to Information Act, Privacy Act, Personal Information Protection and Electronic Documents Act*) provides a strong overall framework within which the InterPARES principles could be implemented. One minor caveat relates to the definition of record used in Canadian legislation, which does not associate records with the business processes they relate to, nor does it include the archival sense of “fixed and set aside.”

- **Administrative** - the overall administrative context of the GOC should be conducive to the implementation of the InterPARES findings, given the presence of policies on the management of government information; the management of information technology, and government security; and of record-keeping metadata requirements for the Government of Canada. However, due to a gap between the framework established by these policies and actual practice in government institutions, the federal government does not currently have an effective information management regime for its records. Initiatives are now under way to address this problem, but it will be several years at least before the state of information management will support a general adherence to the principles enunciated in the InterPARES framework.
Provenancial / Procedural / Documentary contexts

As a result of the overall weakness of information management in government, these are contexts in which InterPARES principles would not currently fit. However, improvements in the administrative context would have a direct beneficial effect here, particularly in strengthening the procedural context to link record creation more closely to the business functions it supports.

Technological context

A technological context appropriate to the realization of InterPARES principles exists (the Records, Documents and Information Management System or RDIMS—an integrated suite of software available to all government departments and agencies), but needs to be more fully implemented across government. Separation between the information technology (IT) and the information management (IM) spheres within government institutions, which frequently leads to IM considerations being inadequately reflected in IT implementations, is another concern, even with a full implementation of RDIMS.

Individual Principles

Any records preservation policy, strategy, or standard should:

- **address records specifically rather than digital objects generally** - this principle is not explicitly reflected in the overall government context, but is directly supported by the electronic records policy and the appraisal methodology of the National Archives of Canada.

- **focus on authentic electronic records** - the management of government information (MGI) policy, although it covers records in all forms, reflects an awareness of the challenge to records’ authenticity presented by electronic systems; the National Archives’ appraisal methodology supports this principle, particularly in the early years of the twenty-first century, when the relative authenticity of parallel paper and electronic versions of records must be considered.

- **recognize and provide for the fact that authenticity is most at risk when records are transmitted across space or time** - this fact is certainly recognized in the MGI policy, although the extent to which it is reflected in current implementations of RDIMS or other record-keeping systems is uncertain; the National Archives both recognizes and provides for this fact in its procedures for transferring, processing, and preserving electronic records.

- **recognize that preservation of authentic electronic records is a continuous process that begins with . . . records creation** - this requirement is reflected in the record-keeping guidance that the National Archives provides to government, in the terms and conditions of transfer that govern the transfer of electronic records to the institution’s custody, and in the NAC’s internal procedures for electronic records.

- **be based on the . . . concepts of a trusted record-keeping system and the role of the preserver as a trusted custodian** - RDIMS fulfills, in theory, the requirement for a trusted record-keeping system in the Government of Canada, but actual current implementations may fall short of fully ensuring records’ authenticity—this gap will be addressed in the coming years with increased NAC involvement in information management across government; the National Archives of Canada fulfills its practices the role of trusted custodian.

- **be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object: it is only possible to preserve the ability to reproduce the record** - the NAC fully understands this principle, and its preservation strategy focuses on maintaining the ability to reproduce records in authentic form (while still accounting for the need to maintain the various physical carriers on which records are stored over time).

- **recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form** - the former point is generally recognized among those who deal with electronic records, but it is highly unlikely that the diplomatic concept of "element of
documentary form,” *per se*, will figure in the federal government’s or the NAC’s policies, strategies, or standards.

- **specify the requirements a copy of a record should satisfy to be considered equivalent to an original** - in the case of the National Archives of Canada, the requirements to be satisfied will relate primarily not to the copy itself but to the system used to transfer, store, and maintain the record’s digital components and to reproduce the record. But the system used will be selected only once it has demonstrated that it can be used to reproduce accurately both the content and documentary form of the record.

- **integrate records appraisal in the continuous process of preservation** - preservation considerations play an important role in the National Archives’ appraisal methodology; the information thus generated will increasingly be reflected in archival descriptions as the institution’s automated information systems become more closely integrated.

- **explicitly state that the entire process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term** - this requirement is not reflected in the management of government information policy; it is, however, met by the NAC’s interpretation of chapter 9 of the Canadian Rules for Archival Description, which calls for the documentation of all transfer and preservation practices.

- **explicitly recognize that the traditional principle that all records relied upon in the usual and ordinary course of business can be presumed to be authentic needs to be supplemented . . . by evidence that the records have not been inappropriately altered** - the management of government information policy does not explicitly include such recognition, but implicitly supports this requirement in its repeated emphasis that information must be maintained in a manner that preserves its authenticity. Similarly, there is no NAC document on electronic records that explicitly states this, but it is certainly institutional practice to document the continuing authenticity of its electronic records over time.

- **recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records** - the National Archives of Canada assesses the authenticity of records before they are transferred; maintenance of authenticity is a critical part of all preservation activities following transfer.

- **draw a clear distinction between the preservation of authenticity . . . and . . . authentication** - it is not certain that this distinction is well understood in government generally, but it is recognized by the NAC, which has reflected the distinction in its guidelines for records created in a PKI environment.

### Processes

This section relates to the work of the National Archives of Canada with historically valuable records.

**Appraisal**

Overall there is very good fit between the NAC’s approach to appraisal and the activity as modelled in the InterPARES findings. At a more detailed level, however, it must be admitted that the NAC does implicitly or somewhat informally what InterPARES requires explicitly and formally—the assessment of authenticity against the benchmark requirements, and the determination of feasibility at the level of record elements. There is nothing in the appraisal process as modelled by InterPARES, however, which the NAC would be unable to do.

**Preservation**

As with the process of appraisal—there is nothing in the InterPARES model of the preservation process that NAC either doesn’t already do or could not begin doing.
Effects and stakeholders

The principal direct effect of the implementation of the InterPARES requirements in the Government of Canada context would be a stronger archival record in electronic form, with a more explicit focus on determining the authenticity of the records being transferred to the NAC, and greater confidence by the NAC in asserting their authenticity when reproduced. The principal indirect effects would be government-wide, and considerable, namely both the costs and the benefits (better current records, improved accountability, increased access) associated with the implementation of the record-keeping infrastructure necessary to meet the benchmark authenticity requirements.

Conclusion

There is nothing intrinsic in the InterPARES principles that would prevent their incorporation in policies, strategies, and standards at the national level in Canada. To the extent that there are obstacles, they relate primarily to the current administrative context in the GOC, and the weak information management found there, which is neither an intrinsically Canadian phenomenon, nor unique to Canada.
Introduction

This report addresses the feasibility of applying InterPARES principles in China. The overall conclusion is that there is nothing in the InterPARES intellectual framework that is strongly at odds with archives management and administration in China. There exists, however, a significant gap between the current Chinese laws and regulations and the practical situations of archival work.

The following assessment is based on the Chinese contexts.

Principles overall

Juridical context

The Archives Law of the People’s Republic of China provides an overall framework within which most of the principles and processes developed by the InterPARES Project could be implemented. But there are some exceptions. For instance, there is no inexorable connection between the definition of an archived record and the life-cycle process of a record in the Chinese archival legal system. Therefore, one or two InterPARES principles might not be suitable to the Chinese context.

Administrative context

In view of the actual situations and future demands of the management of electronic records and archives in China, the findings of the InterPARES research should be fully used for reference in the future development of policies and standards of archives and records management.

Technological context

In China, basically there exists a technological context appropriate to the realization of InterPARES principles, but it is not a universal phenomenon. Starting in 2002, China will integrate information technology with records management systems to guarantee an information management supported more efficiently by information technology through relevant government regulations.

Individual Principles

- *address records specifically rather than digital objectives generally* - This principle is reflected in the Archiving and Archival Management of Electronic Records issued as a
national standard in China.

- **focus on authentic electronic records** - This principle is a basic requirement for developing management methods and standards. At present, the State Archives Administration of China (SAAC) adopts a policy requiring that hard copies be made for important electronic records in order to protect their authenticity.

- **recognize and provide for the fact that authenticity is most at risk when records are transmitted across space or time** - This fact is fully recognized by those government agencies concerned, and archival departments in China, which have taken relevant measures to prevent and reduce the risk.

- **recognize that preservation of authentic electronic records is a continuous process** - The process of continuous preservation of electronic records is explicitly stipulated in standards issued by the SAAC.

- **be based on the . . . concepts of a trusted record-keeping system and the role of the preserver as trusted custodian** - The archival institution is vested with the role of the preserver as a trusted custodian by Chinese law. However, due to the different extent of application of information technology in various record-creating agencies, the capability of fully ensuring the reliability of records in their record-keeping system is still lacking. Where needed, the archival department will take the responsibility of supervision to improve the credibility of records in the record-keeping system.

- be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object: it is only possible to preserve the ability to reproduce the record - There is a different view of this principle in China. Based on the above understanding, we believe that at present we should maintain as much as possible all the physical media of electronic records so as to preserve their authenticity.

- recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form - The former point is increasingly recognized. However, it is not easy to allow the diplomatic concept of “element of documentary form” to be accepted and adopted in Chinese archival practice.

- **specify the requirements a copy of a record should satisfy to be considered equivalent to an original** - In the case of current management situation in China, the requirements to be satisfied will not depend on the copy of a record itself but on the technical acceptance and legal confirmation of a series of processes used to create, transfer, and migrate electronic records. The key factor primarily relates to legal confirmation.

- **integrate records appraisal in the continuous process of preservation** - This principle is clearly reflected in the SAAC standards, which require that archival description should be developed at the same time that appraisal is made for selecting electronic records for archival preservation.
• explicitly state that the entire process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term - This principle is reflected in the SAAC standards.

• explicitly recognize that the traditional principle . . . that the records have not been inappropriately altered - This principle is not carried out properly in China. However, it has been stressed in the administrative practice.

• recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records - Archival departments assess and confirm the authenticity of electronic records when the records are filed for archival preservation. The departments should also maintain the authenticity of the records after they have been transferred to their custody. This maintenance is a mission of archival departments.

• draw a clear distinction between the preservation of the authenticity of records and the authentication of a record - Drawing such a distinction is considered unnecessary in China. An archival department is not entrusted with the authority to authenticate a particular electronic record.

Conclusion

The InterPARES findings provide a helpful reference for the research on the archiving and archival management of electronic records being conducted in China. Most of the InterPARES principles will find their reflection in the Chinese practice; for the few that do not apply to China, the main obstacles relate to the different cultural and social contexts.
Global Industry Research Team Report

Prepared by Rich Lysakowski, CENSA, and Bill Rhind, Pharmacia Corporation
Acronym Glossary

AIIM  Association for Information and Image Management
ARMA  Association of Records Managers and Administrators
CENSA  Collaborative Electronic Notebook Systems Association
CIPO  Canadian Intellectual Property Office
DoJ   U.S. Department of Justice
EMEA  European Medicines Evaluation Association
EPA   U.S. Environmental Protection Agency
ERLS  Electronic Records Life-cycle Specification
FDA   U.S. Food and Drug Administration
GALPs  Good Automated Laboratory Practices promulgated by EPA
GCPs  Good Clinical Practices promulgated by FDA
GERA  Global Electronic Records Association
GLPs  Good Laboratory Practices—promulgated by both EPA and FDA
GMPs  Good Manufacturing Practices promulgated by FDA
cGMPs  current Good Manufacturing Practices promulgated by FDA
GxPs  refers collectively to GCPs, GLPs, GMPs
InterPARES  International Research on Permanent Authentic Records in Electronic Systems
IT    Information Technology
JPO   Japanese Patent and Trademark Office
LAGER Legal Acceptability Guide for Electronic Records
NARA  U.S. National Archives and Records Administration
NHPRC National Historical Publications and Records Commission
NRC   Nuclear Regulatory Commission
OECD  European Office of Economic and Cooperative Development
OJT   on-the-job training
OMB   U.S. White House Office of Management and Budget
Predicate Rule a regulatory agency rule governing business area(s) to which a subsequent rule applies (e.g., Good Manufacturing Practices take precedence over US FDA 21 CFR Part 11)
PTO   U.S. Patent and Trademark Office
QERPs Quality Electronic Records Practices
RIM   Records and Information Management
SAA   Society of American Archivists
SLA   Special Libraries Association
UBC  University of British Columbia
UCLA  University of California Los Angeles
Executive Summary

Scope of this Report
This report looks at the situation of the global industry companies in the United States, Canada, Europe, and Australasia. Our discussion of the various record contexts defined by the InterPARES Project generally relates to the regulated industries where intellectual property protection (patents, trademarks, copyrights, trade secrets, etc.) and regulatory compliance (high-consequence health, safety, and potential liability records) are important. The analysis of the individual InterPARES principles focuses primarily on the corporate or R&D records of management and archival groups within the global companies. We reviewed the processes of continuum and life-cycle approaches of records creation, indexing, management, appraisal, preservation, long-term access, and reuse as these practices exist within the industries we surveyed. Our surveys were informal and based largely upon the companies within the Collaborative Electronic Notebook Systems Association (CENSA) and others in the same industries, not just CENSA association members. Specific company identities and survey results are confidential and results and conclusions were made anonymous in this summary report.

For the Global Industry Research Team’s work, the initial findings of InterPARES Project 1 relating to authenticity, appraisal, and preservation (as they became available) were constantly assessed against pre-existing juridical (legal), administrative (regulatory), provenancial/procedural/documentary, and technological contexts that are relatively uniform throughout the developed world.

We studied the applicability of the InterPARES 1 findings in more than twelve global industries, including high technology, software, pharmaceutical, chemical, biotechnology, hospital products, medical devices, clinical diagnostics, food, beverage, nutritional, and oil and gas. They all have common ties within the context of R&D record keeping. They also all have an environmental impact or a direct impact on human health, safety, comfort, and cleanliness, and quality of human or animal life (summarized by the slogan “better living through chemistry”).

The State of the Art at the Beginning of InterPARES
The Global Industry Research Team needed requirements, tools, techniques, and strategies from InterPARES to assess authentic electronic records preservation capabilities of current and future systems. The global industry team’s primary focus was on very high commercial value records for intellectual property protection, demonstrating regulatory compliance, and historically significant records. However, when we started, no good examples existed of any organization using fully electronic records systems—where no paper was used.

Progress of the Global Industry Team Research
In our investigations, we first spent considerable time understanding the rigorous theoretical background of the global library and archival science communities. We were already familiar with contemporary records management as practised in U.S.- and European-headquartered global companies. We also had to spend a significant amount of time understanding how to adapt and apply the new concepts of contemporary archival diplomacy as initially defined by the University of British Columbia group, and modified and codified by the InterPARES 1 group.

Within approximately eighteen months we were able to use the concepts and terminology of contemporary archival diplomacy because we were knowledgeable enough about the new concepts, vocabulary, and processes required for assessing authenticity, doing appraisal, and long-term preservation and access. However, we could not use the formal knowledge from the detailed process models for appraisal and preservation in initial assessments because they were not close enough to completion until the third year of the InterPARES Project.
The global industry team would have liked to do case studies within industrial settings, which would have provided the InterPARES Project some “real-world” examples of industrial systems. However, we found it was simply not possible to find able and willing test sites to be subjects for the case studies. This is despite the fact that CENSA was willing to completely fund the first case study as carried out by InterPARES-trained graduate students with the assistance of Bill Rhind, Rich Lysakowski, or other trained InterPARES investigators. Initial estimates of the time it took academic and government archives to do case studies ran from at least two full weeks to in some case several person-months of work. Frankly, this scared industrial companies and the return on investment was perceived as very small.

Even when we had refined the case study analysis tools to require only approximately two days of industrial company contact time, the time constraints of this amount of interviewing time were prohibitively expensive given the other responsibilities. Rich Lysakowski made serious requests at least four or five times at CENSA meetings for volunteer companies to step forward and offer case study sites. Lysakowski also drafted a “case study marketing brochure” for the case studies in order to outline “What’s In It For Me? (WIIFM)” if an industrial company in CENSA wanted to do a case study. The benefits of the case studies appeared to be simply too small to justify the investment of time, when most of the companies already knew they did not have formal record-keeping systems as defined by InterPARES or CENSA. The answers expected from the InterPARES case studies were in most cases already known as a result of comparison with the more than twenty-six case studies done by academic and governmental institutions.

**Executive Summary Conclusions**

From extensive surveys within and outside the CENSA membership, we concluded that *in virtually all cases, global industrial companies are not doing fully electronic record keeping where no paper is required*. Fewer than 5 percent of our members claim to be using fully electronic records where no paper is required. The 5 percent that are doing so are in non-regulated areas of the companies. The rest of the industrial population relies on paper, microfilm, or optical image storage for archiving records. In all cases, “hybrid systems” are still being used for the most valuable or consequential records. We found this to be true across the majority of regulated industry, in Europe, Australia, Canada, and the United States.

In general, industry exhibits enormous levels of ignorance and confusion about the formal concepts and definitions of contemporary archival diplomatics or archival science. Most key issues are understood and articulated in colloquial terms, but not formal, rigorous, or consistent terms. All except two of the corporations in the more than fourteen industries we studied did not even employ a corporate archivist, and those two do so only because they are about one hundred and fifty years old and have a large number of physical artifacts of historical significance to the company. These corporate archivists are beginning to worry about what they need to do and are going to do about electronic records.

**Benefits to the Global Industrial Team from InterPARES 1 Project**

We found the knowledge from the InterPARES deliverables to be very useful in educating industry and even the government regulatory and administrative agencies (besides NARA) on the proper usage of concepts and terminology. The InterPARES deliverables offer a preciseness and rigour that clearly distinguish the concepts and terminology from an archival perspective; they also define what is meant by the legal perspectives quite clearly, and help to eradicate the confusion of terminology that information technology has introduced into the general technical vocabulary (e.g., a record as a legal document, rather than as “database object” or “collection of database fields”).
We incorporated our findings from the InterPARES Project into conference presentations, CENSA documents, specifications, and papers in the various scientific and technical literature, and the Legal Acceptability Guide for Electronic Records (LAGER). The production of the LAGER would have been very difficult had we not acquired a lot of the rigorous foundations from InterPARES 1 Project deliverables and especially Heather MacNeil’s book, which served to document many of the key diplomatic principles succinctly for laypersons.\(^1\) InterPARES provided the formality of terms, concepts, and processes that were necessary for people to think more completely and accurately about what needs to be done for long-term preservation and access to authenticated records. These included the InterPARES glossary, the Requirements for Assessing and Maintaining the Authenticity of Electronic Records, the initial and refined Template for Analysis, the case study analysis tools (CSIP and TEDGI), the many case studies done in other non-industrial domains, and other deliverables from the project.

Our own case study—using the InterPARES methodology for the current paper notebook-keeping process and its extension into future electronic record books—helped us formalize and document our understanding of the process. It also gave InterPARES a concrete high-value example of an electronic system that is used to assemble documents, and creates records from them, but in nearly all cases does not keep them in authentic electronic form for their full retention period. The final electronic data are assembled and either bound in a book or printed to special security paper, signed and date-stamped by an author and witness, and then managed the traditional way, although their location and contents may be tracked electronically within a commercial database or, typically, one developed in-house.

We hosted a special “Legal and Regulatory Symposium” at a CENSA quarterly meeting where Heather MacNeil from UBC/InterPARES attended as an academic expert, and was available to provide commentary on a “mock trial” that we held at the symposium to illustrate the state of unreadiness of industry to deal with the complexities of fully electronic records systems.

As a result of InterPARES participation and the deliverables being shared with these global industrial companies via the various meetings and conferences that CENSA is involved in, the level of ignorance or misunderstanding of terms is beginning to drop sharply for those CENSA member companies that participate. Knowledge is diffusing rapidly out from CENSA members to industry at large. This is an important first step that will facilitate the introduction and usage of the remaining output from InterPARES 1 and InterPARES 2.

True electronic records management systems as defined by the CSIP and TEDGI are simply not found outside of industries that are not heavily regulated. Although many companies have employed electronic document management systems (EDMS), most corporations still do not use commercially available electronic records management (ERM) software systems. Some EDMS systems vendors have used third-party EDMS vendors’ software (PC-DOCS, iRIMS/OpenText, Documentum, Provenance, and others) and ported their ERM application on top of the EDMS and had it DoD 5015.2 certified.

Many EDMS systems in the pharmaceutical industry have been rigorously validated against a predefined system functional requirements document. This includes a rigorous internal audit of the vendor’s internal development and documentation practices. Such validation is an expensive process, estimated by the Society of Quality Assurance to cost more than U.S.$300,000 when all labour and documentation costs are included. Once validated against a rigorous requirements set, such as those produced by CENSA or the U.S. Department of Defense, these systems have been found to be good for the creation and online maintenance of records. However, these validated or certified systems still do not archive records in the formal sense (by achieving long-term preservation and access regardless of the underlying technology changes). Most frequently documentary records are either printed from these systems and signed and witnessed, or the plan is to simply keep “the records” online indefinitely until a good long-term archiving technology

system product will be developed. There is significant risk inherent in this strategy, because it does not follow the findings of planning for preservation before creating records.

In general, all the industries we studied are heavily regulated by the U.S. Food and Drug Administration, the Environmental Protection Agency, the Nuclear Regulatory Commission, or their international counterparts in Canada, Europe, Japan, and elsewhere. We studied the processes for assembly and submission of product licensing applications (e.g., new drug applications, new medical applications, new pesticide applications) to the FDA, EPA, and Health Canada. The companies we studied are impacted heavily by the records creation, records management, and records retention requirements specified by these regulatory agencies.

We also studied records systems in the same companies from the perspective of submission of patent and trademark applications to government administrative agencies such as the U.S. Patent and Trademark Office (PTO), the World Intellectual Property Organization (WIPO), or the European Patent Office (EPO). All companies considered are impacted, and held back by their desire to protect and sell their products in the U.S. marketplace first (by far the largest and most profitable market in the world for most of their products). As a result, European and Asian divisions of these companies still feel bound by the U.S. rules of “first to invent” that is the patent approval practice of the U.S. Patent Office, rather than the “first to file” rules of the rest of the world. This means that the original records of invention (usually held on paper) must be kept and produced in the case of a patent interference or patent infringement.

Global Industrial Team Findings

We found in our investigations that, except for a few rare exceptions, there are still too many things missing in many contexts (juridical context, regulatory and administrative law, technological context and others) that negatively affect the implementation of the InterPARES Project’s findings regarding long-term preservation and access right now. There is nothing wrong with the InterPARES 1 findings. In fact, they will be very useful for influencing and eventually making or lobbying for specific changes of these contexts themselves. If we had had the full InterPARES results many years ago, the pharmaceutical and chemical industries could have saved billions of dollars in complying with regulations that are now or soon to be law and ignore many basic archival science principles and findings of InterPARES. However, these are now sunken costs that cannot be recovered. At best, we can learn from the mistakes of the past.

One important overall conclusion is that nothing in the InterPARES intellectual framework is at odds with the various contextual aspects affecting record keeping and archiving in global industry. However, currently there are significant gaps between government and industrial policies, laws, regulations, and practices (which are largely inconsistent with InterPARES requirements) and the state of the art of technology and program implementation—as implemented “on the ground” in current in-house technology systems. This means that implementation of InterPARES requirements will be possible only once some remedial measures have been taken by getting regulatory or administrative law changes or additional guidance from key government agencies. For example, 1) we must get additional guidance or implementation guidelines from the relevant administrative or regulatory agencies (e.g., PTO, OMB, OECD, EPA, and EMEA), and 2) where they are at odds with the InterPARES intellectual framework, we must get changes made in various laws or regulations to achieve or allow consistency with the InterPARES framework. Knowledge gained in InterPARES and via our own observations has clarified the best avenues to pursue to start making such changes.

There is nothing in the InterPARES model of the preservation process that industry either doesn’t already do in best cases or could not begin doing across the board, though this will take many years to be commonplace. In general the work of the global industry team in InterPARES has focused on records of value for the purposes of intellectual property protection and regulatory compliance, but the principles will apply equally well to records for product liability, commercial or civil litigation, or preserving corporate or organizational history.
One important realization that the Global Industry Team working on the Legal Acceptability Guide for Electronic Records made was that the organization must plan and budget for, and technologically be ready to convert the record to a long-term preservation format at the time of creation, or else it will have difficulty migrating records to future records formats. This plan and commitment must be done before the first record is created, or the organization will likely run into problems when it comes time to do migration sometime in the future.

The changes required to comply with the InterPARES requirements are starting to take place now as new government laws and regulations are introduced and implemented. For example,

- U.S. Government Paperwork Elimination Act (GPEA).
- U.S. White House’s Office of Management and Budget “E-sign Act”—The Electronic Signatures in Global and National Commerce Act, 15 U.S.C. Sections 7001–7031 covers electronic signatures for contracts and documents involving interstate and foreign commerce. In recent implementation guidance advanced by the U.S. Department of Justice, the records retention requirements are included as important to be considered when creating the records.
- Various “e-government” initiatives in the developed world (United Kingdom, Australia, New Zealand, Ireland, United States, and others).

The various e-government laws and initiatives are forcing the hand of government agencies to plan, budget for, and start to implement systems that permit citizen-to-government commerce.

As these new laws and regulations eventually take full effect across the private and public sectors, organizations are grappling to understand and build first-, second-, and even third-generation systems that will be capable of integrated creation, management, and archiving (long-term preservation and access) for authenticated electronic records.

The electronic records software systems marketplace is also beginning to recognize and address the needs of industry for long-term archival preservation and access for electronic records. The Australian government has helped to catalyze the creation of this software market by investing more than four million dollars in developing a conceptual model, software system functional specification for an entire electronic records archive, and an XML/PDF-based file format for permanent electronic records collections. Several commercial software companies and end-users (e.g., CSIRO) collaborated in Australia to build a prototype. The prototype was tested for four to six months, modified, and then put in place as the production version of The Electronic Records Archive for the Victoria State Archives. The Victoria government has declared and mandated that its XML/PDF-based file format for permanent electronic records collections will be relied upon for a minimum of one hundred years. This hundred-year long-term accessibility requirement was one of Victoria’s essential business requirements and design criteria for its system.

CENSA has decided recently that it will piggy-back onto this Victoria government standard approach for the near term, meaning at least the next twenty-five years. The exact approach for “piggy-backing” is still being decided upon, but we expect to finalize our technical strategy and announce it by the end of 2002. We are working with U.S. NARA and Australia’s Victoria State Archives to ensure that their long-term strategies are functionally compatible.

**Assessment of InterPARES Principles Against Industrial Contexts**

In this section, the InterPARES principles as defined in the report of the Strategy Task Force are first considered as a set, and assessed against the various contexts that define the record-keeping (creation, management, and preservation) regimes within regulated industries where
intellectual property protection is important. Where appropriate, individual principles are then assessed against those contexts as well.

In many cases we found that industry has identical concerns as government and academic institutions. However, in a few cases, the emphasis of concerns of industry mandate faster for less expensive solutions to an immediate problem, sometimes because of a regulatory or legal mandate with a deadline that industry must meet regardless of whether government or academia has the same deadline. In government and academia, it is not a primary concern to meet environmental, food, drug, or medical-device product licensing regulations, or achieve rock-solid intellectual property or patent protection. In the case of government or academia, other government or institutional acts or policies—such as right-to-privacy, freedom-of-information, the Government Paperwork Elimination Act (GPEA), or e-government initiatives—become primary concerns the institution must pay attention to implementing.

**Principles Overall—Findings in Critical Contexts**

**Juridical/Administrative Context**

This is the key context defining the possibilities for successful implementation of the InterPARES findings. The juridical and administrative elements must be dealt with separately, as their effects and impact are not uniform.

**Juridical Context**

U.S. information legislation (The U.S. courts’ Federal Rules of Evidence, the U.S. government Paperwork Reduction and Paperwork Elimination Acts, the E-Sign Act, Freedom of Information Act (FOIA), the Privacy Act, and regulatory rules by the FDA, EPA, NRC, PTO, and other agencies) provides a strong overall framework within which the InterPARES principles could be implemented.

A huge problem that exists is that U.S. regulatory agencies, the U.S. Code of Federal Regulations, and U.S. congressional code define records and electronic records to be any information in any format that is stored for later evidential, business, or historical purposes. They thus equate with records, all evidence or data of any type created by anyone anywhere within the business. They also do not associate records with the business processes they relate to, nor do they include the archival requirement of the record to be “fixed and set aside under the care of a qualified custodian with the responsibility of ensuring the ongoing authenticity of the record.” There is no measure of the quality of the evidence that defines whether something is a record, such as author’s and/or witness signature(s), or the custodial care or evidence about authenticity that is generated during custodial care.

Equating all data into “records” turns retrieval for litigation and criminal prosecution purposes into an evidentiary discovery “circus” whereby intellectual property attorneys, regulatory inspectors, product liability litigators, and other attorneys can treat any data or information as though they were all records created and stored within a record-keeping system, when in reality they are simply just evidence used to satisfy the data collection needs of the business. There is so much uncontrolled data/evidence floating all over corporations and government agencies that is it is nearly impossible to control it all right now. E-mail is the worst/best-case example that gets cited repeatedly as poorly managed evidence that the creators don’t consider as records, but litigators absolutely have a field day using.

The U.S. Federal Rules of Evidence provide specific guidance about how electronic records may be admitted as “hearsay” into court proceedings. There are industry practices and standards promulgated over the years by the Association of Records Managers and Administrators (ARMA), Association for Information and Image Management (AIIM), Society of American Archivists (SAA), and Special Libraries Association (SLA), but no firm single set of rules exists in the United
States like the Canadian and Australian governments’ Requirements for Management of IT, Security, and Record Keeping Metadata Requirements.

Because of this lack of common detailed guidance from either the U.S. executive branch of government (OMB), or detailed NARA directives, various administrative and regulatory agencies create their own regulations to fill in the gaps. This is not a fault of NARA, which has simply chosen to work at higher levels of policy setting and let individual regulatory and administrative agencies develop detailed directives that work best in their own agency. Some of these agency regulations are performance-based and less murky about specific requirements.

Because of the inconsistencies in scope, policies, standards, and knowledge across the set of regulations, laws, guidances, and recommendations issued by the OMB, NARA, the U.S. Department of Justice (DoJ), FDA, EPA, PTO, and other agencies, and the actual record management and archival practice in corporations—at least in the global industry team members represented in CENSA—CENSA had to create its own standard guides and specification and knowledge transfer symposia to educate at least eight audiences that explicitly have electronic records among their concerns and responsibilities. These audiences are:

- end users
- attorneys
- regulatory affairs
- quality assurance
- executive and R&D managers
- records managers
- archivists (where they exist)
- information technologists.

Other functions within the corporations are just now beginning to be examined (administrative, financial, personnel, and other types of records).

**Administrative Context**

The overall administrative context of industrial work is conducive to the implementation of the InterPARES findings, given the presence of record management policies of NARA, though they do not apply directly to industry. Nearly all government regulatory agencies have recommended retention schedules. Most industries have specific records retention schedules imposed upon them by industry regulators, or they have general guidelines on how to determine these retention schedules. Good businesses have implemented and regularly update their retention schedules to achieve better business record keeping in support of managing their intellectual asset portfolios.

Other than the FDA and EPA regulated industries that implement the Good Laboratory Practices, Good Manufacturing Practices, or Good Clinical Practices, and the Nuclear Regulatory Commission (NRC) Recordkeeping Guidelines, global industry does not currently have effective standards for the entire records management and archival regimes. Many initiatives and policy creation efforts are now under way to address this problem, but it will be many months to years before widespread creation occurs for management policies, information technology, and standards operating procedures (SOPs) to support a general adherence to the principles enunciated in the InterPARES framework. That is one reason the InterPARES work is so important to CENSA from both industrial and international governmental perspective.

CENSA and the Global Electronic Records Association (GERA) wish to base their standards, guides, and tools upon the most rigorous and best time-tested principles of archival science and records management. The worst that could happen is that uninformed and ignorant regulatory or administrative agencies promulgate inconsistent laws and regulations without this rigorous knowledge, resulting in massive efforts to interpret what poorly written regulations really mean, filling in the gaps to include what they think is missing—without direct specification and input from the regulatory or administrative agency—and then building systems that may or may not comply with a rule that was poorly written in the first place. The net negative economic result is that it costs business hundreds of millions or billions of dollars to comply with such rules, and the costs get passed along to either consumers or taxpayers.
In order to address the needs for rigorous, consistent tools for electronic records program implementers “on the ground” (not simply corporate-level policy makers), CENSA spun off the non-profit research institute GERA. Its mandate includes facilitating interaction and collaborative true research, authoring, and standards work with government regulatory and administrative agencies, academia, and non-profit organizations that are stakeholders in establishing quality electronic records programs. The purpose of this institute for electronic records is to produce standards called the Quality Electronic Records Practices (QERPs) to assist implementers in knowing the detailed requirements (we call them “80/20 blueprints or templates”) of what electronic records programs, systems, and components to build, as well as how to validate and audit them; and legal defence procedures and case law with recommendations on how to effectively litigate using electronic records. The “core” deliverables include professional education and training materials to teach professionals about the QERPs. In addition, the goal is to put these deliverables in an online, integrated knowledge base that supports records managers, archivists, cyberterrorist response teams, and others who must use electronic records over any stage of their life cycle. We will continue to update and feed key knowledge from InterPARES and other sources into these standards as it becomes available, fully acknowledging the origin as we integrate content from InterPARES and other sources. The work of GERA centres on the R&D, creation, publication, and dissemination of these standards (and maintenance when the time comes.)

GERA (via a jointly funded CENSA/NHPRC project) is developing a general-purpose life-cycle standard that includes a framework model covering all basic program design, policy, procedural, personnel, technological, and other programmatic and technological aspects of QERPs across government agencies and industries. The QERPs are a set of integrated documents that include how to validate and litigate using records drawn from a system implemented against the requirements identified in the Electronic Records Standard Lifecycle Specification. However, it will take at least three more years from the time of writing before all of the standards are done. The e-Records Lifecycle Specification (ERLS), the Validation Guide for Electronic Records Systems, and the Legal Acceptability Guide for Electronic Records (LAGER) were to be finished by October 2002. The Auditing Guide for Electronic Records; the Certification Guide for Electronic Records Programs, Systems, and Archivists; and some other companion guides and tools were to be created after October 2002 as funding is secured from various federal and corporate sources. These QERPs deliverables—as they are applied—will help to prove out the concepts, terminologies, and process models defined in InterPARES 1.

As a point of emphasis, at the time this report was written the global industry team badly needed InterPARES 2 to produce an integrated life-cycle model spanning the entire electronic records continuum of electronic records creation, appraisal, maintenance, long-term preservation, access, and reuse. It was especially hoped that this would be ready in time for version 1.0 of the ERLS global standard. If the integrated InterPARES model was not ready by mid-2002, we planned to use what was available as part of the ERLS draft global standard.

**Provenancial/Procedural/Documentary Contexts**

In regulated industries, the provenancial context is established by default, rather than by design, as a result of the implementing regulations and complying with law, for example, complying with FDA 21 CFR Part 11 on electronic records and signatures. The EPA Good Automated Laboratory Practices, the FDA current Good Laboratory, Clinical, and Manufacturing Practices.

Detailed documented procedures or performance criteria must be specified and tested before putting data, information management, and record-keeping systems for labs, clinics, and manufacturing processes into “production” usage. The systems must be “validated” to ensure compliance with the predicate rules of the agency. In these cases, the procedural context is linked to the record creation process that is linked closely to the business functions it supports. This appears to establish the *provenancial/procedural/documentary contexts*. The White House’s Office of Management and Budget—The Electronic Signatures in Global and National Commerce Act, 15 U.S.C. Sections 7001–7031 “The E-Sign Act” covers electronic
signatures for contracts and documents involving interstate and foreign commerce. In recent E-Sign "Implementation Guidance" advanced by the U.S. Department of Justice, records retention requirements are clearly spelled out to be included as important when creating the records, and the guidance makes it quite clear that it does not just cover e-commerce records but all records resulting from government transactions with citizens and businesses. The implementation of E-Sign will affect all government agencies and their interactions with industrial companies, especially regulatory and administrative agencies, and commercial transactions.

**Technological Context**

A technological context appropriate to the realization of InterPARES principles exists (the many commercially available records, documents and information management system products—an integrated suite of softwares available to all corporate, government departments and agencies). Many of these systems have been implemented and validated, but these are not formal record-keeping systems because they are missing the capacity for long-term record keeping. These commercial systems' records and information management (RIM) needs to be more fully implemented across corporate and government enterprises. Our survey of thirty-eight corporations and at least sixty-five federal agencies indicate that EDMS systems are starting to be implemented widely, but in nearly every case, records for archiving are being kept on paper.

Separation between the information technology and the records, and RIM responsibilities and reporting structures within global corporate institutions, has caused many problems for records managers and archivists. This frequently leads to RIM (especially archival considerations) not being inadequately reflected in IT implementations. This also raises a more important concern: the view that records managers or archivists are filing clerks, and should not be considered in the design of all new IT systems. This is a huge concern, particularly with what are proposed are “the full implementations of RDIMS [records, documents, and information management systems].”

The only government that has accomplished the research, engineering, construction, prototype, fixing problem, validation, and production implementation of technology systems for full electronic records management and archiving (ERMA) systems is the Australian government with the Victoria Electronic Records Strategy (VERS). They are also constructing a policy framework to permit the cross-government users of the VERS software, which will round out and complete the fully electronic records archiving system, by providing the requisite policy and program guidance. Various other state governments within Australia have produced policy frameworks that support full implementation of electronic records programs and systems.

**Individual Principles Extracted from the Research**

We generally concur with the Canadian team’s principles extracted from the research. We made some minor modifications to fit some broader, immediate needs for quality electronic records practices.

However, in general, any records preservation policy, strategy, or standard should:

- address records specifically rather than digital objects generally.

This principle is not explicitly reflected in the overall global industry or government context, but is directly supported by various local agency rules such as FDA 21 CFR Part 11, SEC rules, EPA GLPs and GALPs, OECD Rules for Electronic Records. In the United States, the most comprehensive support is the Government Paperwork Elimination Act (GPEA) and the E-Sign Rule, which will have sweeping effects throughout industry and government.

- focus on authentic electronic records.

Several industrial initiatives are being driven by regulatory and legal concerns on product and environmental quality and liability. The long-term preservation and access to electronic records is emerging as a by-product of new industry regulations, e-government initiatives, and last, but not least, the need to improve business productivity overall in all types of businesses, including
government, industrial, and academic. Regulatory agencies needed to speed up the product application safety review process, which could take as long as seven to ten years, cutting years off the useful patent protection period at the end of a product’s life.

This resulted in regulatory agencies starting to specify electronic records business requirements and regulations, but without the restraint that comes from informed knowledge of the current state of the art of information technology and the total lack of design of IT systems to meet even near-term (25+ years) records retention requirements. Thus the agencies specified the full length of time of product research and development studies plus the full amount of time that products are sold on the market.

Among the best-known U.S. industry regulations that affect industrial companies worldwide that wish to sell their products in the United States is the FDA 21 Code of Federal Regulation (21 CFR Part 11) rule on electronic signatures and electronic records.

Originally, the pharmaceutical industry asked the US FDA for guidance and regulations that would permit fully electronic submission and review processes in lieu of paper applications processing. The FDA's response is that it also wanted to be able to inspect original data and records to ensure that required quality levels are met. This requires companies to retain all supporting records and source (raw) data and make them available to inspectors during audits. The FDA equated all data to records, whether or not they were of a quality capable of being preserved or accessed in authentic form for long periods of time, and whether or not they were given to a trusted party with a duty of care for preservation.

In fact, though the FDA 21 CFR Part 11 rule is about retention of authenticated electronic evidence, it is written in a language that emphasizes “electronic signatures” and “electronic records.” The FDA took licence and mistakenly equated all computerized data to “records” as defined by the U.S. Congressional Code (44 U.S.C. 3301), which reads “All books, papers, maps, photographs, machine readable materials, or other materials, regardless of physical form or characteristics, made or received by an agency of the United States under federal law in connection with the transaction of public business and preserved or for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, or other activities of the Government or because of the value of the data in them.” Building upon this U.S. Congressional Code was the U.S. Code of Federal Regulations 36 CFR 1234.1, which defines electronic record as “Information on any electronic medium that can be read by using a computer or any other electro-magnetic device that satisfies the definition of a Federal record in 44 U.S.C. 3301,” which in turn served to equate all electronic evidence to electronic records. In the broadest sense this may be true, because any data may serve as a record and be used as evidence in court or for historical purposes; in practical everyday terms, however, it erroneously gives data and record creators the duty of care for long-term preservation and access normally given to trained records managers and archivists. Most data and record creators are improperly trained, and are not given the responsibility, budgets, or the job spans to take on the long-term duty of care for their data and records. Strictly speaking, even if records creators were given responsibility for long-term care, this violates the principle of good accounting practices that one should give responsibility for records maintenance and auditing to a trusted third party. Thus, a few oversights made in this historic rule that affected FDA-regulated industries.

The archival community and records management was not given the opportunity to intervene, or even have significant influence to develop criteria or standards for archival electronic records that include “capable of being preserved in authentic form for as long as needed by the business.” There was also the practical matter that not enough technology standards existed yet to make long-term preservation via migration mechanisms facile or straightforward.

FDA-regulated companies were unsuccessful in presenting an effective financial justification case against the FDA demonstrating that the FDA rules, as written, simply could not be implemented without extreme cost and undue burden. FDA-regulated industry spent several years just to interpret the rule, and is still working hard to comply. The FDA Part 11 rule simply left too many details of records management and archival science unclearly specified or simply unstated. From
the time the rule was proposed (early 1992) until it was finally implemented as law (early 1997), industry and government were just beginning to understand the full cost of implementing basic electronic records creation, records management, and the implications of long-term records preservation and access.

Many people in industry believe the FDA regulation simply violated the Government Paperwork Elimination Act (GPEA)'s goals for reducing paperwork burdens, and that it caused undue financial burden on the companies forced to comply with record-keeping requirements. However, the FDA-regulated industries failed to respond fast enough or early enough with compelling financial research results to show to the contrary that, while the goals to accelerate electronic submission and review process requirements were reasonable and achieved, the records retention burdens required retrofitting or in many case total replacements of otherwise modern equipment. By late 2001, it was estimated that the 21 CFR Part 11 compliance costs for large FDA-regulated companies was already between US$100 million and US$200 million per company. This translates into billions of dollars industry-wide, just to get compliant, and then additional costs for maintaining more complex record-keeping systems, in addition to information management systems used to gain competitive advantage.

NARA guidance on electronic signatures and electronic records retention for GPEA purposes did not appear until October 2001, and by then the pharmaceutical industry had already spent huge sums of money to comply with FDA 21 CFR Part 11. It was not until 25 September 2000 that the U.S. Department of Justice provided guidance to help with all federal agencies' interpretation on GPEA with regards to long-term electronic commerce transactions.

The U.S. Patent and Trademark Office is a branch of the Department of Commerce. The Patent and Trademark Office applications can be viewed as long-standing commercial licensing application transactions (records) that are covered under E-Sign.

The recently proposed EPA Cross-Media Reporting and Recordkeeping Regulation (CROMMERR) may also be in danger of trying to do too much too fast, trying to combine GPEA reporting requirements, while at the same time imposing internal records retention requirements on the entities that it regulates (estimated to be between 1.2 million to 1.7 million entities.) The public comment period for this rule ended 27 February 2002. Many companies in industry have asked that the CROMMERR rule be withdrawn and evaluated against the current state of readiness of regulated entities and technologies to actually implement the rule.

By 2002, the Japanese Patent Office, the European Patent Office, and the U.S. Patent and Trademark Office (PTO) online electronic submissions processes had been in place for more than ten, exactly five, and less than one year, respectively. All of these patent offices recognize that they do not yet have supportable and supported long-term retention and access systems for fully electronic records, even though they accept the submissions in fully electronic form, complete with the patent agent's digital signature. Patent and trademark submission regulations cover all parts of the submission form. By 2003, the PTO wanted to have a fully electronic application submission process; it has already achieved that goal. The challenge remains on the back-end long-term records retention and access systems that the PTO will use in-house.

Challenges exist for both the sending and receiving parties involved in the patent application submission process. For example, the PTO is required to retain submissions online as records, in secure and confidential format, for a minimum of forty years, before final disposition occurs, where the records are either submitted to NARA or destroyed. NARA has admitted openly that it does not know how it will meet these retention and access requirements, but that it is working on a solution. On the other hand, the submitting organization must keep records from the moment of invention and reduction to practice that can go back as far as two decades earlier. Thus, the submitting organization in an FDA-regulated industry must preserve evidentiary records of invention and reduction to practice that may need to be kept for as long as twenty-seven years or longer. Drugs can be granted up to twenty years of patent protection; some records of R&D must be kept for seven years beyond the last date of manufacture for pharmaceutical products. For medical devices embedded in a patient at an early age, the retention periods can last from the
time of invention, and then for the life of the patient and several years beyond in case of class-
action product liability lawsuits, potentially many decades.

The FDA, PTO, and EPA initiatives are being watched closely by equivalent agencies outside the
United States because frequently these agencies set the pace for product safety and
environmental protection in many other parts of the world, and the United States represents such
a profitable market that many companies outside the country want to meet FDA or EPA
requirements in order to be able to sell products within the United States.

Overall, the problems of long-term electronic record preservation and access remains unsolved
for both parties involved in submitting and receiving patent applications, and in trying to defend or
attack intellectual property rights if infringement or interferences occur; those entities submitting
or receiving product licensing applications; or those demonstrating that they “did the right thing”
with potentially hazardous chemical or biological materials.

- recognize and provide for the fact that authenticity is most at risk when records are
  transmitted across space or time.

This fact is certainly recognized in the polices of the PTO, FDA, EPA, Canada’s public key
infrastructure (PKI), and other regulations worldwide. However, appropriate PKI initiatives are
underway to help guarantee the transmission of records between agencies and to national
archives in many countries. Industry will be forced to follow the lead of local national governments
in adopting the rules and practices for digital signatures, identity cards, and public key
infrastructures. There is still much resistance in the United States and elsewhere to universal
electronic identity cards that can be used for electronic signatures.

- recognize that preservation of authentic electronic records is a continuous process that
  begins with the process of records creation.

This requirement is reflected in the record-keeping guidance that CENSA provided to its
members in the “Legal Acceptability Guide for Electronic Records.” This view is consistent
throughout global industry, the National Archives of Canada, NARA, Australia, and elsewhere.
This unites the life-cycle and continuum viewpoints of electronic records, rather than invalidating
either one. It also provides guidance to government and industry that govern the transfer of
electronic records to the custody of an institution’s records managers or archivists (commonly
called “terms and conditions of transfer”).

Industry eventually came to realize that records must be converted at the time of creation to a
format that facilitates preservation—but only after observing the inability of many projects to
access “records” in systems that were never designed for preservation in the first place. An
enormous expense is required if you do not plan for preservation before you create the record.
We found a significant number of cases where IT groups designed systems or vendors sold
systems that did not provide sufficient documentation of file formats, software interfaces, or
export processes to permit the migration of critical assets. All too often this fact was not
discovered until a system was obsolete, really ready for the graveyard, and its records were
needed for pending litigation. The result is that a lot of money and time were spent resurrecting
systems to extract their records and then trying to prove those records’ authenticity.

The Global Industry Research Team findings strongly support this recognition that one does best
to plan for preservation before the first record is created. CENSA has been promulgating this
view through its many meetings with industry software implementers.

- be based on the . . . concepts of a trusted record-keeping system and the role of the
  preserver as a trusted custodian.

CENSA and its global industry members achieved consensus that “custodianship” is the central
theoretical concept that fulfils the requirement for a trusted record-keeping system, but actual
current implementations routinely fall short of fully ensuring records’ authenticity—sometimes
because records are left under the control of their creators, sometimes because trained electronic
records managers and archivists are not consulted when systems expected to keep records are
being specified and procured. These gaps will be addressed in the coming years with increased training throughout industry, government, and the archival and records management communities. The role of trusted custodianship will grow as organizations learn more about it, but also experience failures and losses from not implementing it properly.

- be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object: it is only possible to preserve the ability to reproduce the record.

The InterPARES researchers fully understand this principle, and its preservation strategy focuses on maintaining the ability to reproduce records in authentic form (while still accounting for the need to maintain the various physical carriers on which records are stored over time).

- recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form.

The former point is generally recognized among those who deal with electronic records, but it is highly unlikely that the diplomatic concept of “element of documentary form,” per se, will figure in industrial or government policies, strategies, or standards unless or until these diplomatic concepts become commonplace via more widespread adoption and education.

- specify the requirements a copy of a record should satisfy to be considered equivalent to an original.

In the case of the industry and the court systems with which industry deals, the requirements to be satisfied will relate primarily not to the copy itself but to the system used to transfer, store, and maintain the record’s digital components and to reproduce the record. But the system used will only be selected once it has demonstrated that it can be used to reproduce accurately both the content and documentary form of the record. Courts must evolve their preference for paper as the “best original form” to include reproduced authentic electronic records as the best original form.

- integrate records appraisal in the continuous process of preservation.

Appraisal for both economic or historical value is done formally in only the best industrial organizations. Appraisal methodologies must play an increasingly important role in industrial records and archives management. This would be facilitated if somehow processes for record classification, archival descriptions, and appraisal could become more closely integrated and more automated, because archival description and appraisal are all time-consuming processes. Again, these must be viewed as continuous processes that are integrated into a continuum operating model for preservation.

- explicitly recognize that the traditional principle that records all relied upon in the usual and ordinary course of business can be presumed to be authentic needs to be supplemented . . . by evidence that the records have not been inappropriately altered.

The information management policies do not always explicitly include such recognition, nor recognition of the purpose and value of audit trails in proving both reliability and ongoing authenticity of electronic records. It must be put into system and policy requirements that archiving audit trail records (in preservable, authentic, accessible form, of course) should become an institutional practice to document the continuing authenticity of the institution’s electronic record collections over time.

- recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records.

An archival group must assess the authenticity of records before they are transferred, and be responsible for maintenance of their authenticity as a critical part of preservation activities following transfer into the archive.

- draw a clear distinction between the preservation of authenticity . . . and . . . authentication.
It is not certain that this distinction is well understood in industry or government generally, but it is recognized by the groups participating in InterPARES. It is reflected in some local country guidelines for records created in a PKI environment, but not globally in countries or companies that do not rely on PKI or electronic national identity cards yet.

**Effects on Stakeholders**

*In short, the primary benefit will be that meeting the benchmark authenticity requirements will substantially improve the quality of record-keeping processes throughout government and industrial institutions, whether local, state, federal, or global.* The principal direct effect of the implementation of the InterPARES requirements in the global industry context will be much stronger archival records in electronic form, with a more explicit focus on determining the reliability and authenticity of the records being created and transferred to the corporate archival repositories, and greater confidence by companies in asserting their authenticity when reproduced. The principal indirect effects will be considerable, both industry-wide and government-wide. The costs and the benefits of implementation of the record-keeping infrastructure necessary to meet the benchmark authenticity requirements must be better understood. Right now these are understood qualitatively as “better current and future records, improved accountability, and increased access and reuse of records.” Some administrations, agencies, or corporations that wish to keep their operations covert view this heightened quality and accountability as a liability. However, this must not impede those entities with a desire for better quality and accountability.

**Conclusion**

Everything in the InterPARES principles can be incorporated into policies, strategies and standards at the international level at which the global industry team works. There are still significant administrative, legal, regulatory, and technological barriers, but these are not insurmountable. However, it will require completion of the InterPARES life cycle or continuum process models and their incorporation into the Quality Electronic Records Practices and other standards, their worldwide promulgation, followed by extensive training, implementation, and certification. Many of the remaining administrative, legal, regulatory, and technological barriers will need to be removed by legal, regulatory, government administrative law, and technology market development processes. These activities are simply outside the scope of an international collaborative research project. InterPARES should retain its focus on research, codification, and documentation of key concepts, definitions, and processes from an integrated viewpoint of records management and archival science for electronic records.

In general, the current situation and “digital gap” is because of the poor understanding of formal records management and archival science by the masses, the subservient position in which records managers and archivists are put relative to information technologists, and their strong dependence on information technologists for system implementation. There is also a serious lack of training of information technologists and computer scientists to include “design for preservation” for authenticated information as an important design centre for current and future systems. Dependencies on IT may force records managers and archivists to simultaneously become “IT-savvy,” become good team players and negotiators on IT systems requirements analysis and specification teams, diplomatically bring the attorneys to their rescue when necessary, become good electronic records systems project implementation managers, and become good at financial justification and business management. They must become truly multidisciplinary themselves. This is not unique to any one country, but is a worldwide phenomenon. As the need for “bridging the digital gap” continually grows, multi- or cross-disciplinary professionals will emerge via on-the-job training or via formal university or post-professional training and certification programs. This process may take another five years to complete, because it is truly required to complete the paradigm shift from paper to electronic records.
Italian Research Team Report

Prepared by Maria Guerico, Italian Research Team
This report analyzes the main findings of the InterPARES Project as expressed and summarized in the “Strategy Task Force Report” and their impact within the Italian juridical system and the archival national tradition. Specific attention will be dedicated to the principles, criteria, and standards identified as relevant in the course of the InterPARES Project.

- address records specifically rather than digital objects generally; that is, it should address documents made or received and set aside in the course of practical activity.

This is a principle already established and in place within the Italian juridical system. New legislation (decrees of president of the republic 445/2001) has approved a general definition of record and electronic record whose basic concepts are consistent with the research findings. With reference to this aspect, the main issue should concern the effort to create a sufficient degree of awareness within the public administrations regarding the specific requirements connected to the electronic records identification and to management for their long-term preservation.

- focus on authentic electronic records.

Many aspects of the criteria described in the report should require further detailed analysis both in relation to the national archival tradition and juridical system and the ways to implement them in the electronic record-keeping system.

A basic tool for guaranteeing a correct and adequate implementation of the principles and the criteria here expressed (baseline and benchmark requirements) already exists in the Italian jurisdiction in the form of a manual of procedures where each creator should define and describe the elements and activities necessary for the maintenance of the electronic records' authenticity.

- recognize and provide for the fact that authenticity is most at risk when records are transmitted across space (i.e., when sent between persons, systems, or applications) or time (i.e., either when they are stored offline, or when the hardware or software used to process, communicate, or maintain them is upgraded or replaced).

The transmission across space is seen as a risk in the new Italian and European legislation, while transmission across time is not seen as a problem. Detailed and complex rules for authenticating electronic records to be transmitted are in place (rules on electronic and digital signature). A very low level of awareness exists with reference to the transmission across time. With this respect the Italian legislation is too generic.

- recognize that preservation of authentic electronic records is a continuous process that begins with the process of records creation and whose purpose is to transmit authentic records across time and space.

In the Italian context this principle is clearly established at the level of the archival tradition and it has been renewed in the recent archival theory developments and also confirmed in the new legislation on electronic records keeping in the public sector (dpr 445/2000): standard and rules for the electronic record-keeping system have been approved with reference to the registry system and digital classification. Best practices are already in place and presented on the Web by the authority for information technology in the public administration (<http://www.aipa.it>) and by the Scuola superiore della pubblica amministrazione (<http://www.sspa.it>).

- be based on the concept of trust in records keeping and record preservation and specifically on the concepts of a trusted record-keeping system and the role of the preserver as a trusted custodian.

This is a new concept in the Italian tradition, specifically with reference to the necessity for the preserver to maintain the possibilities for verifying the authenticity of the records transferred and preserved, and to offer future users the conditions and the tools for doing this verification.

- be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object: it is only possible to preserve the ability to reproduce the record.
recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form.

The principles and the criteria present here are crucial within any tradition and any juridical system. Some of these criteria have been already accepted by the rules for the reproduction of paper records in digital form and for guaranteeing the migration process of the digital records (deliberazione Aipa 24/1998), but the concrete procedure established by the legislator is very complex and for the moment not efficient and not easily implemented.

specify the requirements a copy of a record should satisfy to be considered equivalent to an original.

In the Italian legislation this principle is already established through the rules and the procedures for the authentication process: each records migration requires an authentication process, which has different levels of complexity. The officer entrusted with such a responsibility can be appointed from within its staff by a public administration, while in the private sector a notary is for the moment the only officer entrusted. In the course of the revision of the legislation related to the electronic records authentication, simpler and less expensive mechanisms will be provided, even if the risk is to create new concepts not consistent with the whole juridical tradition: for instance, the ambiguity of the concept of original—if not clearly identified—risks creating serious difficulties for the preservation of the integrity and authenticity of the records.

integrate records appraisal in the continuous process of preservation.

For the moment appraisal is not conceptualized enough in the Italian tradition. Also the training process needs to be further detailed and specified. It is correct to focus on the change of timing in the appraisal function, which in the digital environment has to start at the very beginning.

explicitly state that the entire process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term.

This concept, new in the national archival tradition, is very complex to implement because implementation would require lots of collaboration with creators plus attention to practice from the archivists who at the moment are not properly documented and trained. What should be explicitly present in the legislation, even if the general principles already exist in the dpr 445/2000, is the obligation for the public administrations to control the quality, the arrangement, and the descriptive tools of the current records before their transfer to the semi-current or historical repositories.

explicitly recognize that the traditional principle that all records relied upon in the usual and ordinary course of business can be presumed to be authentic needs to be supplemented in the case of electronic records by evidence that the records have not been inappropriately altered.

The principle that records relied upon in the usual course of business can be presumed authentic is not active in the Italian juridical system. Many tools (classification and registration, control on the records/files flows, etc.) are required also in the paper world to guarantee authenticity. Of course all these tools are required to be further refined in the digital environment.

recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records. The assessment of the authenticity of electronic records takes place before records are transferred to the custody of the preserver as part of the process of appraisal, while the maintenance of the authenticity of copies of electronic records takes place once they have been transferred to the preserver’s custody as part of the process of long-term preservation.

This principle is really new and constitutes a big change for Italian national archives and for any national archival programs. See previous comments.
• draw a clear distinction between the preservation of the authenticity of records and the authentication of a record.

With reference to this principle there is big confusion in the Italian legislation, which dedicates detailed rules with reference to the authentication process (electronic, digital signature), but not a word with reference to the authenticity problems. Only the archival community is aware of the difference and the complexity related to the authenticity, while officers (also in the public sector) and IT people do not seem to care about the complexity and the difficulty of this issue.
Introduction

European context

The Dutch National Archives participated in the InterPARES Project both on its own merit and as part of the multinational European team, which also included the Public Records Office of England, University of Glasgow (HATII), Dutch Archives School, Archives Nationales of France, National Archives (Riksarkivet) of Sweden, and the National Archives of Ireland. Europe has a long history of archival practice. Currently, each European country has its own legislative framework for records management and archives, incorporating a variety of cultural variations. As a result, despite commonalities, each country will ultimately differ in its assessment of the principles for ensuring authenticity of electronic records as identified by the InterPARES Project.

Dutch context

Records management and archives in the Netherlands are governed by the Archives Act (adapted in 1995) and subordinate regulations such as the Archives Ordinance (1995) and ministerial regulations on archives buildings; accessibility and arrangement of archival records; and sustainability of archival records. The ministerial regulation on accessibility and arrangement of archival records, in force since 1 January 2002, most notably provides guidelines for preserving authenticity; metadata; and the arrangement, accessibility and preservation of records (including the identification of standards for XML, TIFF, and PDF).1

Aside from the National Archives, the Ministry of the Interior plays a coordinating role for information management in central government, and the provincial "Eldermen" and the municipal councils play similar roles at their respective levels of administration. In 2001 the Dutch government adopted ISO Record Management Standard 15489 as a framework for proper record keeping, including electronic records. This created a need for further requirements and practical guidelines to help organizations meet the standard. Encouragement and support will be provided by other more generic legislation—covering such areas as freedom of information, privacy, and information security—that was being prepared and promulgated in 2002.

Legislation, policies, and approved standards that were already in place in 2002 provide a firm basis for the principles for ensuring authenticity of electronic records as identified by the InterPARES Project. As in most other countries, the changes and developments due to the increasing use of information technology raise many questions and create some confusion about the way in which traditional archival concepts have to be applied now. In several projects within the Dutch government, especially the Digital Longevity project, those issues are being addressed, and support and guidelines are provided for implementing proper records management in an electronic environment.

1 In the Dutch context no distinction is made between records and archives. Dutch legislation uses the word archiefbescheiden to indicate both. The Archives Act articulates the responsibility of government organizations of all administrative layers (including provinces and municipalities) for managing records and the related processes.
<table>
<thead>
<tr>
<th>Principle</th>
<th>Dutch comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. address records specifically rather than digital objects generally; that is, it should address documents made or received and set aside in the course of practical activity.</td>
<td>This is already a widely accepted principle. However, theoretically it can be stated that records are records because of their nature, even if they are not set aside. It is in order to guarantee their authenticity as long as the records are needed that they are managed and preserved in a safe and controlled environment. The principle is firmly based on the existing legislation and applies to government organizations and the records management community alike.</td>
</tr>
<tr>
<td>2. focus on authentic electronic records.</td>
<td>In a digital environment the authenticity of records raises questions, in the sense that it is not always clear what is meant. The issue is especially addressed, be it implicitly, in the new ministerial regulation on accessibility and arrangement (2002), which says in Article 2:  “The responsible authority ensures that it will be possible at all times to establish for each record the content, structure and form at the moment that it was received or drawn up by the administrative authority as by its nature destined to be kept by this organization, to the extent that the content, structure and form had to be recognizable for the accomplishment of the task or action by reason of which it was received or drawn up; at which time and by reason of which task or action it was received or drawn up by the administrative authority; the relationship with other records received and drawn up by the government organization.” The Dutch government’s adoption of ISO RMS 15489 supports the requirement of authenticity too.</td>
</tr>
<tr>
<td>3. recognize and provide for the fact that authenticity is most at risk when records are transmitted across space (i.e., when sent between persons, systems, or applications) or time (i.e., either when they are stored offline, or when the hardware or software used to process, communicate, or maintain them is upgraded or replaced).</td>
<td>The main risk at the moment is that, in Dutch government organizations, electronic records are not really managed. The most obvious example is e-mail messages. In an electronic environment, records management requires a completely new infrastructure (organizationally, technically, culturally, etc.). In the emerging world of electronic service delivery, this is even more so and it will be necessary to integrate records management into those processes. There is an increasing awareness of this issue and it is included in guidelines for developing electronic services. Apart from that, the Digital Longevity project offers practical tools (<a href="http://www.digitaleduurzaamheid.nl">www.digitaleduurzaamheid.nl</a>) and the Dutch Testbed project provides guidelines for preserving specific types of records and migration strategies (<a href="http://www.digitaleduurzaamheid.nl/">www.digitaleduurzaamheid.nl/</a>). Dealing with those issues, has created a huge need for training and education. There is a lack of knowledge, skills,</td>
</tr>
</tbody>
</table>
and experience in the whole area of digital preservation (and electronic records management) and in understanding the new world of digital information/records. New courses and curricula are being developed to meet this need.

4. recognize that preservation of authentic electronic records is a continuous process that begins with the process of records creation and whose purpose is to transmit authentic records across time and space. The concept of records continuum is gaining support. The question is, how will it be realized organizationally? There is still a gap between government organizations and archival institutions (life-cycle thinking). The National Archives currently focuses on making cultural sources accessible on the Web, but there is an increasing awareness of the responsibilities for electronic records to have archival value. The current point of view of the National Archives is that government organizations in the first instance are responsible for their own records and that the NA should not be involved, except for appraisal and transfer of these records. The task of coordination and guidance in this area of government organizations has been taken up by the Ministry of the Interior through the already mentioned program of Digital Longevity (of which also the NA is a partner, in the area of digital preservation).

5. be based on the concept of trust in records keeping and records preservation and specifically on the concepts of a trusted record-keeping system and the role of the preserver as a trusted custodian. This principle is already included in existing legislation and normal (traditional) practice. Archival institutions especially have fulfilled that role for 200 years.

6. be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object: it is only possible to preserve the ability to reproduce the record. This is an issue that has to be explained. It requires thinking in a way that many people, including members of the archival community, are not used to. Many people still look at records as physical entities. This is closely related to the issue of authenticity as well. It implies the necessity of having metadata for describing the essential characteristics of authenticity. That is included in the ministerial regulation on accessibility and arrangement (see above under 2).

7. recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form. As under 6.

8. specify the requirements a copy of a record should satisfy to be considered equivalent to an original. The Archives Act and the Archives Ordinance address the issue of substitution. For electronic records, it will be necessary to identify the essential characteristics for maintaining authenticity. That is addressed in the ministerial regulation on accessibility (2002).

9. Integrate records appraisal in the continuous process of preservation. This is the consequence of principle 4 (records continuum) and is already a long-standing tradition in Dutch archival legislation. Also, according to ISO RMS 15489, appraisal should be incorporated in records management applications.

10. Integrate archival description in the continuous process of preservation. As above, under 9.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11. explicitly state that the entire process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term.</td>
<td>The Netherlands government has adopted <em>ISO RMS 15489</em>, which also identifies this as a requirement. In the case of electronic records, one impact will be the creation of stricter requirements for a records management application or preservation function. The ministerial regulation on accessibility (2002) includes this requirement as well.</td>
</tr>
<tr>
<td>12. explicitly recognize that the traditional principle that all records relied upon in the usual and ordinary course of business can be presumed to be authentic needs to be supplemented in the case of electronic records by evidence that the records have not been inappropriately altered.</td>
<td>This is stated explicitly in <em>ISO RMS 15489</em> and it is included in the requirements of the ministerial regulation on accessibility (2002) of archival records.</td>
</tr>
<tr>
<td>13. recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records. The assessment of the authenticity of electronic records takes place before records are transferred to the custody of the preserver as part of the process of appraisal, while the maintenance of the authenticity of copies of electronic records takes place once they have been transferred to the preserver’s custody as part of the process of long-term preservation.</td>
<td>The issue of assessing the authenticity of electronic records needs more attention, particularly the implication of implementing records management systems that could support it. This is therefore included in policies for implementing appropriate records management in an electronic environment as now emerging in government organizations in the Netherlands.</td>
</tr>
<tr>
<td>14. draw a clear distinction between the preservation of the authenticity of records and the authentication of a record.</td>
<td>This principle is already accepted, though some people still need help to understand that a digital signature is not required for a record to be authentic. Authentication is seen as a means to secure the authenticity of a record, e.g., while preserving it. This is the case for both government organizations and the records management and archival communities.</td>
</tr>
</tbody>
</table>
Introduction

A defining characteristic of the United States' record-keeping context is its heterogeneity. Individual sectors, most notably the federal government and certain industries, are relatively homogeneous and controlled by specific records legislation, industry standards and regulations, and accreditation or licensing requirements; however, the United States consciously has not embraced a national information policy that requires uniform approaches to electronic records management. This is the result of several factors, including the common law juridical base of the United States; the distributed structure of the federal, state, and local systems of government; the traditional autonomies of the academic and religious sectors; and the increasing emphasis on enterprise and digital government facilitated through the implementation of state-of-the-art technology.

Records produced by federal, state, and local agencies are generally retained by a designated preserving agency in accordance with statutory requirements and associated records retention policies. Private sector records (e.g., those of businesses, religious organizations, museums, and private universities) are either retained by the creating agency in accordance with statutory, regulatory, and organizational records retention requirements or are, usually after a period of time, deposited in or donated to archival or manuscript repositories for the purposes of historical research.

The range of practices employed by the archival community reflects this pluralism in record-keeping requirements and responsibilities (e.g., both life-cycle and continuum approaches and custodial and non-custodial management of electronic records). This pluralism is also reflected in the differences in the resources available for archival management and in the level of expertise of institutional archivists in electronic records management. As a result, electronic records preservation policies have been developed and implemented primarily on an ad hoc basis as needed within individual organizational contexts. Few systematic electronic records programs exist outside of the National Archives and Records Administration (NARA) and certain state archives, despite the National Historical Publications and Records Commission’s funding emphasis on developing electronic records management principles, practices, and programs, and a measurable increase in the number of graduate education programs offering coursework in electronic records management.

This report reviews existing national and international legislation and standards that have implications for electronic records management, and specifically the preservation of authentic electronic records within one or more sectors in the United States. The review indicates the extent to which these standards or pieces of legislation address the principles for preservation policies, standards, or strategies identified by InterPARES. It offers some commentary about the current situation and how it might be improved. For example, the principles could be used by different sectors and interest groups to augment, qualify, or tighten the legislation and standards as sources of warrant; to suggest new legislation, standards, and policies; or to recognize and nurture best practices through professional education. It should be noted, however, that the preferred approach in many non-governmental sectors has been to enhance professional education in order to inculcate best practices, and to work in concert with professional archival associations to develop and support professional standards, rather than to respond to externally imposed standards.

1 For example, responsibility for the archival management of federal records resides with the U.S. National Archives and Records Administration, while in some states responsibility for local government records from clusters of counties is devolved to archival repositories at universities situated in those counties. The records of individual state universities are usually managed by their archives according to state and institutional records management requirements.
International and National Legislation and Standards

There are few national standards that relate specifically to the authenticity and long-term preservation of electronic records, although the corporate sector has warrant in the form of the ISO 15489 Records Management Standard as well as regulatory requirement, such as those of the Food and Drug Administration that affect approval of new products. The following legislation and standards are referred to in the subsequent analysis:

*Digital Millennium Copyright Act (DMCA)* 17 USC Section 101 et seq. (title IV amending §108, §112, §114, chapter 7 and chapter 8, title 17, United States Code)

President Bill Clinton signed the Digital Millennium Copyright Act into law on 28 October 1998. The legislation implements the 1996 World International Treaty and two World International Property Organization (WIPO) treaties: the WIPO Copyright Treaty and the WIPO Performances and Phonograms Treaty. Key provisions of the DMCA concern the circumvention of copyright protection systems, fair use in a digital environment, and online service provider (OSP) liability (including details on safe harbours, damages, and “notice and takedown” practices).

*Electronic Communications Privacy Act (ECPA)*, Title 18 of the United States Code, Section 2701 et seq.

This was adopted to address the legal privacy issues that were evolving with the growing use of computers and other innovations in electronic communications. The ECPA updated legislation passed in 1968 that had been designed to clarify what constitutes invasion of privacy when electronic surveillance is involved. The ECPA extended the privacy protection outlined in the earlier legislation to apply to radio paging devices, electronic mail, cellular telephones, private communication carriers, and computer transmissions.

*E-Sign (Electronic Signatures in Global and National Commerce Act)*, Title 15 of the United States Code, Section 7001 et seq. See also state digital signature legislation—for example, the Electronic Uniform Electronic Transactions Act (“UETA”).

E-Sign (Public Law 106–229), enacted on 30 June 2000, eliminates legal barriers to the use of electronic technology to sign and form contracts, collect and store documents, and send and receive notices and disclosures. E-Sign applies broadly to federal and state statutes and regulations governing private sector (including business-to-business and business-to-consumer) activities. E-Sign authorizes the substitution of electronic notices for paper notices, including most, but not all, types of consumer notices. E-Sign also includes a number of important protections to ensure that consumers can receive, keep, and use electronic notices provided to them.

*Government Paperwork Elimination Act*, Title 44 of the United States Code, Section 3504 note (GPEA).

The GPEA, enacted on 21 October 1998, requires that by October 2003, all executive branch agencies are to provide for the use and acceptance of electronic signatures in communications with the public, where practicable.


---

2 Though not specifically referenced in the chart that follows, the DMCA and other copyright legislation can pose significant challenges to the design and implementation of the long-term preservation of electronic records. Long-term preservation of authentic electronic and digital records may require copying that is outside the scope of current copyright protection. Future InterPARES policy research will include the examination of existing copyright and intellectual property regimes and their relationship to proposed digital preservation strategies and implementation.
Management). Records Management by Federal Agencies. See also state records and information management legislation.

*Presidential Records Act*, Title 44 of the United States Code, Sections 2201 et seq.

The *PRA*, enacted in 1978, changed the legal ownership of the official records of the President from private to public.

*Federal Rules of Evidence (FRE)* govern admissibility of evidence in administrative proceedings in federal courts. The general requirements address relevance, authentication, and hearsay aspects of evidence. While the *FRE* do not apply to suits in state courts, the rules of many states have been closely modeled on these provisions.

*Food and Drug Administration (FDA)* Title 21 of the CFR Part 11: Electronic Records; Electronic Signatures.

The Electronic Records and Electronic Signature Rule (21 CFR Part 11) was established by the U.S. Food and Drug Administration and put into effect on 20 August 1997. The rule defines the requirements for controlling electronic records and submitting documentation in electronic form, and the criteria for approved electronic signatures. It is designed to assist laboratories in the areas of improved data management, simplified regulatory compliance, and increased data security and integrity. The final rule relating to this title provides criteria under which *FDA* will consider electronic records to be equivalent to paper records, and electronic signatures equivalent to traditional handwritten signatures. Part 11 (21 CFR Part 11) applies to any paper records required by statute or agency regulations and supersedes any existing paper record requirements by providing that electronic records may be used in lieu of paper records. Electronic signatures that meet the requirements of the rule will be considered to be equivalent to full handwritten signatures, initials, and other general signings required by agency regulations. Section 11.2 provides that records may be maintained in electronic form and electronic signatures may be used in lieu of traditional signatures. Records and signatures submitted to the agency may be presented in an electronic form provided the requirements of Part 11 are met and the records have been identified in a public docket as the type of submission the agency accepts in an electronic form. Unless records are identified in this docket as appropriate for electronic submission, only paper records will be regarded as official submissions.


Provides that any person has the right to request access to federal agency records or information, and that agencies shall make reasonable efforts to search for records in electronic formats and provide to requesters records in any format (including electronic). All agencies of the United States government are required to disclose records upon receiving a written request for them, except for those records that are protected from disclosure by the nine exemptions and three exclusions of the *FOIA*. This right of access is enforceable in court. The federal *FOIA* does not, however, provide access to records held by state or local government agencies, or by private businesses or individuals. All states have their own statutes governing public access to state and local records; state agencies should be consulted for further information about them.

*International Organization for Standardization (ISO) 15489 Records Management Standards*

ISO 15489 focuses on the business principles behind records management and how organizations establish a framework to enable a comprehensive records management program. The new standard identifies key issues involved in retaining the information and making it available in a usable and reliable way. ISO 15489 is aimed at individuals responsible for setting policies, standards, and guidelines for information management within organizations. These include records managers, archivists, librarians, knowledge
management professionals, database managers, and business administrators within organizations who are responsible for the oversight of record-keeping processes.

*Reference Model for an Open Archival Information System (OAIS)—CCSDS*

The OAIS-Reference Model drafted by the National Aeronautics and Space Administration (NASA), is an ISO technical recommendation relating to the preservation of digital information by digital archives and their producers and consumers. *Referencing Model for an Open Archive Information System (OAIS)*, White Book, Issue 4, Don Sawyer / NASA and Lou Reich / CSC. Among the components of OAIS are the following: the reference model identifies a minimum set of responsibilities for an archive to claim it is an OAIS; establishes common terms and concepts for comparing implementations, but does not specify a specific implementation; provides detailed models of both archival functions and archival information; and discusses OAIS information migration and interoperability among OAISs.

*United States Department of Defense (DoD) 5015.2 Records Management Standard*

The DoD standard was created for use by agencies of the United States government. The standard is designed and expressed in terms of compliance with U.S. laws and regulations. The purpose of the DoD standard is to prescribe “mandatory baseline functional requirements, and to identify non-mandatory features deemed desirable for Records Management Application (RMA) Software." Within the context of the U.S. government, 5015.2 is a procurement standard requiring government agencies to purchase RMAs that are compliant with at least the minimum specifications.

The authenticity of archival records (i.e., that those records are indeed what they purport to be) is an aspect largely ignored in the legal context of the United States. Issues of system integrity and data reliability for active records are more common areas of concern when evidentiary value is an issue. The notion of authenticity of records in the sense used in diplomatics is largely alien to the corporate and legal records management communities in the United States.

**State and Local Context**

State and local authorities have also not systematically addressed the preservation of authentic electronic records, although legislation that parallels federal legislation often exists at the state level. The National Historical Publications and Records Commission has funded several initiatives to address electronic preservation issues at the state and local government level (e.g., in Minnesota, Mississippi, and the City of Philadelphia); and at individual academic institutions (e.g., Indiana University) that have sought to develop model solutions and policies in the absence of more specific legislation and standards.

The InterPARES research outcomes (i.e., principles, requirements, and models) will have an impact in the United States only to the extent that the authenticity and preservation of electronic records are considered universally pressing issues by archival, records management, and legal professionals. Although system integrity and access to reliable information are critical components of an effective electronic documentary record, the ability to establish and document the continued authenticity of electronic records is crucial to implementing an effective preservation plan.

---

\(^{4}\) The InterPARES glossary defines an *authentic record* as "a record that is what it purports to be and that is free from tampering or corruption."
The following table presents the principles that should govern any preservation policy, standard, or strategy for ensuring the long-term preservation of authentic electronic records. The principles are drawn from the report of the InterPARES Strategy Task Force. Each principle is paired with references to relevant legislation or standards that affect the application of the principle in the U.S. environment. Commentary on the application of the relevant legislation or standards or on the absence of any such legislation or standards is provided as appropriate.

<table>
<thead>
<tr>
<th>Principle</th>
<th>U.S. References and Commentary</th>
</tr>
</thead>
<tbody>
<tr>
<td>address records specifically rather than digital objects generally; that is, [any preservation policy, strategy, or standard] should address documents made or received and set aside in the course of a practical activity.</td>
<td>The U.S. National Archives is specifically charged with the archival management of the records of the federal government as defined by the <strong>Federal Records Act</strong> 44 U.S.C. Chapter 31. See also <strong>National Archives and Records Administration</strong> 44 U.S.C. Chapter 21; <strong>Records Management by the Archivist of the United States</strong>, 44 U.S.C. Chapter 29; <strong>Disposal of Records</strong> 44 U.S.C. Chapter 33; <strong>Coordinator of Federal Information Policy</strong> 44 U.S.C. Chapter 35; <strong>Information Technology Management Reform Act (ITMRA)</strong> 40 U.S.C. Section 1401 et seq.; <strong>Paperwork Reduction Act</strong> 44 U.S.C. Chapter 35; <strong>Administrative Procedure Act</strong>, 5 U.S.C. Chapter 5, the <strong>Freedom of Information Act</strong>, 5 U.S.C. Section 552, the <strong>Privacy Act</strong>, 5 U.S.C. Section 552a. In common practice, however, records and digital objects are typically undifferentiated in litigation and in business activities. Moreover, in non-federal repositories such as those of universities and local historical societies, records, manuscripts, and sometimes other library or artifact collections are often co-administered without explicitly addressing the distinctive preservation and authenticity needs of electronic records. The &quot;Open Archival Information System (OAIS) Reference Model&quot; refers only to information objects and not to records. See also <strong>FRE</strong>, <strong>DoD</strong>, and <strong>ISO</strong>.</td>
</tr>
<tr>
<td>focus on authentic electronic records.</td>
<td><strong>E-Sign (Electronic Signatures in Global and National Commerce Act)</strong> digital signature legislation addresses aspects of the reliability of electronic records; however, digital signatures only provide a means for assuring authenticity in time, and not preserving authenticity over time. See also <strong>FRE</strong>, <strong>DoD</strong>, and <strong>ISO</strong>.</td>
</tr>
</tbody>
</table>
recognize and provide for the fact that authenticity is most at risk when records are transmitted across space (i.e., when sent between persons, systems, or applications) or time (i.e., when they are stored offline, or when the hardware or software used to process, communicate, or maintain them is upgraded or replaced).

This is not currently addressed in the U.S. context.

<table>
<thead>
<tr>
<th>recognize that preservation of authentic electronic records is a continuous process that begins with the process of records creation and whose purpose is to transmit authentic records across time and space.</th>
</tr>
</thead>
</table>

be based on the concept of trust in records keeping and record preservation and specifically on the concepts of a trusted record keeping system and the role of the preserver as a trusted custodian.


be predicated on the understanding that it is not possible to preserve an electronic record as a stored physical object; it is only possible to preserve the ability to reproduce the record.

This is a new construct in the U.S. context.

recognize that the physical and intellectual components of an electronic record do not necessarily coincide and that the concept of digital component is distinct from the concept of element of documentary form.

Archival practice in the United States has not traditionally examined elements of documentary form in establishing record-keeping protocols and requirements. The concept of a digital component is a new construct in the U.S. context.

specify the requirements that a copy of a record should satisfy to be considered equivalent to an original.

See 17 USC 101 et seq.

integrate records appraisal in the continuous process of preservation.

U.S. archivists are increasingly involved in the design of record-keeping systems as well as scheduling electronic records. Both of these activities provide opportunities to integrate appraisal and description requirements into electronic record keeping at a pre-archival stage. U.S. archivists need increased education and training in how best to effect this integration in their own institutional contexts.
<table>
<thead>
<tr>
<th>Explicitly state that the entire process of preservation must be thoroughly documented as a primary means for protecting and assessing authenticity over the long term.</th>
<th>This principle underlies FRE. As with the previous principle, this is in part an issue of ensuring best practices through increased archival education and training in electronic records management. However, there is no current metadata framework that U.S. archivists could impose on record-keeping system design, or require of record-keeping procedures.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicitly recognize that the traditional principle that all records relied upon in the usual and ordinary course of business can be presumed to be authentic needs to be supplemented in the case of electronic records by evidence that the electronic records have not been inappropriately altered.</td>
<td>Not explicitly recognized in the U.S. context.</td>
</tr>
<tr>
<td>Recognize that the preserver is concerned with both the assessment and the maintenance of the authenticity of electronic records. The assessment of the authenticity of electronic records takes place before the records are transferred to the custody of the preserver as part of the process of appraisal, while the maintenance of the authenticity of copies of electronic records takes place once they have been transferred to the preserver’s custody as part of the process of long-term preservation.</td>
<td>This could in part be ensured through increased archival education and training in electronic records management and the development of professional best practices. Since individual U.S. archival repositories in the United States espouse both life cycle and continuum models of archival management, archivists need to understand how to apply this principle within their own institutional contexts.</td>
</tr>
<tr>
<td>Draw a clear distinction between the preservation of the authenticity of records and the authentication of a record.</td>
<td>This distinction is not currently made in the U.S. context.</td>
</tr>
</tbody>
</table>

---

1. The Federal Records Act, 44 U.S.C. 3301, defines federal records to include 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.

2. The Freedom of Information Act, 5 U.S.C. 552(f)(2), defines record to include “any information that would be an agency record subject to the requirements of this section when maintained by an agency in any format, including an electronic format.” In general, the definition of “agency record” under FOIA is broader than the definition of “record” under the Federal Records Act.

3. The Privacy Act, 5 U.S.C. 552a(a)(4) defines record to mean “any item, collection, or grouping of information about an individual that is maintained by an agency, including, but not limited to, his education, financial transactions, medical history, and criminal or employment history and that contains his name, or the identifying number, symbol, or other identifying particular assigned to the individual, such as a finger or voice print or a photograph.”
The FDA Electronic Records: Electronic Signatures, 21 CFR Part 11 defines electronic records as “any combination of text, graphics, data, audio, pictorial, or other information representation in digital form that is created, modified, maintained, archived, retrieved, or distributed by a computer system.”

Records required to be kept because of a law or regulation may be maintained in electronic format in lieu of paper, and electronic signatures in lieu of traditional signatures, in whole or in part, shall be regarded as equivalent to paper records and traditional signatures provided they meet the requirements of these guidelines except where specifically prohibited by law or regulation. 21 CFR 11.1 (c).

Ohio Administrative Code Section 149.01.1 (G) defines records to include “any document, device, or item, regardless of physical form or characteristic, created or received by or coming under the jurisdiction of any public office of the state or its political subdivisions, which serves to document the organization, functions, policies, decisions, procedures, operations, or other activities of the office.”

Ohio Administrative Code Section 1306.21 (1) defines “the minimum requirements of creation, maintenance, and security of electronic records and electronic signatures; (2) If electronic records must be signed by electronic means, all of the following: (a) the type of electronic signature required; (b) the manner and format in which the signature must be affixed to the electronic record; (c) the identity of, or criteria that must be met by, any third party used by the person filing a document to facilitate the process. (3) Control processes and procedures as appropriate to ensure adequate preservation, disposition, integrity, security, confidentiality, and auditability of electronic records; (4) Any other required attributes for electronic records that are specified for corresponding non-electronic records or reasonably necessary under the circumstances. (B) (1) The department of administrative services may adopt rules in accordance with section 111.15 of the Revised Code to ensure consistency and interoperability among state agencies with regard to electronic transactions, electronic signatures, and security procedures.”

New Mexico Title 1 General Government Administration Chapter 13 Public Records Part 70 Section 7 (A) defines records as “information preserved by any technique in any medium, now known, or later developed, that can be recognized by ordinary human sensory capabilities either directly or with the aid of technology.”
Committee Members

Ken Hannigan (Chair)
Representing the Authenticity Task Force: Maria Guercio
Representing the Appraisal Task Force: Philip Eppard
Representing the Preservation Task Force (successively): Bill Rhind, Hans Hofman, Torbjorn Hornfeldt (Vice-Chair),
Bruce Walton, Michele Cloonan, Rich Lysakowski
Ex officio as project director: Luciana Duranti
Glossary technical coordinator (successively): Peter Van Garderen, Jean-Pascal Morghese
Glossary Coordinator: Tahra Fung
Research Assistants: Yau Min Chong, Robert Edwards,
Prisca Giordani, Monica Greenan, Claire Vessaire
The InterPARES glossary is a controlled vocabulary of terms used in the InterPARES research. Its purpose is: to facilitate the communication of ideas and findings by offering clear definitions of the key terminology used in the InterPARES research; to maintain consistency in the use of such terminology within the project itself; and to ensure that the meanings of such terms will be communicated effectively to the world outside the project. The fact that the glossary in its entirety amounts to just over a hundred terms should serve to indicate that it is not intended as a comprehensive vocabulary of archival or electronic records terminology. It is hoped, however, that the glossary will have applications beyond being an aid to understanding of the work of the project. It is offered as a contribution to the dialogue that must take place across those disciplines containing the knowledge and expertise required to effect the long-term preservation of authentic records in digital systems. The representatives of those disciplines who have contributed to the InterPARES research, whether as investigators or as record creators, have included archivists, records managers, IT systems managers, computer scientists, and members of the pharmaceutical industry. It is a fact that distinct disciplines tend to develop their own currency of language, which, while easing the flow of knowledge and understanding within that discipline, may conversely impede the flow of ideas and understanding between disciplines. It is most important, therefore, that an interdisciplinary research project such as InterPARES should strive for precision in the use of key terms.

For archivists in particular, the need for precision in defining key concepts governing their work has become more important since the advent of digital technology. Traditional diplomatics and archival science informs much of the InterPARES research and one of the underlying principles of that research has been that much of what archivists need to allow them to identify records in electronic systems and to articulate their requirements for preserving those records over time is already available to them in traditional diplomatics and archival science. Within the digital landscape, many features of the record-keeping environment that seem unfamiliar at first glance begin to take on familiar shape when viewed through this prism. In drafting this glossary, therefore, the emphasis in the first instance has been on using where possible the language of diplomatics and archival science.

The compilation of the glossary has been overseen by a committee drawn from the task forces within and around which the main work of InterPARES has been organized. One representative each from the Appraisal Task Force, the Authenticity Task Force, and the Preservation Task Force served on the Glossary Committee along with the project director and a neutral chair. The committee conducted its business at face-to-face meetings held during the thrice-yearly InterPARES workshops, and electronically on a dedicated section of the InterPARES Web site which functioned as a glossary forum. Research work undertaken on the nominated terms to ensure consistency of use and compliance with the agreed criteria was undertaken by postgraduate research assistants from the School of Library, Archival and Information Studies at the University of British Columbia.

As each task force undertook research, it identified terms that it believed should be entered into the InterPARES glossary. Terms were eligible for nomination and inclusion in the glossary if they had been used in any part of the InterPARES research—whether in meetings, workshops, forum discussions, models, or other products of the research—or if they appeared in supporting literature used in the research. For a term to be eligible for inclusion it also had to meet one at least of the following seven nomination criteria demonstrating that it:

- represented an entity or concept that was key to understanding the research questions or findings of an InterPARES research domain (e.g., electronic record, appraisal)
- was specific to one discipline or sector and InterPARES researchers from another discipline or sector would be unlikely to be familiar with it
- had conflicting definitions and/or meanings between disciplines or sectors
• had conflicting definitions and/or meanings within a discipline or sector
• had multiple meanings
• had one or more synonyms
• was an obscure word or phrase that the majority of InterPARES researchers would find unfamiliar.

Terms for inclusion in the glossary were submitted from the task forces, by the task force representatives on the committee, from the International Team by the project director or by any member of the Glossary Committee.

The work flow of the committee was regulated by a clear set of procedures detailing the precise manner in which new terms or proposed revisions should be submitted and voted on and how, once accepted, work on them should progress through the various stages. Time limits were set for all stages. A software-based system monitored this work flow and held the documentation created during this process.

Proposed new terms or proposed revisions of exiting terms, along with their proposed definitions, were submitted formally to the Glossary Committee on standard template forms called "Term Proposal, Research and Nomination Forms" or "Term Record Revision Requests," which were then posted to the glossary forum. Minimum information required on the "Term Proposal, Research and Nomination Form" before it could be accepted for nomination included the name of the term, the proposed definition, the source for the definition, the context in which the term was used in the project, and how the term complied with stated nomination criteria.

Each nominated term was automatically accepted into the Glossary System, to be researched by the research assistants, unless an objection to its inclusion was signalled by a member of the Glossary Committee during a "quick vote" period. The quick vote period comprised the ten days following the first date of posting of a term to the glossary forum and was designed to effect a steady work flow by ensuring that undisputed terms advanced quickly by default to the next stage of the research process. Members of the Glossary Committee who were unhappy with a term could register an objection by replying to the posting of the term within the quick vote period, using the single word "opposed." A term disputed in this way came up for consideration at the next meeting of the Glossary Committee. If no opposition was signalled during the quick vote period, the term passed into the Glossary System and was researched by the research assistants who checked for previous occurrences of the term in the archival literature or other definitions that might conflict with the InterPARES definition.

Before a term that had passed into the Glossary System became a part of the published InterPARES glossary, a number of other steps were necessary. As already mentioned, each term that passed into the system was then researched by a research assistant, to ensure consistency of use and compliance with the agreed criteria. The research assistant also ensured that the term and its definition complied with agreed rules of style governing grammatical form, use of the singular and plural, spelling, capitalization, and use of acronyms. The style format rules were based on ISO 5964: Documentation—Guidelines for the Establishment and Development of Multilingual Thesauri and on the Oxford English Dictionary.

Based on this research, the research assistants made recommendations which were added to the appropriate section of the "Term Proposal, Research and Nomination Form."

Researched term were then resubmitted to the Glossary Committee using the quick vote system once again and, once again, terms failing to pass the quick vote were discussed at the next face-to-face meeting of the Glossary Committee. Although any member of the Glossary Committee could veto a term, the aim of the process throughout was to achieve consensus and thereby unanimity on all terms to be included in the glossary. The terms that are included in the final
glossary as it appears can be considered, therefore, as having passed through a rigorous process of research and examination. It is worth noting also that while the InterPARES glossary comprises merely the alphabetical list of terms and definitions, the *Glossary System* that has managed the process will continue to hold all the data on term histories, usage notes, and research on alternative definitions that have been generated in the course of the glossary's compilation.
Conclusion

The acronym InterPARES was chosen as the working title for this project because the Latin expression *inter pares* means “among peers” and there was a strong desire among the researchers to build a truly collaborative research project. It was easier said than done. Academics of diverse disciplines, professionals from a variety of fields, industry representatives, institutional representatives, and students often look at research from radically different perspectives. They have specific individual interests, may have organizational agendas, and engage in research often using contrasting approaches and methods. It took great care and nurturing on the part of myself and the national and multinational teams’ chairs, and even greater open-mindedness, flexibility, intellectual curiosity, and sheer effort on the part of all researchers to achieve the impressive level of integration that made our granting agency’s site visit committee state in its report: “the team took the time . . . to build a state of mind that emphasized ‘we’ rather than ‘me.’ This is a significant accomplishment, bringing together a wide variety of disciplines and cultures.”

Several activities supported the development of a common mind and a strong feeling of ownership on the part of each and every researcher, unit chair, and student research assistant. First, the project began with the writing of an organizational policy, which established quite clearly the categories of membership in the project and the respective responsibilities; the structure and composition of the research units, the management units and the administrative centre; the function of the research workshops, their schedule and their procedures (including the voting procedure); the publication policy, the Web site policy, the intellectual property policy, the collaborative research and authors’ guidelines, and the human subjects and data protection guidelines; and the procedures for maintaining the InterPARES research material, data, and records. The organizational policy was discussed in face-to-face meetings of all researchers, till consensus was reached on all clauses.

Second, a great deal of time was spent defining the terms and concepts on which the research had to be founded, such as *record, authenticity, archives, file, attributes, components, and elements*. At the outset, there were almost as many interpretations of those terms and concepts as there were researchers, given the fact that the same term could have different meanings from one culture to another, one discipline to another, and one organization to another, and conversely, that different terms could have the same meanings across cultures, disciplines, and organizations.

Third, face-to-face one-week-long research workshops, held three times a year, allowed all researchers to engage in scholarly debate; to review and discuss research activities, methodologies, draft documents, and preliminary findings and reconcile them; to distribute among themselves tasks and responsibilities; to adjust the direction of the research according to the findings; to resolve any conflict that might arise; and to deliberate on any issue brought forward by any researcher. These workshops were the most effective mechanism for promoting team integration across geographical and cultural distances and often competing conceptual frameworks.

As a result of all the above, each researcher brought unique qualities, skills, knowledge, and perspectives to research results that truly are everyone’s work, as the individual contributions are so blended together and indistinguishable that the whole is indeed much more than the sum of its parts.

---

In addition to a shared sense of the importance of their work, all researchers have demonstrated an exceptional ability to change the orientation of the research when it became clear that a certain direction was not viable. They have not clung rigidly to the framework established at the outset, nor have they sought or settled for the easy, but incomplete or disjunctured answer, and they have not hesitated to put into question again early findings, regardless of the significant risk of being unable to deliver the results by the scheduled deadlines. In other words, the intellectual rigour and courage of InterPARES researchers has been as remarkable as the collaborative spirit and the commitment to the research. And it has rubbed off on the students assisting in the research, who have been exceedingly enthusiastic about working on the project and, in the process, have developed an abundance of new skills and self-confidence. It is my belief that these characteristics of InterPARES researchers are evident in the work contained in the InterPARES findings.

Of course, the problems presented by the preservation of the authenticity of electronic records have not been solved by the InterPARES research, but a significant step forward has been made by having developed a shared view of the issues, concepts, principles, and criteria on which solutions must be based, and of the methodologies that are viable with existing technology, and by having built a solid foundation from which further research can be launched without having to reinvent the wheel. It is also important that both the shared view and the foundation are international, interdisciplinary, and valid across sectors and organizations.

The findings and products of InterPARES contained in *The Long-Term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* are complete and finished for the purposes of the first phase of the research project. However, as InterPARES continues into a second phase investigating more complex records generated in interactive, experiential, and dynamic digital environments and by activities that are yet to be regulated, that is, artistic activities (i.e., music, still and moving images, theatre, choreography), scientific activities (i.e., in the natural and social sciences), and electronic delivery of services by governments, these findings and products will be re-examined, refined, elaborated, and further developed. Moreover, several additional findings and products will be delivered. Therefore, stay tuned for a sequel!

Luciana Duranti
Director, InterPARES Project
Appendix 1

Template for Analysis

Authenticity Task Force

7 November 2000
Documentary Form

Definition: The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated. Documentary form possesses both extrinsic and intrinsic elements.

Extrinsic Elements of Documentary Form

Definition: The elements of a record that constitute its external appearance.

1. Presentation Features

Definition: A set of perceivable features (graphic, aural, visual) generated by means of encoding and program instructions, and capable, when used individually or in combination, to present a message to our senses.

Overall Presentation

Definition: The record's overall information configuration, i.e., the manner in which the content is presented to the senses.

Text

Definition: Words, numbers, or symbols.

Graphic

Definition: A representation of an object or outline of a figure, plan, or sketch by means of lines. A representation of an object formed by drawing.

Image

Definition: An artificial imitation or representation of the external form of any object, or an optical appearance or counterpart of an object, such as is produced by rays of light, refracted as through a lens, or falling on a surface after passing through a small aperture. A subset of image is moving images which are visual images, with or without sound, that, when viewed, present the illusion of motion.

Sound

Definition: Aural representation of words, music, or any other manifestation of sound.

Combination of More than One of the Above

Specific Presentation Features

Definition: Specific aspects of the record’s formal presentation that are necessary for it to achieve the purpose for which it was created.

Examples: Specific presentation features might include but are not limited to the following:

- special layouts
- deliberately employed type fonts
- deliberately employed colours
- hyperlinks
- graphic indication of attachments
- sample rate of sound files
- resolution of image files
- scales of maps
2. Electronic Signature

*Definition:* A digital mark having the function of a signature in, attached to, or logically associated with a record, which is used by a signatory to take responsibility or give consent to the content of that record, and which may be used to verify its authenticity.

**Electronic Seal**

*Definition:* Specific electronic means of authenticating a record or ensuring that it is only opened by the intended addressee. It is a distinct type of electronic signature.

*Example:* An electronic seal might include but is not limited to the following:

- digital signature, i.e., an electronic signature based on public key cryptography.

**Authentication Certificate of Trusted Third Party (TTP)**

*Definition:* An attestation issued by a TTP for the purpose of authenticating the ownership and characteristics of a public key. Such attestation appears in conjunction with the digital signature of the author of a record and is itself digitally signed by the TTP.

3. Digital Time-Stamp Issued by a Trusted Third Party (TTP)

*Definition:* An attestation by a TTP that a record was received at a particular point in time.

4. Special Signs

*Definition:* Symbol that identify one or more of the persons involved in the compilation, receipt, or execution of the record.

*Examples:* Special signs might include but are not limited to the following:

- digital watermarks
- organization crest
- personal logo
- originator identifier

**Intrinsic Elements of Documentary Form**

*Definition:* The elements of a record that convey the action in which the record participates and its immediate context.

1. Name of Author

*Definition:* Name of the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

*Note:* In traditional records, the name of the author typically appears as the name expressed in the letterhead (*entitling*), in the initial wording of the record (*superscription*), and/or at the bottom of the record (*subscription*). It may be the same name as that of the writer, and, with records that are electronically transmitted, may correspond to the name of the originator. However, the name of the author only validates the record when it has the function of an attestation.

2. Name of Originator

*Definition:* Name of the person assigned the electronic address in which the record has been generated and/or sent.
Note: When the name of the originator is different from the name of the author of the record, the law usually considers the originator’s name as the indication of the person responsible for issuing the record.

3. Chronological Date
Definition: The chronological date is the date, and possibly the time, of the record's compilation included in the record by the author or the electronic system on the author’s behalf.

4. Name of Place of Origin of Record
Definition: The name of the geographic place where the record was generated, included in the content of the record by the author or the electronic system on the author’s behalf.

5. Name of Addressee(s)
Definition: The name of the person(s) to whom the record is directed or for whom the record is intended.

Note: In traditional records this element corresponds to the inscription and usually occurs at the top of the record. With electronic mail records, the name of the addressee(s) continues to appear in the top portion of the record (i.e., in a header).

6. Name of Receiver(s)
Definition: The name of the person(s) to whom the record is copied for information purposes.

7. Indication of Action or Matter
Definition: The subject line(s) and/or the title at the top of the record.

8. Description of Action or Matter
Definition: Presentation of the ideal motivation (preamble) and the concrete reason (exposition) for the action as well as the action or matter itself (disposition).

9. Name of Writer
Definition: The name of the person having the authority and capacity to articulate the content of the record.

Note: In traditional records, the name of the writer usually appears at the bottom of the record and is typically constituted by the subscription. The name of the writer may be the same as the name of the author (and perhaps of the originator).

10. Corroboration
Definition: Explicit mention of the means used to validate the record.

Note: To validate means to make legally valid; to grant official sanction to by marking; to support or corroborate on a sound or authoritative basis.

11. Attestation
Definition: The written validation of a record by those who took part in the issuing of it (author, writer, counter-signer) and by witnesses to the action or to the signing of the record.

Note: In traditional records, the attestations usually appear as signatures at the bottom of the record (the eschatocol). However, some records have the attestation in the protocol. For example, memoranda may be signed or initialled beside the superscription. With electronic records, such as electronic mail messages, the attestation appears in the header of the message.
12. Qualification of Signature

Definition: The mention of the title, capacity and/or address of the persons signing a record.

Note: Qualification of signature may follow either a subscription or a superscription.

Annotations

Definition: Additions made to a record after it has been created.

Annotations Made in the Course of Executing the Record

Definition: Additions made to a record after it has been created as part of the formal execution phase of an administrative procedure.

Note: Normally this sort of annotation is used only for the authentication and registration of legal records whose form is required by law, e.g., the registration number added to a land deed by the land registry office, or the statement of the authenticity of the signatures in a will.

Examples: Such additions might include, but are not limited to the following:

- **Priority of Transmission**
  Definition: Indication of the priority in which a record is to be transmitted.

- **Transmission Date, Time and/or Place.**
  Definition: The date, time, and/or place when the record leaves the space in which it was generated.

  Note: Transmission date, time and/or place is usually added by the electronic system.

- **Indication of Attachments**
  Definition: Mention of autonomous items that have been linked inextricably to the record before transmission (i.e., added during its execution) in order for it to accomplish its purpose.

Annotations Made in the Course of Handling the Business Matter to which the Record Relates

Definition: Additions made to the record in the course of handling the business matter in which the record participates and reflecting actions taken subsequent to the creation of the record for the purpose of handling the action or matter in which the record participates.

Such additions might include, but are not limited to the following:

- Received Date and Time
- Name of Handling Office
- Action Taken
- Dates and Times of Further Action or Transmission

Annotations Made in the Course of Managing the Record for Records Management Purposes

Definition: Additions made to a record for the purpose of handling the record itself and reflecting actions taken subsequent to the creation of the record for the purpose of managing it as part of the agency’s records.

Such additions might include, but are not limited to the following:
• Archival Date
  *Definition:* The date on which a record is officially incorporated into the creator’s records.

• Draft or Version Number
  *Definition:* The unique identifier assigned to sequential drafts or versions of the same record, added to the record when it is saved.

• Record Item Identifier
  *Definition:* The progressive number of the record within the dossier or, in the absence of dossiers, within the specific class.

• Dossier Identifier
  *Definition:* The identifier for the dossier in which the record belongs.
  
  *Note:* It may be constituted by the name of a person or organization, a symbol, a progressive number, a date, or a specific topic within the class’s general subject.

• Class Code
  *Definition:* The code of the class to which the record belongs, as it appears in the classification scheme, thus connecting it to other records in the same class.

• Registration Number
  *Definition:* The consecutive number added to each incoming or outgoing record in the protocol register, which connects it to previous and subsequent records made or received by the creator.

• Name of Creator
  *Definition:* The name of the person in whose archival fonds the record exists.

**Medium**

*Definition:* The physical carrier of the message.

*Note:* The medium is considered an essential component of the record inasmuch as a record does not exist until it has been affixed to a physical carrier.

**Context**

*Definition:* The framework of action in which the record participates.

**Juridical-Administrative Context**

*Definition:* The legal and organizational system in which the creating body belongs.

*Note:* Indicators of juridical-administrative context are laws, regulations, etc.

**Provenancial Context**

*Definition:* The creating body, its mandate, structure, and functions.

*Note:* Indicators of provenancial context are organizational charts, annual reports, the classification scheme, etc.

**Procedural Context**

*Definition:* The business procedure in the course of which the record is created.

*Note:* In some organizations, the business procedures are integrated with documentary procedures. Indicators of procedural context are work-flow rules, codes of administrative procedure, classification schemes, etc.
Documentary Context

*Definition:* The fonds to which the record belongs and its internal structure.

*Note:* Indicators of documentary context are classification schemes, record inventories, indexes, registers, etc.

Technological Context

*Definition:* The characteristics of the technical components of the electronic system in which the record is created.

**Hardware**

1. **Storage**

   *Definition:* The medium that stores data in the system.

   **Main Memory (aka primary memory)**
   
   *Note:* This type of storage is fast, different parts of it can be accessed randomly (rather than sequentially) and directly by the CPU/microprocessor. Thus, for a process to run or a file to be accessed, it must be loaded, at least partially, into the main memory. Main memory is provided via integrated circuit chips and does not involve mechanical movements. It is "volatile" in that its contents will be lost when a computer system is shut down.

   *Example:* random access memory (RAM), cache memory.

   **Secondary Storage (aka secondary memory)**
   
   *Note:* This type of storage is slower than main memory and is cheaper. It involves mechanical parts and movements that contribute to its low speed of access. It is non-volatile in that shutting down the system will not result in loss of data on the secondary storage. Compared to magnetic tapes, secondary storage devices are randomly accessible.

   *Examples:* hard disks, magnetic or optical disks, CD-ROMs, DVDs.

   **Tertiary Storage**
   
   *Note:* This type of storage is sequentially accessible only, and is used for long-term file preservation.

   *Examples:* magnetic and digital tapes.

   **Storage for Security/Recovery Purposes**
   
   *Note:* This type of storage is used as a protective measure against the possibility of catastrophic loss. It tends to be overwritten at regular intervals and is not intended to serve the purpose of long-term file preservation.

   *Examples:* magnetic and digital tapes.

2. **CPU/Microprocessor**

   *Definition:* The primary resource for job/instruction execution.

   *Note:* This resource can be broken down further into its own sub-systems (e.g., registers and logic units). Its speed of executing instructions is considerably higher than the speed of accessing main memory. It interfaces directly with main memory, so a record must be loaded into main memory from secondary or tertiary storage before it can be readable.

3. **Network**

   *Definition:* The primary source of communication between systems or components thereof.

   *Note:* Network encompasses its own types of hardware, software, and architectures.
4. Peripheral Devices  
*Examples*: Mouse, monitor, keyboard, printer.

5. Architecture  
*Definition*: The configuration of hardware components and their interfaces.

*Examples*: CPU architecture, motherboard architecture, system architecture (i.e., serial, pipelined, parallel, distributed, client-server), network architecture.

### Software

1. **Operating System**  
*Definition*: The system that manages, controls, protects and facilitates the use of hardware resources in the electronic system.

*Note* The following can be identified as functions and main modules of an operating system: process management (scheduling, switching), deadlock management, memory management, secondary storage management, storage scheme (data mapping), disk scheduling, virtual memory management, file system (distributed, file format, directories), interrupt handling, user interface, device and network interface. The way an operating system is configured (parameterized), may affect certain aspects of data and files in the system. For example, there may be a limit imposed on the size of a data file.

2. **System Software**  
*Definition*: Software that creates an environment for application programs to be created, executed, and maintained, typically through system calls to the operating system.

*Note*: System software is sometimes referred to as system utilities or system tools.

*Examples*: languages (machine language, high-level languages), compilers, interpreters and translators, coding (compression, encryption), system utilities (i.e., hard disk defragmentation tools, virus detectors, etc.).

3. **Network Software**  
*Definition*: Network software manages networks and their resources in order to meet the communication requirements of one or more applications.

*Examples*: protocols, routing, and switching software.

4. **Application Software**  
*Definition*: Software that constitutes any type of program that is tailored to satisfy real-world needs and requirements.

*Note*: Application software varies widely in nature and complexity, as the range of applications using this type of software is quite diverse. Application software may be developed in-house by the organization that uses it, custom-made by another company or contractor for the organization that uses it, or purchased as an off-the-shelf package. It is important to know whether the software includes source code, documentation, and other components, in addition to the executables. As in the operating system, a set of parameters or characteristics may be associated with the application software whose values affect the number, format and size of the records that are handled.

*Examples*: Microsoft Word, Lotus 1-2-3, Netscape Communicator, database management system (DBMS) software, computer-aided design (CAD) software.

### Data

*Definition*: numbers, characters, images or other methods of recording that represent values that can be stored, processed, and transmitted by electronic systems.
1. File Structure  
**Definition:** The relationship and organization of files within a system.  

*Note:* File structure includes the directory structure of a file system. The physical structure and organization of files in a file system may also constitute an aspect of the file structure and data format. This can include the mapping of files onto disk blocks of each disk plate, and among a set of disks.  

2. Data Format/File Format  
**Definition:** The organization of data within files. These are organizations that are usually designed to facilitate the storage, retrieval, processing, presentation, and/or transmission of the data by software.  

*Note:* Data format is concerned with the representation of each piece of data and the relationship between pieces of data. Within a file, it includes standardized data formats such as ASCII text, as well as proprietary file formats such as Microsoft's Word97 and Adobe's PDF file formats. It also includes structures such as the tabular format of data files in a database management system, and the format (using tags) of data files used by markup languages.  

*Examples:* portable document format (PDF), rich text format (RTF), ASCII text.  

**System Models**  
**Definition:** System models are abstractions that represent the entities, activities and/or concepts in the system as well as their attributes, characteristics, and the functional relationship between them.  

*Note:* "Functional relationship" refers to a relationship involving two or more entities/objects that it is important to represent explicitly in order for the application to function correctly. System models contrast with data format and file structure in that they represent behavioural, procedural, and/or functional aspects of a system or software application. They may, however, affect directly or indirectly the way files are conceived in an application and the way data are organized within the files in an application. A model is usually represented graphically (e.g., as in entity-relationship, object-hierarchy, data-flow, control-flow, and state-transition diagrams). Modelling languages (e.g., IDEF, UML) and their associated software tools serve as aides in creating model specifications. The model usually becomes part of an application’s requirements, specifications, and/or design document. Parts of the model can also be represented and used in an application’s data dictionary.  

*Examples:* entity-relationship models, object domain diagrams, IDEF(0) process models, UML use-case models, data-flow diagrams.  

**System Administration**  
**Definition:** System administration is a set of procedures that ensure correct, secure, reliable, and persistent operation of the system.  

*Examples:* Providing access privileges; ensuring security, availability, reliability and integrity of the system over time; configuring the system; backing up files; system maintenance; and upgrading hardware, software and storage systems.
Appendix 2

Requirements for Assessing and Maintaining the Authenticity of Electronic Records

Authenticity Task Force

March 2002
The requirements that are identified in this document fall into two groups: requirements that support the presumption of the authenticity of electronic records before they are transferred to the custody of the preserver,¹ and requirements that support the production of authentic copies of electronic records that have been transferred to the custody of the preserver. The report is organized into the following sections:

- Conceptual Framework for the Requirements for Assessing and Maintaining the Authenticity of Electronic Records
- Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records
- Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records
- Commentary on the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records
- Commentary on the Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

Conceptual Framework for the Requirements for Assessing and Maintaining the Authenticity of Electronic Records

Introduction

Authenticity is defined as “the quality of being authentic, or entitled to acceptance.”² Authentic means “worthy of acceptance or belief as conforming to or based on fact” and is synonymous with the terms genuine and bona fide. Genuine “implies actual character not counterfeited, imitated, or adulterated [and] connotes definite origin from a source.” Bona fide “implies good faith and sincerity of intention.”³ From these definitions it follows that an authentic record is a record that is what it purports to be and is free from tampering or corruption.

In both archival theory and jurisprudence, records that the creator⁴ relies on in the usual and ordinary course of business are presumed authentic. However, digital information technology creates significant risks that electronic records may be altered, either inadvertently or intentionally. Therefore, in the case of records maintained in electronic systems, the presumption of authenticity must be supported by evidence that a record is what it purports to be and has not been modified or corrupted in essential respects. To assess the authenticity of an electronic record, the preserver must be able to establish its identity and demonstrate its integrity.

The identity of a record refers to the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records. From an archival-diplomatic perspective, such attributes include: the names of the persons concurring in its formation (i.e., its author, addressee, writer, and originator); its date(s) of creation (i.e., the date it was made, received, and set aside) and its date(s) of transmission; an indication of the action or matter in which it participates; the expression of its archival bond, which links it to other records participating in the same action (e.g., a classification code or other unique identifier); as well as an indication of any attachment(s) since an attachment is considered an

¹ The preserver is the juridical person whose primary responsibility is the long-term preservation of authentic records. The preserver’s responsibilities include appraisal.


³ Merriam-Webster Online Collegiate Dictionary, s.v. “authentic.”

⁴ The creator is the physical or juridical person in whose archival fonds the record exists. The fonds is the whole of the records created (meaning made or received and set aside for action or reference) by a physical or juridical person in the course of carrying out its activities.
integral part of a record. The attributes that establish the identity of a record may be explicitly expressed in an element of the record, in metadata related to the record, or they may be implicit in its various contexts. Those contexts include: its **documentary context**, that is, the archival fonds to which a record belongs, and its internal structure; its **procedural context**, that is, the business process in the course of which the record is created; its **technological context**, that is, the characteristics of the technical components of an electronic computing system in which records are created; its **provenancial context**, that is, the creating body, its mandate, structure, and functions; and its **juridical-administrative context**, that is, the legal and organizational system in which the creating body belongs.

The **integrity** of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects. This does not mean that the record must be precisely the same as it was when first created for its integrity to exist and be demonstrated. Even in the paper world, with the passage of time, records are subject to deterioration, alteration and/or loss. In the electronic world, the fragility of the media, the obsolescence of technology, and the idiosyncrasies of systems likewise affect the integrity of records. When we refer to an electronic record, we consider it essentially complete and uncorrupted if the message that it is meant to communicate in order to achieve its purpose is unaltered. This implies that its physical integrity, such as the proper number of bit strings, may be compromised, provided that the articulation of the content and any required annotations and elements of documentary form remain the same. The integrity of a record may be demonstrated by evidence found on the face of the record, in metadata related to the record, or in one or more of its various contexts.

**Assessing the Authenticity of Electronic Records**

The records of the creator belong to one of two categories. The first category comprises those records that exist as created. They are considered authentic because they are the same as they were in their first instantiation. The second category comprises those records that have undergone some change and therefore cannot be said to exist as first created; they are considered authentic because the creator treats them as such by relying on them for action or reference in the regular conduct of business. However, the authenticity of electronic records is threatened whenever they are transmitted across space (i.e., when sent to an addressee or between systems or applications) or time (i.e., either when they are in storage, or when the hardware or software used to store, process, or communicate them is updated or replaced). Given that the acts of setting aside an electronic record for future action or reference and of retrieving it inevitably entail moving it across significant technological boundaries (from display to storage subsystems and vice versa), virtually all electronic records belong to the second category. Therefore, the preserver’s inference of the authenticity of electronic records must be further supported by evidence—provided in association with the records—that they have been maintained using technologies and administrative procedures that either guarantee their continuing identity and integrity or at least minimize risks of change from the time the

---

5 An attachment is a document that constitutes an integral part of the whole record, notwithstanding the fact that it exists as a linked, but physically separate entity.

6 The use of the terms *attribute* and *element* in this report should not be confused with the way the terms are used in other contexts, such as the various Standard Generalized Mark-up Languages (SGML). In this report, a record attribute is a defining characteristic of a record or of a record element. A *record element* is a constituent part of the record's documentary form and may be either extrinsic or intrinsic. An attribute may manifest itself in one or more elements of a record's documentary form. For example, the name of the author of a record is an attribute, which may be expressed as a superscription or a signature, both of which are intrinsic elements of documentary form. For a more detailed explanation of the extrinsic and intrinsic elements of documentary form see the Authenticity Task Force's *Template for Analysis*, in Appendix 1. An attribute may also manifest itself in the form of an annotation(s) to a record, in metadata linked to it, or in one or more of its various contexts.

7 For example, for an electronic mail message, an authentic copy of a complete message may include only the text. Provided it clearly indicated the author, addressee, receivers, and date as well as the content, it would not need to appear in the same way in which it was seen by the author or addressee. In contrast, an authentic copy of a map would have to retain its original presentation features, including colour and feature presentation. Provided these requirements were met, an authentic copy could be produced in GIF, JPEG, or GML format.
records were first set aside to the point at which they are subsequently accessed. The requirements for assessing the authenticity of the creator’s electronic records concern this evidence.

The Presumption of Authenticity
A presumption of authenticity is an inference that is drawn from known facts about the manner in which a record has been created and maintained. The evidence that supports the presumption that the record creator created and maintained them authentic are enumerated in the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records (Requirement Set A). A presumption of authenticity will be based upon the number of requirements that have been met and the degree to which each has been met. The requirements are, therefore, cumulative: the higher the number of satisfied requirements, and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity. This is why these requirements are termed “benchmark” requirements.

The Verification of Authenticity
In any given case, there may be an insufficient basis for a presumption of authenticity, or the presumption may be extremely weak. In such cases, further analysis may be necessary to verify the authenticity of the records. A verification of authenticity is the act or process of establishing a correspondence between known facts about the record and the various contexts in which it has been created and maintained, and the proposed fact of the record’s authenticity. In the verification process, the known facts about the record and its contexts provide the grounds for supporting or refuting the contention that the record is authentic. Unlike the presumption of authenticity, which is established on the basis of the benchmark requirements, this verification involves a detailed examination of the records themselves and reliable information available from other sources about the records and the various contexts in which they have been created and maintained. Methods of verification include, but are not limited to, a comparison of the records in question with copies that have been preserved elsewhere or with back-up tapes; comparison of the records in question with entries in a register of incoming and outgoing records; textual analysis of the record’s content; forensic analysis of the medium, script, and so on; a study of audit trails; and the testimony of a trusted third party.

Maintaining the Authenticity of Electronic Records
After the records have been presumed or verified authentic in the appraisal process, and have been transferred from the creator to the preserver, their authenticity needs to be maintained by the preserver. In order to do so, the preserver must carry forward the records in accordance with the baseline requirements that apply to the maintenance of records, producing copies according to procedures that also maintain authenticity. The production of authentic copies is regulated by the Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records (Requirement Set B). Unlike the benchmark requirements, all of the requirements included in the baseline requirements must be met before the preserver can attest to the authenticity of the electronic copies in its custody. This is why the requirements for the production of authentic electronic copies are termed “baseline” requirements.

Satisfaction of these baseline requirements will enable the preserver to certify that copies of electronic records are authentic. Traditionally, the official preserver of the records has been the person entrusted with issuing authentic copies of such records. To fulfill that role, the preserver needed simply to attest that the

8 In common usage, verify is synonymous with the terms validate, confirm, corroborate, and substantiate. According to Merriam-Webster Online Collegiate Dictionary, “validate means to attest to the truth or validity of something; confirm implies the removing of doubts by an authoritative affirmation or by factual proof; corroborate suggests the strengthening of something that is already partly established; substantiate implies the offering of evidence that sustains the contention.”

9 It is understood that the records maintained by the preserver exist only as copies of the creator’s records.
copy conformed to the record being reproduced. With electronic records, and the accompanying difficulties related to preservation, the prudent path would be for the preserver to produce and maintain documentation relating to the manner in which it has maintained the records over time as well as the manner in which it has reproduced them to support its attestation of authenticity.

A copy is the result of a reproduction process. A copy can be made from an original or from a copy of either an original or another copy. 10 There are several types of copy. The most reliable copy is a copy in the form of an original, which is identical to the original although generated subsequently. An imitative copy is a copy that reproduces both the content and form of the record, but in such a way that it is always possible to tell the copy from the original. A simple copy is a copy that reproduces only the content of the original.

Any of these types of copy is authentic if attested to be so by the official preserver. By virtue of this attestation, the copy is deemed to conform to the record it reproduces until proof to the contrary is shown. Such attestation is supported by the preserver’s ability to demonstrate that it has satisfied the applicable baseline requirements for maintenance and all of the requirements for the production of authentic copies.

Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records

Preamble

The benchmark requirements are the conditions that serve as a basis for the preserver’s assessment of the authenticity of the creator’s electronic records. Satisfaction of these benchmark requirements will enable the preserver to infer a record’s authenticity on the basis of the manner in which the records have been created, handled, and maintained by the creator.

Within the benchmark requirements, Requirement A.1 identifies the core information about an electronic record—the immediate context of its creation and the manner in which it has been handled and maintained—that establishes the record’s identity and lays a foundation for demonstrating its integrity. Requirements A.2–A.8 identify the kinds of procedural controls over the record’s creation, handling, and maintenance that support a presumption of its integrity.

10 In common language, copy and reproduction are synonyms. For the purposes of this research, the term reproduction is used to refer to the process of generating a copy, while the term copy is used to refer to the result of such a process, that is, to any entity which resembles and is generated from the records of the creator. An original record is the first, complete record, which is capable of achieving its purposes (i.e., it is effective). A record may also take the form of a draft, which is a temporary compilation made for purposes of correction.
Benchmark Requirements (Requirement Set A)

To support a presumption of authenticity the preserver must obtain evidence that:

REQUIREMENT A.1: Expression of Record Attributes and Linkage to Record

the value of the following attributes are explicitly expressed and inextricably linked to every record. These attributes can be distinguished into categories, the first concerning the identity of records, and the second concerning the integrity of records.

A.1.a Identity of the record:

A.1.a.i Names of the persons concurring in the formation of the record, that is:

- name of author
- name of writer (if different from the author)
- name of originator (if different from name of author or writer)
- name of addressee

A.1.a.ii Name of action or matter

A.1.a.iii Date(s) of creation and transmission, that is:

- chronological date
- received date
- archival date
- transmission date(s)

A.1.a.iv Expression of archival bond (e.g., classification code, file identifier)

---

11 The name of the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

12 The name of the physical or juridical person having the authority and capacity to articulate the content of the record.

13 The name of the physical or juridical person assigned the electronic address in which the record has been generated and/or sent.

14 The name of the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

15 The date, and possibly the time, of compilation of a record included in the record by the author or the electronic system on the author’s behalf.

16 The date, and possibly the time, when a record is received by the addressee.

17 The date, and possibly the time, when a record is officially incorporated into the creator’s records.

18 The date and time when a record leaves the space in which it was generated.

19 The archival bond is the relationship that links each record, incrementally, to the previous and subsequent ones and to all those participate in the same activity. It is originary (i.e., it comes into existence when a record is made or received and set aside), necessary (i.e., it exists for every record), and determined (i.e., it is characterized by the purpose of the record).
A.1.a.v Indication of attachments

A.1.b Integrity of the record:
   A.1.b.i Name of handling office
   A.1.b.ii Name of office of primary responsibility (if different from handling office)
   A.1.b.iii Indication of types of annotations added to the record
   A.1.b.iv Indication of technical modifications;

**REQUIREMENT A.2:**
Access Privileges
the creator has defined and effectively implemented access privileges concerning the creation, modification, annotation, relocation, and destruction of records;

**REQUIREMENT A.3:**
Protective Procedures:
Loss and Corruption of Records
the creator has established and effectively implemented procedures to prevent, discover, and correct loss or corruption of records;

**REQUIREMENT A.4:**
Protective Procedures:
Media and Technology
the creator has established and effectively implemented procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change;

**REQUIREMENT A.5:**
Establishment of Documentary Forms
the creator has established the documentary forms of records associated with each procedure either according to the requirements of the juridical system or those of the creator;

---

20 The office (or officer) formally competent for carrying out the action to which the record relates or for the matter to which the record pertains.

21 The office (or officer) given the formal competence for maintaining the authoritative record, that is, the record considered by the creator to be its official record.

22 Annotations are additions made to a record after it has been completed. Therefore, they are not considered elements of the record's documentary form.

23 Technical modifications are any changes in the digital components of the record as defined by the Preservation Task Force. Such modifications would include any changes in the way any elements of the record are digitally encoded and changes in the methods (software) applied to reproduce the record from the stored digital components; that is, any changes that might raise questions as to whether the reproduced record is the same as it would have been before the technical modification. The indication of modifications might refer to additional documentation external to the record that explains in more detail the nature of those modifications.
REQUIREMENT A.6: Authentication of Records
if authentication is required by the juridical system or the needs of the organization, the creator has established specific rules regarding which records must be authenticated, by whom, and the means of authentication;

REQUIREMENT A.7: Identification of Authoritative Record
if multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative;

REQUIREMENT A.8: Removal and Transfer of Relevant Documentation
if there is a transition of records from active status to semi-active and inactive status, which involves the removal of records from the electronic system, the creator has established and effectively implemented procedures determining what documentation has to be removed and transferred to the preserver along with the records.

Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

Preamble
The baseline requirements outline the minimum conditions necessary to enable the preserver to attest to the authenticity of copies of inactive electronic records.

Baseline Requirements (Requirement Set B)
The preserver should be able to demonstrate that:

REQUIREMENT B.1: Controls over Records Transfer, Maintenance, and Reproduction
the procedures and system(s) used to transfer records to the archival institution or program; maintain them; and reproduce them embody adequate and effective controls to guarantee the records’ identity and integrity, and specifically that

B.1.a Unbroken custody of the records is maintained;
B.1.b Security and control procedures are implemented and monitored; and
B.1.c The content of the record and any required annotations and elements of documentary form remain unchanged after reproduction.

REQUIREMENT B.2: Documentation of Reproduction Process and its Effects
the activity of reproduction has been documented, and this documentation includes

B.2.a The date of the records’ reproduction and the name of the responsible person;
B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver;

B.2.c The impact of the reproduction process on their form, content, accessibility and use; and

B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user;

REQUIREMENT B.3: Archival Description

the archival description of the fonds containing the electronic records includes—in addition to information about the records’ juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the electronic records of the creator have undergone since they were first created.

Commentary on the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records

The assessment of the authenticity of the creator’s records takes place as part of the appraisal process. That process and the role of the benchmark requirements within it are described in more detail in the "Appraisal Task Force Report." This assessment should be verified when the records are transferred to the preserver’s custody.

A.1: Expression of Record Attributes and Linkage to Record

The presumption of a record’s authenticity is strengthened by knowledge of certain basic facts about it. The attributes identified in this requirement embody those facts. The requirement that the attributes be expressed explicitly and linked inextricably24 to the record during its life, and carried forward with it over time and space, reflects the task force’s belief that such expression and linkage provide a strong foundation on which to establish a record’s identity and demonstrate its integrity. The case studies undertaken as part of the work of the task force revealed very little consistency in the way the attributes that specifically establish the identity of a record are captured and expressed from one electronic system to another. In certain systems, some attributes were explicitly mentioned on the face of the record; in others they could be found in a wide range of metadata linked to the record or they were simply implicit in one or more of the record’s contexts. In many cases, certain attributes (e.g., the expression of the archival bond) were not captured at all. The task force’s concern is that, in the absence of a precise and explicit statement of the basic facts concerning a record’s identity and integrity, it will be necessary for the preserver to acquire enormous, and otherwise unnecessary, quantities of data and documentation simply to establish those facts.

The link between the record and the attributes listed in Requirement A.1 is viewed by the task force as a conceptual rather than a physical one, and the requirement could be satisfied in different ways, depending on the nature of the electronic system in which the record resides. For example, in electronic records management systems, this requirement is usually met through the creation of a record profile.25 In other types of systems, the requirement could be fulfilled through a topic map. A topic map expresses the characteristics (i.e., topics) of subjects (e.g., records or record attributes) and the relationships between and among them.

When a record is exported from the live system, migrated in a system update, or transferred to the preserver, the attributes should be linked to the record and available to the user. When pulling together the

24 For the purposes of this requirement, inextricable means incapable of being disentangled or untied, and link means a connecting structure.

25 If the attribute values contained in the profile are also expressed independently as entries in a register of all records made or received by the creator, then, in addition to establishing the identity and supporting the inference of the integrity of the record, they would corroborate such identity and strengthen the inference of integrity.
data prior to export, the creator should also ensure that the data captured are the right data. For example, in the case of distribution lists, the creator must ensure that if the recipients specified on "List A" were changed at some point in the active life of records, the accurate "List A: Version 1" is exported with the records associated with the first version, and that the second version is sent forward with those records sent to recipients on "List A: Version 2."

A.2 Access Privileges
Defining access privileges means assigning responsibility for the creation, modification, annotation, relocation, and destruction of records on the basis of competence, which is the authority and capacity to carry out an administrative action. Implementing access privileges means conferring exclusive capability to exercise such responsibility. In electronic systems, access privileges are usually articulated in tables of user profiles. Effective implementation of access privileges involves the monitoring of access through an audit trail that records every interaction that an officer has with each record (with the possible exception of viewing the record). If the access privileges are not embedded within the electronic system but are based on an external security system (such as the exclusive assignment of keys to a location), the effective implementation of access privileges will involve monitoring the security system.

A.3 Protective Procedures: Loss and Corruption of Records
Procedures to protect records against loss or corruption include: prescribing regular back-up copies of records and their attributes; maintaining a system back-up that includes system programs, operating system files, etc.; maintaining an audit trail of additions and changes to records since the last periodic back-up; ensuring that, following any system failure, the back-up and recovery procedures will automatically guarantee that all complete updates (records and any control information such as indexes required to access the records) contained in the audit trail are reflected in the rebuilt files and also guarantee that any incomplete operation is backed up. The capability should be provided to rebuild forward from any back-up copy, using the back-up copy and all subsequent audit trails.

A.4 Protective Procedures: Media and Technology
Procedures to counteract media fragility and technological obsolescence include: planning upgrades to the organization’s technology base; ensuring the ability to retrieve, access, and use stored records when components of the electronic system are changed; refreshing the records by regularly moving them from one storage medium to another; and migrating records from an obsolescent technology to a new technology.

A.5 Establishment of Documentary Forms
The documentary form of a record may be determined in connection to a specific administrative procedure, or in connection to a specific phase(s) within a procedure. The documentary form may be prescribed by business process and work-flow control technology, where each step in an administrative procedure is identified by specific record forms. If a creator customizes a specific application, such as an electronic mail application, to carry certain fields, the customized form becomes, by default, the required documentary form. It is understood that the creator, acting either on the basis of its own needs or the requirements of the juridical system, not an individual officer, establishes the required documentary form(s) of records.

When the creator establishes the documentary form in connection to a procedure, or to specific phases of a procedure, it is understood that this includes the determination of the intrinsic and extrinsic elements of form that will allow for the maintenance of the authenticity of the record. Because, generally speaking, that determination will vary from one form of a record to another, and from one creator to another, it is not possible to predetermine or generalize the relevance of specific intrinsic and extrinsic elements of documentary form in relation to authenticity.

A.6 Authentication of Records
In common usage, to authenticate means to prove or serve to prove the authenticity of something. More specifically, the term implies establishing genuineness by adducing legal or official documents or expert opinion. For the purposes of the benchmark requirements, authentication is understood to be a declaration of a record’s authenticity at a specific point in time by a juridical person entrusted with the authority to make such declaration. It takes the form of an authoritative statement (which may be in the form of words or

---

26 The extrinsic and intrinsic elements of form are defined and explained in the Authenticity Task Force’s Template for Analysis, Appendix 1.
symbols) that is added to or inserted in the record attesting that the record is authentic.\textsuperscript{27} The requirement may be met by linking the authentication of specific types of records to business procedures and assigning responsibility to a specific office or officer for authentication.

The authentication of copies differs from the validation of the process of reproduction of the digital components of the records. The latter process occurs every time the records of the creator are moved from one medium to another or migrated from one technology to another.

A.7 Identification of Authoritative Record

An authoritative record is a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other copies. The identification of authoritative records corresponds to the designation of an office of primary responsibility as one of the components of a record retention schedule. The Office of Primary Responsibility is the office given the formal competence for maintaining the authoritative (that is, official) records belonging to a given class within an integrated classification scheme and retention schedule. The purpose of designating an Office of Primary Responsibility for each class of record is to reduce duplication and to designate accountability for records.

It is understood that in certain circumstances there may be multiple authoritative copies of records, depending on the purpose for which the record is created.

A.8 Removal and Transfer of Relevant Documentation

This requirement implies that the creator needs to carry forward with the removed records all the information that is necessary to establish the identity and demonstrate the integrity of those records, as well as the information necessary to place the records in their relevant contexts.

Commentary on the Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

The establishment and implementation of the baseline requirements take place as part of the function of managing preservation. The preservation function and the role of the baseline requirements within it are described in more detail in the "Preservation Task Force Report."

B.1 Controls over Records Transfer, Maintenance, and Reproduction

The controls over the transfer of electronic records to archival custody include establishing, implementing, and monitoring procedures for registering the records’ transfer; verifying the authority for transfer; examining the records to determine whether they correspond to the records that are designated in the terms and conditions governing their transfer; and accessioning the records.

As part of the transfer process, the assessment of the authenticity of the creator’s records, which has taken place as part of the appraisal process, should be verified. This includes verifying that the attributes relating to the records’ identity and integrity have been carried forward with them (Requirement A.1), along with any relevant documentation (Requirement A.8).

The controls over the maintenance of electronic records once they have been transferred to archival custody are similar to several of the ones enumerated in the benchmark requirements. For example, the preserver should establish access privileges concerning the access, use, and reproduction of records (Requirement A.2); establish procedures to prevent, discover, and correct loss or corruption of records (Requirement A.3), as well as procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change (Requirement A.4). Once established, the privileges and procedures should be effectively implemented and regularly monitored. If authentication of the records is

\textsuperscript{27} The meaning of authentication as it is used by the Authenticity Task Force in this report is broader than its meaning in public key infrastructure (PKI) applications. In such applications, authentication is restricted to proving identity and public key ownership over a communication network.
required, the preserver should establish specific rules regarding who is authorized to authenticate them and the means of authentication that will be used (Requirement A.6).

The controls over the reproduction of records include establishing, implementing, and monitoring reproduction procedures that are capable of ensuring that the content of the record is not changed in the course of reproduction.

B.2 Documentation of Reproduction Process and its Effects
Documenting the reproduction process and its effects is an essential means of demonstrating that the reproduction process is transparent (i.e., free from pretence or deceit). Such transparency is necessary to the effective fulfillment of the preserver’s role as a trusted custodian of the records. Documenting the reproduction process and its effects is also important for the users of records since the history of reproduction is an essential part of the history of the record itself. Documentation of the process and its effects provides users of the records with a critical tool for assessing and interpreting the records.

B.3 Archival Description
Traditionally it has been a function of archival description to authenticate the records and perpetuate their administrative and documentary relationships. With electronic records, this function becomes critical. Once the records no longer exist except as authentic copies, the archival description is the primary source of information about the history of the record, that is, its various reproductions and the changes to the record that have resulted from them. While it is true that the documentation of each reproduction of the record copies may be preserved, the archival description summarizes the history of all the reproductions, thereby obviating the need to preserve all the documentation for each and every reproduction. In this respect, the description constitutes a collective attestation of the authenticity of the records and their relationships in the context of the fonds to which the records belong. This is different from a certificate of authenticity, which attests to the authenticity of individual records. The importance of this collective attestation is that it authenticates and perpetuates the relationships between and among records within the same fonds.

---

28 Although, technically, every reproduction of a record that follows its acquisition by the preserver is an authentic copy, it is the only record that exists and, therefore, should normally be referred to as “the record” rather than as “the copy.”
Appendix 3

Appraisal of Electronic Records: A Review of the Literature in English

Appraisal Task Force
Prepared by Terry Eastwood, chair, Appraisal Task Force
Contributors
Shadrack Katuu, research assistant, University of British Columbia
Jacqueline Killawee, research assistant, University of British Columbia

30 May 2000
Purpose and Scope
The purpose of this paper is to review the literature in English written on the subject of appraisal of electronic records in order to refine the research questions in Domain 2 of the InterPARES Project. As a preliminary measure, those questions have been articulated as follows:

- What is the influence of digital technology on appraisal?
- What is the influence on appraisal of retrievability, intelligibility, functionality, and research needs?
- What are the influences of the medium and the physical form of the record on appraisal?
- When in the course of their existence should electronic records be appraised?
- Should electronic records be appraised more than once in the course of their existence, and, if so, when?
- Who should be responsible for appraising electronic records?
- What are the appraisal criteria and methods for authentic electronic records?

The final question expresses the overall aim of the work in this domain. That broad question needs to be refined into more specific questions around which pointed investigations can be made. It should also be noted that our work does not address appraisal criteria specific to juridical systems and cultures but rather how to conduct appraisal to ensure that records of enduring value are actually preserved in authentic form. This review aims to begin the process of doing that. Once the questions are agreed upon, the intention is to review policy and procedure documents detailing existing practices in a select number of institutional settings where there has been a significant experience of actually appraising electronic records for long-term preservation. Following upon the completion of the case studies of electronic records undertaken by the Authenticity Task Force of the project, which will provide an empirical base for analysis of appraisal considerations, models of the entities and activities involved in appraisal and a related glossary will be developed. When the knowledge gained in these investigations is done, it will be combined with the knowledge achieved in the work to develop a typology of electronic records and in the study of storage media to develop methodologies and strategies for appraisal of particular classes of electronic records. At this stage, it is an open question what those methodologies and strategies should be.

Because the aim is to refine the research questions, the authors did not attempt to provide a comprehensive review of everything said about the subject, but rather to summarize the important issues. They concentrated on literature specifically on appraisal of electronic records, although some reference is made to the general discussion of management issues for current electronic records. References to the literature are restricted to marshalling some measure of support for the refinements suggested.

Organization of This Paper
This paper is organized into the following sections:

- Changing Environment of Electronic Records. This section aims to characterize how the computing environment has evolved over time and what implications this has for appraisal.
- Factors Affecting the Archival Management of Electronic Records. This section aims to characterize certain of the most prominent factors that have affected the ability of archival institutions to appraise and preserve electronic records.
- Tactical and Methodological Issues and Questions. This section addresses three questions: when should electronic records be appraised, what should be appraised, and who appraises?
- Technical Analysis. This section reviews what the literature says about evaluating technical aspects of electronic records.
• Content Analysis. This section reviews what the authorities say about evaluating content.
• Summary of Conclusions.

Changing Environment of Electronic Records
The Committee on Electronic Records of the International Council on Archives (ICA) summarizes the change in the technological environment in these words in its Guide for Managing Electronic Records from an Archival Perspective.

The evolution of information technology falls into three overlapping phases: the mainframe era, the era of the personal computer (PC), and the networking era. Each succeeding innovation in information technology made new uses for information technology feasible without necessarily displacing older systems. Depending on when computers were introduced into an organization, archivists may encounter electronic records that were accumulated under any of the phases.\(^1\)

In the first phase, “data were entered into the computer system, processed in batches, and then output was used in summaries, bills, accounts, and other business documents or in reports and analysis of scientific research.” For the most part, one administrative entity purchased the expensive hardware required, another programmed it, and yet another decided “the tasks that lent themselves to automation.” The ICA committee concludes that:

The prevailing view of electronic records at the time was they were special media records which were primarily valuable because of their informational content while records that were needed for evidence of actions and decisions were printed on paper and stored in established filing systems.\(^2\)

In the second phase, beginning with the introduction of the first personal computers in 1981, computing rapidly decentralized as action officers acquired their own computers and used them, particularly for word processing. The ICA Guide notes that “the rapid proliferation of text and data files” made inventorying, appraising, and preserving electronic records difficult, and turned archivists’ attention to the question of developing policies and practices to ameliorate this decentralized and uncontrolled situation.\(^3\)

“The next significant advance in computing,” the Guide says, “began in the mid-1980s with the rapid integration of telecommunications and computing into vast computer networks.” Mainframe computers still handled large databases and highly complex operations, but the client-server approach allowed organizations to combine “the autonomy that the PC offers with some of the central controls of the mainframe environment,” with important implications for records creation. As the Guide puts it,

With the growth of networking and the development of paperless transactions, archivists have become increasingly concerned about the long-term preservation of electronic records. These new archival concerns arise out of both the capabilities of the new technologies and the ways in which these technologies are being used in organizations.\(^4\)

These changes in the application of computing to work processes have had a profound influence on the way organizations operate. The result has been “changes in workflow, communications, and formal organizational structures” that can “affect the provenance, ownership, and physical location of records.” Many organizations are rapidly working towards a situation in which “electronic records become the most complete evidence of the business process and paper records begin to function as convenience copies.”


\(^2\) Guide, 14.

\(^3\) Guide, 15.

\(^4\) Guide, 15.
In this rapidly evolving environment, “archivists have been driven to examine a broader set of records management issues in order to carry out the archival function in the digital environment.”

The Guide also discusses the problem of technological obsolescence. Both hardware and software have a relatively short life.

Organizations replace their systems when their supplier ceases to support an obsolete system or when new products offer advantages over older software. To ensure that records created in the old system will remain available, understandable, and usable to users of the new system, the organization must migrate its older records to the new system. Transferring records from older proprietary systems—called legacy systems—to current technology may require substantial reformatting and restructuring of record. As long as information technology continues to evolve and organizations find new ways to apply computers to information handling and communications, archives will have to be prepared to offer advice and guidance in a dynamic environment.

Elsewhere, the Guide says that “in order to preserve electronic records, they must from time to time be migrated to new technological platforms (i.e., be copied to new storage devices and in some cases converted to a format suitable for new computer systems.)” However, a major issue for the research is whether and when conversion or migration will be necessary, and whether and when other strategies of long-term preservation may be appropriate.

The implications for appraisal of this situation are many. First, because the products of the various phases differ greatly, somewhat different approaches may be needed for each. An objective of this part of the research should be to detail the ways in which these approaches need to vary. This would constitute one of the ways in which the technology has an influence on appraisal. Of course, it may be that there is no fundamental difference, only a different strategy. This remains to be seen. Certainly, it is an important task to distinguish the common elements of appraisal of electronic records from the particularities in given classes of cases, assuming that those classes can be identified and characterized. It is also clear that the problem of technological obsolescence impinges on appraisal of records for long-term preservation as it does on everything else to do with electronic record keeping. What is not clear are the methods to be employed in given cases for long-term preservation and what implications this decision-making has for appraisal.

### Factors Affecting the Archival Management of Electronic Records

Almost all writers on appraisal of electronic records begin, as Harold Naugler did in his ICA Records and Archives Management Program (RAMP) study, *The Archival Appraisal of Machine Readable Records*, published in 1984, by identifying “a number of factors which could have a major impact on [appraisal] of electronic records.” He identifies them as follows:

- Legislation may prevent or inhibit archives from acquiring electronic records.
- Data held by an agency might belong to another body.
- The data may be encumbered by contractual agreements.
- Source agencies may have poor data management programs.
- It is difficult to schedule records after systems are designed and implemented.
- Archivists and records managers are not trained to appraise electronic records.

Some of these factors are obviously connected with the trends and developments already discussed, but others need elaboration. The legal issues identified by the ICA Guide are:

• the legal definition of a record, especially when it does not encompass records in electronic form;
• laws that do not accept electronic records as legitimate evidence in legal proceedings;
• legislation that defines the role of archives strictly as a custodial one;
• laws and policies that impose long waiting periods before the archives can appraise records or influence their disposition;
• legislation governing privacy and access to records;
• alienation of [public] records from public oversight. 9

The point here is that these legal impediments often make it impossible for archival institutions to conduct appraisal to select and acquire electronic records.

The question of ownership and provenance of records has, if anything, become more complicated since Naugler wrote. In today’s world, as the ICA Guide makes clear, “powerful new networks provide rapid communications and make it possible to share information across geographical boundaries as well as across organizational hierarchies.” 10 Careful analysis of these contextual circumstances, including contractual arrangements, will have to be taken into account in appraisal. The template for analysis developed by the Authenticity Task Force takes into account the various elements of context. The result should be case studies that produce a rich sense of the contextual factors that need to be taken into account during appraisal.

Perhaps by far the greatest concern of archivists has been about records and information management practices as they relate to electronic records. By the late 1980s and early 1990s, archivists were beginning to see that their attempts to integrate electronic records management into the traditional pattern of records inventory, appraisal, accessioning, preservation, and reference were not working, for a variety of reasons not necessarily associated with the applicability of those patterns. These concerns were expressed from many quarters. The experience of the State Archives of New York, as communicated by Margaret Hedstrom and her colleagues in numerous articles, represents them fairly well. As Alan Kowlowitz argues, “the most pressing issues facing electronic records appraisal today are not narrowly technical and methodological but broad program development and information management issue.” He also observes that “progress in addressing these issues has been glacial” in organizations and agencies. 11 Hedstrom and Kowlowitz concluded from their experience that

• long-term preservation depends on improved records and information management programs and on a clear statement of the archives jurisdiction in the matter;
• an integrated system for managing electronic and hard copy records is needed on an organization wide basis;
• schedules have to be developed at the time of design of systems;
• the archival authority needs more resources to tackle electronic records problems;
• policies and procedures must be in place to regulate the making and receiving of records by archives officers. 12

In short, archivists have had to concentrate on getting organizations and their various arms to integrate electronic records management concerns into the broader picture of both records and information management. They have also been concerned to develop a clearly understood role for themselves in the process and to convince the powers that be to devote sufficient resources to the archival task. In many cases, the situation Hedstrom and Kowlowitz describe explains why so few institutions have actually appraised electronic records, data, or information.

10 Guide, 16.
Tactical and Methodological Issues and Questions

Much of the archival literature on electronic records in general and the specific literature on appraisal are concerned with questions about the desirability and nature of the involvement of archivists in design of systems for generating and keeping current electronic records. As the ICA Guide observes, “it seems less clear in the electronic environment that the record creator can be relied upon actually to create a record.” Archivists have therefore looked to insinuate themselves into the design stage of electronic systems, to a time before it has been traditionally assumed the life cycle begins. This early involvement is justified on more than the grounds of appraisal, of course, but it has also been supposed that, as the Guide says, “retention requirements based upon archival considerations should be built into an electronic system at the time of its design.” The ICA document observes that this requirement “suggests that new approaches to appraisal and selections tasks may be warranted,” but that they should be “directed toward the functions of the originating body, the business processes and activities through which those functions are carried out, rather than towards the records themselves.”

Even though there is general agreement on the need to situate appraisal in this manner, it is useful to review some of the discussion, for it raises important issues.

When should electronic records be appraised?

Early in the debate about appraisal of electronic records, Trudy Peterson recognized that with “records of the new technology” the potential to lose information was an aspect of computer systems and thus a practical concern for archivists, if not a theoretical one. She says that “we all know that paper records are lost because records creators throw them away, but it normally takes a certain amount of decision making to haul files from a file drawer and dispatch them to the trash. With machine-readable files, however, the elimination of records may be built into the system.” Because a complete view of the record creation process may not be possible if one “asks for a yearly cutoff of . . . [a] file, all you will get will be a ‘snapshot’ of the operation at the time of cutoff.” As a consequence, she further argues that if “the archivist wants to maintain the records of stages of a project, he must work with the computer programmers to capture it all.”

Peterson is speaking of the kinds of databases or data files common in the first phase of computing. Despite what she says, there are important theoretical issues of whether such databases or data files contained records or not, with important implications for appraisal. In particular, Peterson’s suggestion that the aim is to capture “a complete view of the record creation process” is problematical. There is in fact no way to “capture” a dynamic database, but neither is there a way to capture a “complete view of the records creation process” except insofar as the records selected for retention give it. It is one thing to ensure that records in the electronic form are set aside and controlled properly so that they will be there and can be managed throughout the various stages of their existence. It is quite another to go into the system and extract information ex post facto for preservation purposes.

Catherine Bailey discusses the viability of the life-cycle model for electronic records appraisal in her article "Archival Theory and Electronic Records." She argues that the difficulty with the life cycle concept rests with the common identification of active, semi-active, and inactive with "physical state or activity." With electronic records, the method of storage on computer systems makes the traditional view of the life cycle difficult to apply. It is necessary to view the life cycle in a fashion that will facilitate the scheduling

---

16 Peterson, 386.
17 Peterson, 386.
and appraisal of electronic records. 19 Like Peterson, she argues that archivists will have to intervene early in the process:

They cannot wait until inactive electronic records are offered to them for appraisal, as they might have for paper records; too many computer records have vanished by then, and the documentation necessary for their proper appraisal has been lost, destroyed, or is hopelessly outdated. The sheer volatility of electronic records should be a powerful inducement for archivists to accept increased involvement in the scheduling process, beginning at the systems design stage. Again, however, this is not an issue of new or revised theory or principle, but merely one of timing and strategy.20

She sees an even more serious problem in the growing trend towards data resource management in which many entities “combine their resources to create and maintain a single large system or database which can serve all their diverse but related needs at once,” as Bailey writes.

Electronic information then becomes so fluid that not only does it become difficult to determine the active, semi-active, and inactive stages of records, but it also becomes next to impossible to determine the provenance of records. There is no longer a single application on which to focus attention. . . . Where [and when, we can add] do archivists begin to schedule the contexts of these shared databases? Can they legitimately break them down into smaller units . . . or will this act destroy the true nature of the system? Or will such a system require a scheduling technique completely different from that of the system overview?21

Bailey then outlines a three-step appraisal methodology framed within a life-cycle concept. The first step would require “a greater emphasis on the appraisal of computerized information as soon after its creation as possible,” presumably by some method akin to scheduling. In the second step, or stage of appraisal, “if a machine readable record has already been assessed as being valuable in the first stage of appraisal, then it will be necessary to separate it from the non-essential records around it and much time and energy will be saved.” The third stage outlined is, in essence, a reappraisal step. Bailey reasons that because “records can conceivably lose their value, data files should be reappraised occasionally to ensure that their archival values have not been overemphasized.”22 It should be noted that, like many authors addressing the subject, Bailey considers that the first stage in the appraisal process is most important, because there is no guarantee that all electronic records will survive until the second stage when inactive records are appraised.

However, many of the writers who argue for new appraisal methodologies emphasize the need to abandon the traditional life-cycle concept in favour of the continuum approach to records management. Glenda Acland argues that within the traditional life-cycle approach, the archivist is positioned at the end of a process, and can apply traditional archival theories only to what is passed on by the creator. This is "a passive role, an accepting role," and "the archivist is the undertaker who then acts as keeper for selected 'permanent' material, the selection often being de facto as well as archival." Acland, in fact, asks whether "the management of current records is simply the first stage in archival methodology or whether the archival concern, fundamentally the requirement to preserve permanently valuable records, is merely the first step in a comprehensive records management process." Clearly coming out in favour of the latter, she argues that "the split between the records management and the archival phases of record keeping is

19 She suggests that "the answer to the question lies in treating the life cycle model on a more conceptual level. If archivists consider the life cycle as an abstract expression of the legal authority over a record rather than a designation of its physical state or activity, then the differences between a paper record and electronic record disappear. It does not matter whether a record is located on a disc pack in an organization or department, on storage tapes in a record centre tape library, or on tapes or disks in an archives; its administrative and, especially, legal status is still determined by the amount of use it gets and the jurisdiction that controls it." Bailey, 183.

20 Bailey, 184.

21 Bailey, 184.

22 Bailey, 186.
no longer an acceptable alternative, it is no longer sufficient to exclude archivists from an active role in the process of data or information management.²³

Greg O'Shea, one of Acland's Australian colleagues, argues similarly that "the need for archivists to intervene in the records creation process has never been stronger than it is with electronic records.²⁴

It is precisely at this last point that the principles and practices of Archives and Records Management merge. The need to adopt this interventionist approach at the very outset of the records life cycle, which for electronic records is the systems development phase, in order to preserve the archival record finally kills the notion that archivists are passive spectators at the genesis and over the formative years of the life of the record.

Essentially, O'Shea argues that appraisal decisions will have to be built into the system before the records are created. He suggests that "archivists in the appraisal process for electronic records now need to specify [which] records are [to be] kept."²⁵

According to O'Shea, this involves working closely with information technology managers "who will (a) physically capture the records and (b) develop or redevelop systems to ensure that records are identified and retained for the appropriate period of time."²⁶ The fact that archivists work with systems design experts requires a shift from appraisal of the record to appraisal at the logical level, "i.e. the high-level diagrammatic representation of the system where it is relatively easy to see what functions the systems manages and where records may be kept."²⁷ Through this high-level analysis, O'Shea argues, records worthy of preservation can be identified before creation, and the retention of records can be built into the system.

Another Australian, Michael Hoyle, speaking in the context of a particular case—involving appraisal of reports on cash transactions tendered to a special agency supporting the work of tax authorities—questions how much can be done at an early stage in the development of a system. "It seems that it would be more productive for the Archives to have an advisory role at the early stage. . . . Rather than taking a detailed appraisal . . . perhaps an overview could be prepared . . . with a view to assessing the system's acceptability in terms of the Archives Act." Later, when the system has matured and action officers have a better understanding of its uses, appraisal can be undertaken.²⁸

Charles Dollar also urges archivists to become involved in information systems design to ensure that appraisal concerns are met. "From an archival point of view, the appraisal and retention functionalities should be incorporated into the design of information application systems in order to ensure the identification and retention of records of continuing value." He goes on to say that "one of the most useful contributions archivists can make to information systems design is to incorporate into it the concept of the life cycle management of recorded electronic information." However, Dollar notes that archivists have not done enough to analyze the life-cycle concept in a way that it can be adapted to the electronic


²⁴ Greg O'Shea, "The Medium is not the Message," in Keeping Data, 76.

²⁵ O'Shea, 88.

²⁶ O'Shea, 77.

²⁷ He argues further that "functional/logical level appraisal, is seen as producing simple, integrated and non-redundant definition of the permanent records that is independent of frequent system and software changes." O'Shea, 77.

environment. “Consequently, archivists have not articulated clearly the functional requirements of the life cycle of recorded information that could become part of the design of a complex information system.” 29

More recently, Hans Hofman has argued similarly that, in establishing the groundwork for managing archival records, archivists need to take an integrated approach to the management of electronic records. More specifically, he argues for a management regime based on three interrelated factors or layers. The first layer is "an intellectual infrastructure for inspection, appraisal and intellectual control"; the second "a technological infrastructure for records creation, preservation and research/service delivery"; and the third an organizational infrastructure to facilitate the carrying out of the first two structures. This framework must encompass all agencies. 30 As such, the archives should be involved in the management of records at all stages of the life cycle:

The ideal situation would be if archives [institutions were] involved from the moment that electronic records are created or (even better) when the information system is conceived and developed. This would only be necessary for those records that are of archival value. To know this, the archives have to develop an appraisal method that allows them to determine this as early as possible. 31

An important discussion that has emerged from the appraisal debate, particularly as it relates to life-cycle/continuum concepts, is the continuing relevance of permanent value as a concept in the modern record environment. According to Acland:

Should Archivists “select for permanent retention” as we have all be schooled or “appraise and eliminate” with a shift in axis to the determination of continuing, rather than permanent, value. To the corporate archivist frequently falls the responsibility for determining continuing value because of the direct and integrated relationship that exists with the creators and major users of the records and because they may subsequently be expected to conjure up information or evidence required by their organization on request, irrespective of physical custody or even time lapse.

Acland then goes on to say that "the strength of an integrated corporate archival appraisal program based on continuing value is that it combines systems analysis with cost-benefit efficiency." 32 Kowlowitz agrees. In reference to the United Nations Advisory Committee for Coordination of Information Systems (ACCIS) report, he observes that "appraisal must become a flexible and continuing activity suited to an ever changing automated environment . . . [and] archivists should appraise electronic records in terms of their continuing value rather than their permanent value and that records be reappraised at the time the data is migrated to new media and software environment." 33

This discussion raises several questions about (1) the timing of appraisal, (2) the procedures or methods of appraisal, and (3) its aim. From the discussion, we may suggest some refinements to the research questions:

- Does the life cycle of electronic records differ from that for traditional records?
- When and how should the various classes of electronic records be scheduled?
- Do schedules consider only primary value or both primary and secondary value?
- Is secondary value considered only at the time records become inactive?
- Should electronic records be reappraised, if they are to be converted or migrated?

---


31 Hofman, 6.

32 Acland, 116.

33 Kowlowitz, 37.
What is to be appraised?

Hofman argues that the only sound methodology for electronic records is functional appraisal. Discussing conclusions of the Dutch Project for Implementation Reduction Transfer-period (PIVOT), he states:

> The nature and mass of electronic records make it necessary to approach them from a higher, more abstract level. In other words, it is not the records themselves that need our first attention, but the context in which they are created. In the Netherlands such a method is being developed by PIVOT. . . . The basic principles of this are: identify the spheres of government activity, the organisations involved (the “actors”) and their functions. Based on this overview the functions are appraised.34

This functional approach to appraisal attaches value to the various activities in which the creator engages, and allows appraisal to be conducted across the organizations' spectrum of activity and for large volumes of records, without necessarily engaging in a detailed examination of every records series or system.

The Australian Archives has adopted a similar approach. According to O'Shea, it has focused on developing appraisal methods based on the context of records creation rather than the records themselves. He outlines the three main reasons why functions need to be examined.

> Firstly, the Archives' primary responsibility is to select and preserve archival records. Secondly, the resources devoted to the exercise must produce the most worthwhile outcome in terms of identifying the records with the highest values. Thirdly, it has been recognized, from experience, that a significant proportion of most records and data on systems will be of temporary value. Because of these three factors, agency functions and recordkeeping systems need to be examined at the broadest level. From that point the activities and processes employed to manage these functions are examined in more detail and the values of the records created as a result determined.35

O'Shea argues that the logical extension of this principle implies that the archivist determines which records need to be preserved before they are created. "In the electronic environment, because the content, context and structure are not self evident, experience has led to the conclusion that it is imperative to specify which records are to be captured. As a consequence, to enable the records to be physically selected, more specific details about what data might be needed to make the record needs to be provided linked to good descriptions of the functions to which they relate." 36

The National Archives of Canada also developed a functional approach to appraisal. In this approach, according to Terry Cook, the first and most important question in appraisal concentrates on identifying the functional responsibilities of the person creating the records. Who, he asks, "would have had cause to create a record, what type of record would it be, and with whom would that corporate person cooperate in either its creation or its later use."37 This focus on the function behind the creation of the record leads to a top-down appraisal strategy. According to Bailey, careful functional analysis provides archivists "with an understanding of the numerous factors which will influence their examination of the physical records." 38 In the Australian, Canadian, and Dutch approaches, the emphasis on functional appraisal is meant to provide a practicable means to appraise the large volumes of twentieth-century records in organizations such as governments in which there many functional interrelations.

---

35 O’Shea, “The Medium is Not the Message,” 77.
36 O’Shea, 76.
38 Catherine Bailey, 94.
Cook has also argued that appraisal of electronic records should not be treated as a special project, but rather as part of a strategic acquisition policy that follows traditional rhythms of analysis of the mandate, functions, activities, and record-keeping procedures of all agencies of the organization. This approach, he says, “can only succeed, however, if the organizational and intellectual distinctions between textual (paper) and data (electronic) archivists are obliterated, as well as those between textual and electronic analysis.”

The approach of the National Archives and Records Administration of the United States (NARA) reflects the traditional practice of appraising the record rather than the function. In discussing the work of a task force looking into the appraisal of federal databases, Ken Thibodeau notes that one aspect of the project was to “identify databases with long-term research value.” The task force used a number of experts to establish appraisal criteria based on informational value. It did not use the method of functional analysis. Linda Henry, an archivist at NARA’s Center for Electronic Records, considers the NARA approach sound. In fact, she warns that appraisal by function may be dangerous:

“A professional archivist could easily find this approach troublesome or unworkable. For example, one important function of the U.S. Patent and Trademark Office (PTO) is granting patents. NARA appraised the important electronic patent records a few years ago. In 1996 PTO submitted schedules for 54 additional electronic systems. The appraisal archivist could have considered only function—an important one—and not have looked at the records, presumably appraising all 54 as permanent. Instead the archivist considered the content of all the databases and appraised only one as permanent.”

Henry further argues that archivists “can give advice about creating and managing reliable records”; however, “if archivists usurp the role of creator by defining what records should be created, archivists make records ‘less genuine, less authentic.’” It is very likely, however, that part of the reason for this apparent divergence is that NARA was appraising databases, rather than the kind of record-keeping systems assumed by the other authors. This only points out the need to situate discussion in terms of classes of electronic records that can be assumed to have similar characteristics in different juridical contexts.

Luciana Duranti has also observed that problems develop when archivists attempt to build appraisal decisions into systems before records are created. In reference to the ACCIS report, she argues that building systems that establish which records need to be captured implies that “such an appraisal decision is to be made item by item.” The ACCIS report is not explicit about how this is to be done, or by what criteria. Rather it attempts to facilitate this approach by redefining the record as a business transaction. Duranti observes that the very act of distinguishing those records that are recorded transactions from those that are not is in and of itself an appraisal decision. As she puts it, “something the fact that a piece of information is identified as a ‘record transaction’ means that it must be retained, and indeed, throughout the report there is the sense that the decision that an entity constitutes a record is an appraisal decision.” She notes that there are both difficulties and unresolved ambiguities with this concept when she asks: “On which basis can one segregate a record from a non-record?” Unfortunately, as Duranti notes, little investigation had been conducted in this area despite its obvious significance.

This discussion poses the following questions.

42 Henry, 319.
• Does functional appraisal provide a solution to the determination of value criteria? If so, what is the precise methodology involved?
• On what basis does the archivist decide that certain functions are worth documenting and others not?
• Is appraisal responsible for determining which "recorded transactions" are to be preserved?

Who Appraises?
Another important question raised by current reconsideration of appraisal theory is who does the appraisal at each stage. Linda Henry observes that the records continuum approach tends to blur the distinction between archivists and records managers. Whereas “the traditional life cycle delineates clear responsibilities to creators and records managers for the primary value of records and to the archivists for the secondary value,” in the continuum model “archivists hold responsibility beginning before creation, through maintenance, preservation, and use.” In a sense, then, the question becomes not who is in charge of appraisal but what an archivist is in the electronic world. As Edward Higgs says:

The role of the archivist would, therefore, lie in ensuring that the suitable archival principles are embedded in computer systems at the design stage, ensuring intellectual control, and providing gateways to electronic information. In addition, archivists might cooperate with historians in designing search engines to locate and contextualize relevant records via networks. The archivist appraising, selecting, and listing documents, and placing them in published guides would be a thing of the past.

This discussion suggests the following question.

• Does the role of the archivist/archival institution change in the appraisal of electronic records, and, if so, how?

Technical Analysis
In his RAMP study, Harold Naugler observes that “machine readable records cannot be appraised solely for their content. They must also be examined in terms of their technical requirements.” At the time Naugler was writing, the main technical issues were:

1. Are the materials readable by a computer? This problem, of course, is related to the durability of the medium and to the problems created by the rapid rate of technological change, but, in fact, unreadable electronic records cannot be appraised.
2. The adequacy of documentation is vital. This was a particularly serious problem in the first phase of computing, when a record of programming decisions was necessary to understand the data.
3. Each potential accession had to be assessed as to the internal structure of files and the degree of dependence on hardware and software, and then a determination made as to whether the data could be preserved in that format or needed to be transferred to a standardized format.
4. Each potential accession had to be evaluated considering the cost of preservation and the benefits of preserving the data for continuing research purposes.
5. Certain servicing implications had also to be taken into account because the complexity of the records (or data) and their format affect service to users and the cost of reproduction.
6. In some cases, privacy or confidentiality considerations may require providing a “public use” version of non-restricted data. The cost and viability of this had to be taken into account.

Naugler then identifies a number of other issues:

1. the problem of confidentiality of personal information;
2. the implication of exchange of data across national borders;
3. the viability of sampling electronic records or data; and

---

45 Henry, 318–319.
4. the question of whether initial appraisal decisions need to be reviewed in the light of the cost of continuing maintenance and use.

In the second period of computing, the kind of technical analysis Naugler outlines gave way to systems overview in a first phase of analysis. Greg O’Shea characterizes it as a gathering of information on:

- the title of the system or application
- purpose of it
- an overview of the subject content of the data
- an overview of the major stages of data flows
- the number of logical records or units of measurement associated with the application
- background on its development
- cross references to documentation elsewhere
- data collection procedures.

In a second phase, the following are assessed:

- ability to manipulate the data/records (usually now referred to as functionality)
- level of aggregation of the data in the system
- whether the records themselves can be accessed
- internal arrangement of the data in the system
- frequency at which the data are replaced
- software and hardware of the system/application
- physical condition of the medium
- usability if the data are in their current state
- quantity of material versus its long-term costs of maintenance

The more recent literature avoids discussion of the details of technical analysis. Given that many of the writers on this aspect derived their criteria for technical analysis from Naugler, the questions are:

- Which technical aspects of electronic records need to be taken into account during appraisal?
- How do these aspects vary depending on the type of electronic record?

**Content Analysis**

Naugler uses the traditional notions of legal, evidential, and informational value. The main question as to legal value at the time of his writing was whether electronic records could be admitted to court proceedings. He passes over evidential value without much comment, and concludes that “the main appraisal judgement” concerns informational value, in which the main considerations are:

- the uniqueness of the information
- the importance of the information
- the degree to which researchers can manipulate the information
- the level of aggregation
- the potential for linkage with other data through common identifiers.

He then distinguishes the types of data found in computer systems according to purpose/function as:

1. administrative or housekeeping data
2. personnel data
3. supply data
4. financial data
5. project management data
6. operational data
7. measurement (or instrumentation data)
8. license data
9. survey data
10. registry data
11. automated office information (correspondence, reports, memoranda, and other documents stored in electronic form). Naugler discusses the factors to be taken into account in each case, in order of importance. For instance, for registry data, the factors in order of importance are: "[the character of] the activity registered; the individuals or events being registered; the number of variables of information provided."

These kinds of criteria recommended themselves when it was mostly a matter of evaluating the continuing research utility of data. More recent literature is relatively silent on content analysis. Much of the discussion has given way to consideration of the value of the functional approach, as reviewed earlier in this paper. As archivists recognize that they are in fact dealing with records in electronic form, there seems to be no need to discuss special problems of content analysis such as were considered earlier on.

However, it is evident from Naugler’s discussion of appraisal of the various classes of data that it will be necessary to discuss appraisal of the various classes of electronic record that exist today, for each of them will present special issues. The work in Domain 2 will therefore be instrumentally assisted by the work on an electronic records typology. Therefore the question is:

- Is there any difference in assessing the content of electronic records as compared to traditional records?

**Conclusion**

The literature on the appraisal of electronic records raises many important issues, some of which are relevant to appraisal of records in any medium and form, and some of which relate to the overall strategy and tactics of appraising electronic records in a difficult environment. It is evident that the work of the Appraisal Task Force is primarily to identify the particular issues that apply to long-term preservation of authentic electronic records. To do this, the task force needs to do two things. First, it needs to model the process of appraisal to identify the various activities involved in selection and acquisition. Then it needs to use the results of the work of the Authenticity Task Force in developing a template for analysis, case studies, and a typology to identify the specific issues relevant to appraisal of electronic records.
A Model of the Selection Function

Appraisal Task Force

June 2001
### i) Model Diagrams

**MODEL INFORMATION**

<table>
<thead>
<tr>
<th>TITLE</th>
<th>Select Electronic Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHOR</td>
<td>Appraisal Task Force, InterPARES Project</td>
</tr>
<tr>
<td>MODEL TYPE</td>
<td>IDEF(0) function model. IDEF(0) (Integration Definition for Function Modelling) is a U.S. Federal Information Processing Standard (Publication 183, as issued by the National Institute of Standards and Technology). &quot;A function model is a structured representation of the functions, activities or processes within the modelled system or subject area.&quot; See <a href="http://www.idef.com">http://www.idef.com</a> for more information.</td>
</tr>
<tr>
<td>PURPOSE</td>
<td>The purpose of this model is to define the activities involved in selection of authentic electronic records for long-term preservation.</td>
</tr>
<tr>
<td>VIEWPOINT</td>
<td>The entity (archival institution or program) responsible for long-term preservation of electronic records of an organization (government, corporate body, or institution).</td>
</tr>
<tr>
<td>SCOPE</td>
<td>Covers all the activities conducted by the preserver in appraising and carrying out disposition of electronic records.</td>
</tr>
</tbody>
</table>
Select Electronic Records

- Archival Science
- Legal Requirements
- Societal Needs
- Needs of the Records’ Creator

Electronic Records

- Information About Records’ Contexts
- Information About the Technological Context of Records
- Information From Electronic Records
- Information about Preservation Capabilities

- Transfer of Electronic Records
  - Selected for Preservation
  - Not Selected for Preservation
  - Information About Appraisal Decision
  - Information About Transferred Electronic Records

- Authenticity Requirements

- Computer Equipment and Software
- Facilities
- Persons
Assess Authenticity of Electronic Records

1. Relevant Information About Electronic Records
   - Compile Evidence Supporting the Presumption of Authenticity
     - Evidence for the Presumption of Authenticity

2. Need for Verification
   - Measure Evidence Against Benchmark Requirements
     - Verify Authenticity

3. Assessment of Authenticity
Determine Feasibility of Preserving Authentic Electronic Records

- Relevant Information About Electronic Records
- Valuation Information
- Information about Preservation Capabilities

Appraisal Strategies

List of Record Elements to be Preserved

Identify the Digital Components to be Preserved

List of Digital Components to be Preserved

Reconcile Preservation Req's with Preservation Capabilities

Feasibility Information

NODE: A23
TITLE: Determine Feasibility of Preserving Authentic Electronic Records
NUMBER: v.4.0
PREPARE ELECTRONIC RECORDS FOR DISPOSITION

PREPARE ELECTRONIC RECORDS FOR TRANSFER

TRANSMIT ELECTRONIC RECORDS

Disposition Rules & Procedures

Information About Disposition

Electronic Records Not Selected for Preservation

Electronic Records Selected for Preservation

Information About Electronic Records Selected for Transfer

Electronic Records Prepared for Transfer

Transfer of Electronic Records Selected for Preservation

Prepare Electronic Records for Disposition

Appraisal Decision

Updated Information About Appraised Electronic Records

Prepare Electronic Records for Transfer

Electronic Records

Carry Out Disposition of Electronic Records

A4

v.4.0
ii) Activity Definitions

A0, Select Electronic Records:
To appraise and carry out disposition of electronic records according to the continuing
needs of the creator and society, using the principles of archival science.

A1, Manage Selection Function
To establish, implement, and maintain a framework for the selection function.

A2, Appraise Electronic Records
To evaluate electronic records for the purposes of continuing preservation.

A21, Compile Information about Electronic Records
To collect, organize, and record relevant information from the electronic records and
about their juridical-administrative, provenancial, procedural, documentary, and
technological contexts.

A22, Assess Value of Electronic Records
To analyze and judge (1) the capacity of electronic records to serve the continuing
interests of their creator and society and (2) the grounds for presuming the records to be
authentic.

A221, Assess Continuing Value of Electronic Records
To analyze and judge the capacity of electronic records to serve the continuing interests
of their creator and society.

A222, Assess Authenticity of Electronic Records
To analyze and judge the grounds for presuming electronic records to be authentic.

A2221, Compile Evidence Supporting the Presumption of Authenticity
To collect, organize, and record evidence of the identity and integrity of electronic records
and about the procedural controls applied to them, to support the presumption of
authenticity of electronic records.

This is the compiling of information according the benchmark requirements. Definition
should mention identity, integrity, and procedural control.

A2222, Measure Evidence against Benchmark Requirements:
To compare the evidence compiled about the identity, integrity, and procedural controls
of the records with the benchmark requirements for authenticity.

Benchmark Requirements 1–8. Definition should mention identity, integrity, and
procedural control.

A2223, Verify Authenticity
To establish grounds for presuming the authenticity of electronic records, in cases where
there is insufficient evidence to meet the benchmark requirements, by methods of
verification such as comparing the records with copies or back-up tapes, performing
textual analysis, or examining audit trails.

See footnote 8 of Draft Requirements for Ensuring the Authenticity of Electronic Records

A223, Determine Value of Electronic Records
To establish the value of electronic records based on an assessment of their continuing
value and their authenticity.
A23, Determine Feasibility of Preserving Authentic Electronic Records
To decide whether the record elements conferring authenticity and embodying value can be preserved given the preserver's current and anticipated preservation capabilities.

A231, Determine the Record Elements to Be Preserved
To identify the extrinsic and intrinsic elements of form and the content of electronic records that need to be preserved.

A232, Identify the Digital Components to Be Preserved
Identify the digital components that manifest the record elements that need to be preserved.

A233, Reconcile Preservation Req's. with Preservation Capabilities
To decide whether the digital components manifesting the record elements that need to be preserved can in fact be preserved given the preserver's current and anticipated preservation capabilities.

A24, Make Appraisal Decision
To decide the disposition of electronic records and agree on the terms and conditions of implementation.

A3, Monitor Appraised Electronic Records
To keep track of changes to appraised electronic records or their context that make it necessary to adjust or redo an appraisal, initiate a transfer, or take some other action.

A4, Carry out Disposition of Electronic Records
To effect destruction and/or transfer of custody of electronic records according to the appraisal decision.

A41, Prepare Electronic Records for Disposition
To format and copy records selected for preservation so as to prepare them physically for transfer, or prepare records not selected for preservation for destruction, alienation to another entity, or such other disposition as has been determined in the appraisal decision.

Prepare electronic records for destruction and/or transfer of custody. The latter could include copying, extracting, reformatting, etc ...

A42, Prepare Electronic Records for Transfer
To package records selected for preservation with the necessary information for their continuing preservation, including the terms and conditions of transfer, identification of digital components to be preserved, and associated archival and technical documentation.

A43, Transmit Electronic Records
To send electronic records prepared for transfer, with the accompanying information necessary for continuing preservation, to the office responsible for the preservation function.

Sending transfer packaged with information to those responsible for the continuing preservation.
iii) Arrow Definitions

**Appraisal Decision**
A determination of the disposition of electronic records, including the terms and conditions of transfer, that has been reviewed and revised as necessary in the light of changes in the records and their context.

**Appraisal Strategies**
The rules and conventions of the entity responsible for continuing preservation that govern the appraisal of electronic records.

**Archival Science**
The concepts, principles, and methodologies governing the treatment of records, including the concepts, principles, and methodologies defined by diplomatics.

**Assessment of Authenticity**
A record stating the reasons for presuming electronic records to be authentic in terms of the benchmark requirements for authenticity.

**Assessment of Continuing Value**
A record stating the reasons for continuing preservation of electronic records.

**Authenticity Requirements**
The specification of the elements of form and context that need to be preserved in order to maintain the authenticity of a given type of electronic record.

**Computer Equipment and Software**
Hardware and software to access electronic records.

**Disposition Rules & Procedures**
The rules and procedures governing the process of the disposition of electronic records.

**Electronic Records**
A record that is created (made or received and set aside) in electronic form.

**Electronic Records Not Selected for Preservation**
Electronic records identified for destruction or disposition to an entity other than the one responsible for continuing preservation.

**Electronic Records Prepared for Transfer**
Electronic records formatted and copied for transfer and associated with the information necessary for transmittal and continuing preservation.

**Electronic Records Selected for Preservation**
Electronic records identified for transfer to the entity responsible for continuing preservation.

**Evidence for the Presumption of Authenticity**
Information that has been drawn from electronic records, from metadata related to the record, and/or from their various contexts and that provides evidence to support a presumption of the records' authenticity.

**Facilities**
Material resources need to undertake the selection of electronic records.
Feasibility Information
Information about the cost and technical capability required for continuing preservation of a given body of electronic records.

Information about Appraisal Decision
A record explaining the valuation of electronic records and the feasibility of their continued preservation, and justifying the decision.

Information about Appraised Electronic Records
A record compiled during the appraisal process containing information about the context and content of appraised electronic records.

Information about Digital Components to Be Preserved
Information about the way in which the record elements to be preserved are manifested in the electronic environment, construed for the purposes of instructing preservation activities.

Information about Disposition
Information about the quantity and characteristics of records selected for preservation and records not selected for preservation, and about the process and the cost of disposition of electronic records, utilized for management purposes.

Information about Electronic Records Prepared for Transfer
Information that spells out the terms and conditions of transfer of electronic records, and that identifies the digital components to be preserved together with the archival and technical specifications necessary to guide continuing preservation.

Information about Initial Appraisal Decision
A record explaining the initial valuation of electronic records and the feasibility of their continued preservation, and justifying the decision.

Information about Preservation Capabilities
Information about the preserver's current and anticipated capacity to preserve electronic records, including the state of preservation knowledge, hardware/software capabilities, staff expertise, and financial resources.

Information about Records' Contexts
Information about the juridical-administrative, provenancial, procedural, and documentary contexts of the records.

Information about the Technological Context of Records
Information about the hardware and software environment(s) in which electronic records were created and kept.

Information about Transferred Electronic Records
A record providing the necessary information about electronic records to maintain them continuously in authentic form, including the terms and conditions of transfer.

Information from Electronic Records
Information drawn from reading the form and content of electronic records.

Initial Appraisal Decision
An initial determination of the disposition of electronic records, including the terms and conditions of transfer.

Legal Requirements
The concepts, principles, and specific statements in law relevant to the selection of records.

List of Digital Components to Be Preserved
List of the components in the electronic environment that manifest the record elements that need to be preserved to maintain authenticity.
List of Record Elements to Be Preserved
A list of the extrinsic and intrinsic elements of form that need to be preserved to maintain the authenticity of electronic records.

Need for Verification
The need to employ methods of verification of authenticity as a result of there being weak evidence for the presumption of authenticity.

Needs of the Records’ Creator
The perceived interests of the creator served by continuing preservation of records.

Persons
People who perform the selection function.

Recommendation to Redo Appraisal
Instructions to revise an appraisal decision as a result of substantial changes in the records and their context.

Relevant Information about Electronic Records
Information that is needed to appraise electronic records.

Societal Needs
The perceived interests other than those of the creator served by continuing preservation of records.

Transfer of Electronic Records Selected for Preservation
Electronic records copied and formatted for transfer and sent to the office responsible for the preservation function.

Updated Information about Appraised Electronic Records
A record compiled during the monitoring process containing updated information about the context and content of appraised electronic records.

Valuation Information
Information about the criteria used to assess the value of electronic records and their application in a given case.
Appendix 5

A Model of the Preservation Function
version 6.0

Preservation Task Force

June 2002
i) Model Diagram Descriptions

**Purpose:** The purpose of this model is to articulate the functions, information, and resources required to preserve authentic electronic records.

**Definition:** The "Preserve Electronic Records" model describes the processes required to transmit electronic records over time and over generations of information technology and to produce authentic copies of those records. The model is articulated on the basis of the recognition that it is literally impossible to preserve an electronic record as a physical object that exhibits all the required elements of the documentary form of the record. Physically, it is only possible to preserve digital components that contain binary representations of the contents of electronic records and the information necessary to translate the content from the stored representation into the structure and presentation prescribed for the record.

**Viewpoint:** The model is constructed from the viewpoint of the person responsible for preservation of authentic records. The same person or organization may have other roles or responsibilities, such as appraisal or reference, but coincidental activities are excluded from the "Preserve Electronic Records" model.

When it is determined that a body of electronic records has long-term value, information about those records needs to be provided to the person who will be responsible for preserving them in order for the preserver to determine how the records will be preserved. Feedback from the preserver to the appraiser leads to the articulation of terms and conditions for the transfer of the records. Except in the case of a body of records that will be transferred all at once to the preserver, this feedback loop may extend in time for as long as the records creator continues to create or maintain records belonging to the body of records selected for preservation. During that time, changes in the creator's record keeping may entail changes in the terms and conditions for transfer. One term for transfer is the stipulation of when records should be transferred to the preserver. The preserver's active involvement with the records begins at that point.

The preserver also has a role in responding to requests for copies of records and for information about the records, even though the viewpoint of this model does not include the responsibility for determining access rights, providing information informing potential users about records that may be of interest to them, or otherwise communicating with users. Requests for records and for information about records held in the preservation system are directed to the preserver through the intermediary of the person responsible for reference and access. Nonetheless, the process of reproducing an electronic record from its stored digital components may fail to produce an authentic copy of the record. Therefore, the person responsible for preserving the record must exercise control over reproduction in order to ensure that the record is effectively transmitted over time in authentic form.

**Scope:** The "Preserve Electronic Records" model encompasses activities from the point where it is determined that a body of electronic records has long-term value to the production of a copy of a preserved electronic record. This scope needs to be subdivided into two parts: the first covers information about the records; the second, the records themselves. The role of preserving records includes all of the activities necessary to ensure the transmission of authentic electronic records over time. First, there is interaction with the process by which aggregates of records are selected for preservation. At this stage, the person responsible for preservation needs to determine whether the records can be preserved and given established capabilities and capacities, or alternatively to plan actions to expand those capabilities and capacities. The role of preserving extends to the delivery of copies of preserved records to end-users. At this end stage, the person responsible for preservation needs to control production of copies, because only end-
to-end preservation control—include transfer from the records creators through delivery to end-users—can ensure the authenticity of copies.

The "Preserve Electronic Records" model operates at two different levels. Some of the activities in the model—for example, Bring in Electronic Records—operate on sets of electronic records. Other activities, such as Output Electronic Record, operate at the level of individual records.

The level of detail in the description of processes in the model is limited by the fact that it is intended to apply to any and all electronic records that may be selected for preservation, and that it is intended to be valid for all preservers. Thus, the model provides for the operation of external—legal, societal, stakeholder—requirements as controls on preservation process. It assumes that the preserver is responsible to a specific institution and that external requirements are interpreted by the institution to determine how they apply to the preservation of electronic records. The model does not detail the impact of any specific institutional controls, although it does indicate where such controls impact the process. Similarly, the model does not detail the preservation of any specific record, body of records, class of records, or class of digital component of electronic records. At present, the model does not detail the operation of any specific technique (such as migration, technology preservation, or persistent object preservation) for overcoming technological obsolescence or for maintaining the capability of producing authentic copies of electronic records. However, it does delineate the process both for selecting the approach to be taken in any case and where the technique should be applied.

Source: The "Preserve Electronic Records" model was developed by the Preservation Task Force of the InterPARES Project based primarily on the knowledge of the multidisciplinary experts who were members of the task force, supplemented by information gathered from the members' institutions, by a survey of digital preservation projects, and by the products of the InterPARES Authenticity and Appraisal Task Forces.

### IDEF(0) Diagrams

**A0, Preserve Electronic Records.**

Given Information about Electronic Records Selected for Preservation, and the Transfer of those records, the "Preserve Electronic Records" model transmits electronic records over time. In response to a Request for a Record and/or Information about the Records, the process produces copies of those records and, if requested, certifies the authenticity of the copies. The preservation process can also output the digital components of an electronic record and relevant information enabling the recipient to reproduce the record from those components. If a request is only for information about a record, the available information is delivered. The process also outputs information about how the preservation function has been exercised.

The records preservation function is controlled by Archival Requirements, the State of the Art of Information Technology, Institutional Requirements. It is accomplished by Persons Responsible for Preservation, using Information and Communications Technologies in appropriate Facilities.

**Preserve Electronic Records.**

The process, Preserve Electronic Records, involves four sub-processes: Manage the Preservation Function, Bring in Electronic Records, Maintain Electronic Records, and Output Electronic Record. The first of these functions controls the other three; the other three provide feedback to the "Manage" function.

**Manage the Preservation Function.**

Managing the preservation function involves analyzing Information about Electronic Records Selected for Preservation and about Electronic Records Transferred and Accessioned Records to Determine Preservation Requirements, to Select Preservation Technologies suitable for the Specified Requirements for Preservation; Specify a
Preservation Strategy which determines precisely how and when selected preservation technologies will be used to preserve a target body of records; and related procedures and controls. These processes produce the Preservation Strategies which control all other preservation processes: bringing records in, maintaining them, and outputting them, the Technological Infrastructure necessary to carry out these processes, and the Targeted Preservation Methods applied to the objects within the scope of each Preservation Strategy. In addition, managing the preservation function entails evaluating execution of the three other preservation processes and using the results of this evaluation to improve the management process.

A1.1, Determine Preservation Requirements.
Determining preservation requirements consists of four distinct processes: determining technical specifications for transfer and for storage of records in logical and physical files, determining the specific archival properties of records that must be preserved, determining how these properties are instantiated in the digital components of the records, and producing a synthesis of the requirements articulated in these analyses. There is also an ancillary process of determining the actual basis for presuming that the records creator maintained and transferred authentic records. Its output is also factored into the synthesis of preservation requirements.

The first process considers only archival properties of records and archival aggregates, regardless of how their properties are instantiated in digital objects or what software is needed to execute the methods of each type of digital object. It determines preservation requirements that technology must satisfy and that control the use of technology in preservation. Archival properties are defined for classes of individual records, for archival aggregates, and for the presentation of records and aggregates for use.

The second of the principal processes, determining how archival properties are instantiated, is represented by and decomposed into three sub-processes. First information about classes of records is analyzed to determine how records in each class are composed of their Digital Components. Second, information about how archival aggregates in the body of records are constructed and ordered is reviewed to determine what will be required over time to reinstantiate aggregate structures, with members in proper order. Third, information about how individual records and archival aggregates are presented is analyzed in order to develop specifications for reproducing records. Each of these three steps and also the analysis of the presumption of authenticity produce appropriate requirements for preservation, which are then synthesized in the final process to produce a comprehensive and coherent set of requirements for preserving a body of records.

A2, Bring in Electronic Records.
The process of bringing in electronic records consists of the essentially mechanical step of registering receipt of a transfer, followed by verifying, on the basis of information about the transfer, whether the records purportedly included in the transfer have been selected for preservation. If so, the digital files received are reviewed to ensure that they contain the records authorized for transfer, culminating in a formal decision to accept responsibility for preserving the records.

A2.3, Examine Electronic Records.
Examining electronic records is a process of comparing the records received in a transfer to what was expected about those records, based on information in appraisal and on the documentation received with the transfer, in order to ensure that the transferred records satisfy requirements specified in the applicable Preservation Strategy, to determine if the records can be preserved, and to take any action necessary to ensure their preservation.

A3, Maintain Electronic Records.
Maintaining electronic records entails managing information about them, managing the storage of their digital components, and updating those components when necessary, as indicated by the applicable Preservation Strategy. The process outputs information about
preserved records and the digital components of requested records. If necessary because of updated preservation strategies and methods, it updates the digital components so that records and/or archival aggregates can be reproduced using the updated methods.

A3.1 Manage Information about Records.
Managing information about records entails maintaining information about the records and, on request, retrieving information about the records and, if needed, information about their digital components.

A3.2, Manage Storage of Digital Components of Records.
Managing the storage of the digital components of electronic records starts with putting into storage the components of a body of records accepted for preservation in accessioning. It is completed by retrieving and outputting digital components in response to a request. This process also entails periodically refreshing storage media before current media start to deteriorate and updating the storage system in response to changes in storage technology. This process also involves monitoring stored files to ensure that they are intact, and, if not, taking action to correct any storage problems encountered.

A4, Output Electronic Record.
An electronic record, its digital components, or information about the record are output in response to a request initiated outside of the preservation process. Each request needs to be managed in order to ensure that there is a response and that the response is appropriate. The first step in responding to an external request is to issue a retrieval request for the desired information or digital components. The response to this retrieval request is reviewed to ensure that the retrieved information or components are complete and correct. If the external request was only for information, the requested information is sent to the requester. If the request is for reproduction of a record, the method prescribed in the Preservation Strategy for the records is applied to the digital components and the reconstituted record is then either presented to the external requestor or suitably packaged for delivery via telecommunications. If specified in the request for a copy of the record, the process produces a certificate attesting to the authenticity of a reproduced copy. Alternatively, if the external request was only for digital components, the process reviews for complete and correct retrieval of those components, then packages them for delivery without reproducing the record.

The Output Electronic Record process may also be invoked when records in a transfer are being examined, in order to verify that the records in the transfer can be preserved and reproduced.
### ii) Model Diagrams

#### MODEL INFORMATION

<table>
<thead>
<tr>
<th>TITLE</th>
<th>Preserve Electronic Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHOR</td>
<td>Preservation Task Force, InterPARES Project</td>
</tr>
<tr>
<td>MODEL TYPE</td>
<td>IDEF(0) function model. IDEF(0) (Integration Definition for Function Modeling) is a U.S. Federal Information Processing Standard (Publication 183, as issued by the National Institute of Standards and Technology). &quot;A function model is a structured representation of the functions, activities or processes within the modeled system or subject area.&quot; See <a href="http://www.idef.com">http://www.idef.com</a> for more information.</td>
</tr>
<tr>
<td>PURPOSE</td>
<td>The purpose of this model is to articulate the functions, information, and resources required to preserve permanent, authentic electronic records. The InterPARES Project will use this model to identify and develop the procedures and resources required for the implementation of the conceptual requirements and criteria identified in the project's Authenticity and Appraisal research domains.</td>
</tr>
<tr>
<td>VIEWPOINT</td>
<td>Person responsible for preservation</td>
</tr>
<tr>
<td>SCOPE</td>
<td>This model is constructed within the framework established by the Reference Model for an Open Archival Information System (OAIS), which is an ISO Draft International Standard (DIS). [See <a href="http://ssdoo.gsfc.nasa.gov/nost/isoas/">http://ssdoo.gsfc.nasa.gov/nost/isoas/</a> for more information.] The ‘Preserve Electronic Records’ model includes ‘Preserve Electronic Records’ model activities and related ICOMs specifically required for the preservation and delivery of authentic electronic records. While some of these activities fall within the Ingest, Distribution and Management activities in the OAIS model, the ‘Preserve Electronic Records’ model excludes aspects of those activities not essential for preservation.</td>
</tr>
</tbody>
</table>
Determine Preservation Requirements

- Determine Transfer & Storage Requirements
  - Identify Archival Properties That Must be Preserved
  - Determine Requirements for Reconstituting and Presenting Records
- Synthesize Requirements for Preservation
  - Determine Basis for Authenticity
  - Basis of Authenticity of Records
- Archival Aggregate Requirements
- Record Preservation Requirements
- Information about Digital Components of an Electronic Record
- Information about Presumption of Authenticity of Appraised Records
- Information about Presumption of Authenticity of Transferred Records
- Types of Record Aggregates
- Classes of Records
- State of the Art of Information Technology
- Requirements for Physical and Logical Files
- Information about Electronic Records Selected for Preservation
- Information about Transferred and Accessioned Records
Bring In Electronic Records

A2.1 Register Transfer

A2.2 Verify that the Transfer is Authorized

A2.3 Examine Electronic Records

A2.4 Accession Electronic Records

Technological Infrastructure

Transfer of Electronic Records Selected for Preservation

Registration Procedure

Preservation Strategy

Targeted Preservation Method

Accessioning Policy

Notification of Receipt

Submitter

Request for Information about Authenticity

A3.1

Rejected Transfer

Submitter

Rejected Accession

Accessioned Electronic Records

Record of Accession

Accessioning Dossier

Preservable Records

Retrieved Information about Presumption of Authenticity

A3.1

Conforming Transfer

Registered Transfer

Retrieved Information about Presumption of Authenticity

A3.1

Preservation Strategy

Accessioned Electronic Records

Record of Accession
Examine Electronic Records

A2.3

Map Records and Digital Components Within Transferred Materials

Accessioning Policy

Preservation Strategy

Conforming Transfer

Verifying Records in the Transfer Can be Preserved and Reproduced

Digital Components of a Record That Cannot be Preserved

Take Action Needed to Preserve the Record

Technological Infrastructure

Preservable Records

Rejected Transfer

A4 Output Records

A3.3 Update Digital Components

Conforming Digital Components

Non-Conforming Digital Components

Request for Strategy Decision

A1

Made Records and Digital Components

Within Transferred Materials

Mapped Records and Digital Components

Verify that the Records in the Transfer Can be Preserved and Reproduced

A2.3.1

A2.3.2

A2.3.3

A1

v 6.0
A3.1 Manage Information About Records

A3.2 Manage Storage of Digital Components of Records

A3.3 Update Digital Components

- Basis of Authenticity of Records
- Retrieval Request
- Information about Accessioned Records
- Accessioned Electronic Records
- Information about Updated Digital Components
- Digital Components of Accessioned Electronic Records
- Request for Digital Components
- Updated Storage Information
- Digital Components that Need Updating

Preservation Strategy
Storage Method
Targeted Preservation Method
Method for Updating Components
Retrieved Information about a Preserved Record
Updated Digital Components

Node: A3
Title: Maintain Electronic Records
Number:
A3.2.1 Place Record Components in Storage

A3.2.2 Refresh Storage

A3.2.3 Monitor Storage

A3.2.4 Correct Storage Problems

A3.2.5 Retrieve Components from Storage

Digital Components of Accessioned Electronic Records

Updated Digital Components

Storage Method

Storage Update Method

Monitoring Method

Problem Correction Method

Retrieval Method

Request for Digital Components

Stored Digital File

Recovered File

Updated Storage Information

Refreshed File

Retrieved Digital Components

Manage Storage of Digital Components of Records
iii) Activity Definitions

A0, Preserve Electronic Records
Under the control of Archival and Institutional Requirements, and limited by the possibilities available within the State of the Art of Information Technology, preserve electronic records received in Transfer of Electronic Records. To preserve the records, the process uses Information about Electronic Records Selected for preservation. The process will also accept input of a Request for Record and/or Information about a Record and produce, in response, either the Requested Information about a Preserved Record, a Reproduced Electronic Record, or a Reproducible Electronic Record. The process will produce a Certificate of Authenticity from any Reproduced Electronic Record if requested. The process also produces Information about Preservation, including information about the process itself, the records being preserved, and how the process preserves their authenticity. Electronic records are preserved by Persons Responsible for Preservation using Information and Communications Technology and Facilities. The technology infrastructure and facilities are used in all subprocesses; therefore, they are implicit, rather than explicit, in all decomposition diagrams.

A1, Manage the Preservation Function
Operating under the control of Archival and Institutional requirements and in light of knowledge of the State of the Art of Information Technology, preservation is managed by producing a comprehensive preservation framework consisting of sets of preservation strategies and preservation action plans, each linked to a specific body of electronic records selected for preservation, along with the technological infrastructure and preservation methods needed to implement the action plans. These outputs are all used in execution of the preservation function. The “Manage” process also outputs information about the preservation function and about records being preserved and, on request, will produce a report on the authenticity of one or more records. The “Manage” process produces its output using information received from appraisal about bodies of records selected for preservation, information received from the ingest process about transferred records, and management information received as feedback from all processes involved in execution of the preservation function. Management activities are accomplished by persons responsible for preservation.

A1.1, Determine Preservation Requirements
Determine the archival requirements for preserving and reproducing an electronic record by interpreting and applying external controls on the preservation function to information received from the Select Electronic Records function about the records to be preserved, and information about transfers and accessions received as feedback from the “Bring In,” “Maintain,” and “Output” Electronic Records processes. This entails identifying the classes of objects that must be preserved, including types of records and ordered groups of records, and specifying, for each class, the attributes and methods that must be preserved, as well as the requirements for certifying that any reproduced record is authentic. Determination of archival requirements is guided by evaluation of prior experience in applying such requirements to records that have been transferred to the archives. The evaluation will not alter the external requirements, but will improve their articulation and application. The result of this process will be Synthesized Requirements for Preservation, where the specification consists in identifying what archival and institutional requirements apply to what records and how each applicable requirement is to be implemented.
A1.1.1, **Determine Transfer & Storage Requirements**
Limited by the State of the Art of Information Technology, use information about Electronic Records Selected for Preservation and about Transferred and Accessioned Records to determine how the records will be written in physical and logical files both for transfer and for storage to produce Requirements for Physical and Logical Files.

A1.1.2, **Identify Archival Properties that Must Be Preserved**
Limited by the State of the Art of Information Technology, use Information about Electronic Records Selected for Preservation—and also Information about Transferred Records when, on examination, their properties are found to be different than what had been determined in appraisal—to identify the types of archival properties that must be preserved. This activity will identify the Classes of Records that exist within a body of records selected for preservation. For each Class of Record, determine how Archival Science, the preserver’s Institutional Requirements (tunnelled to this diagram), determinations made in appraisal, and/or the records management practices of the records creator, indicate specific archival properties that must be preserved, including intrinsic or extrinsic elements of form. For the body selected for preservation, specify the Type(s) of Arrangement of records established by the records creator and identify how archival bonds are expressed. For each archival property that must be preserved, determine the parameters, measures, or other evidence that will be used to identify the property and to verify that it has been preserved intact.

A1.1.3, **Determine Requirements for Reconstituting and Presenting Records**
Under the control of Archival and Institutional Requirements (tunnelled to this activity), use Information about a body of Electronic Records Selected for Preservation, Information about previously Transferred and Accessioned Records from that body, and about Classes of Records, to identify the digital components of each class of record, how a record is composed from its digital components, and how it should be presented in order to articulate the requirements for reconstituting a record from its digital components and presenting the record in the documentary form stipulated for that class of record. The information is reviewed in light of the State of the Art of Information Technology, which indicates if current methods of composing a record from its digital components are likely to become obsolete or if there are alternative methods that might be preferable. By revealing possible technical alternatives, the State of the Art of Information Technology also indicates the extent to which archival properties depend on specific technologies. The result of this process is a set of Record Composition Requirements for each class of records, including the parameters, measures, or other evidence that will be used to verify that a record has been reconstituted from its digital components properly.

A1.1.4, **Determine Requirements for Reconstituting and Presenting Archival Aggregates**
Guided by Archival Science and Institutional Requirements (tunnelled to this diagram), use Information about a body of Electronic Records Selected for Preservation—and also Information about Transferred Records when, on examination, their properties are found to be different from what had been determined in appraisal—along with information on Type(s) of Record Arrangement to determine how to identify an archival aggregate, identify the records or other aggregates that belong in it, specify the ordering of the members within the aggregate and of the aggregate within the archival fonds, and how the aggregate should be presented for use. The information is reviewed in light of the State of the Art of Information Technology, which indicates if current methods of arranging records are likely to become obsolete or if there are alternative methods that might be preferable. By revealing possible technical alternatives, the State of the Art of Information Technology also indicates the extent to which archival properties depend on specific technologies. Define criteria that will be used to determine if records and aggregates have been properly arranged when the aggregate is reconstituted. This activity will produce Requirements for Arranging Records.
A1.1.5, Determine Basis for Authenticity
Guided by Archival Requirements and Institutional Requirements (both tunnelled to this diagram), review information collected and created in appraisal about how the records creator satisfied the Benchmark Requirements for Preservation in order to determine how the appraiser’s presumption of authenticity of the records will be documented when the records are transferred, and what impact that presumption will have on both preservation of the records and certification of the authenticity of copies produced by the preserver. Subsequently these conclusions will need to be verified and, if necessary, corrected on the basis of information about the records when they are brought in. Methods for documenting the presumption of authenticity include preserving the appraiser’s report(s) on the subject; preserving documentation produced by the records creator which provides evidence of how the benchmark criteria were addressed; and creating archival description. The information produced in this process is output as the “Basis of Authenticity of Records,” which is sent to activity A3.1 to be maintained and used to verify the established basis when records are brought in and to satisfy requests for such information.

A1.1.6, Synthesize Requirements for Preservation
Guided by Archival Science and Institutional Requirements (tunnelled to this diagram), synthesize requirements concerning physical and logical files used for transfer and storage, information about Classes of Records contained in a body of electronic records selected for preservation, Record Composition Requirements applicable to those classes, Requirements for Arranging Records, Requirements for Presenting Records, and the Basis of Authenticity of the Records as maintained by their creator, producing a comprehensive and coherent set of Synthesized Requirements for Preservation of the body of records. These requirements will guide the functions of technology selection, preservation planning, execution, and evaluation.

A1.2, Select Preservation Technologies
Take into account Information about Electronic Records Selected for Preservation and Information about Transferred and Accessioned Records in order to select Preservation Methods that will be used to preserve the electronic records. If this information indicates that existing Technological Infrastructure and one or more current Preservation Methods are applicable and adequate for preserving the records, they should be used. Otherwise, acquire Information and Communications Technology which will provide the necessary Preservation Methods and/or Technological Infrastructure. Specify the scope of applicability of each Preservation Method in terms of its application to types of digital files, target bodies of records, types of electronic records, and/or classes of digital components, along with conditions for its application to each target group, thus defining a Targeted Preservation Method. The process is controlled by Institutional Requirements, such as the institution’s information technology architecture and standards, data standards, and related procedures, security requirements, access restrictions, and performance objectives and measures. It is also controlled by the Synthesized Requirements for Preservation of a body of electronic records, applicable Authenticity Requirements for Preservation, and Evaluation of the Execution of the preservation methods that have been used to date. The selection is limited by the State of the Art of Information Technology. The scope of this process includes all methods that could impact the preservation of the records from the time of transfer. These include methods for checking the integrity of the transfer process, methods related to storage of digital components in digital files and on physical media, methods to reconstitute and reproduce records, and methods to enable others to reproduce the records.

A1.3, Specify Preservation Strategy
Guided by the Synthesized Requirements for Preservation, Specifications of the technologies selected for addressing those requirements, and Evaluation of the Execution of Preservation processes and methods, use Information about Records Selected for Preservation, and Information about Transferred and Accessioned Records
to develop a Preservation Strategy for preserving a body of records and producing authentic copies of the records. The goal of a Preservation Strategy is to ensure the preservation of authentic electronic records from the point at which they are transferred to the archival system, through the maintenance of their digital components over time, to the delivery of certifiably authentic reproductions of the records. The Strategy will specify action to take as records are brought in, maintained, and output, the methods to use in each action, and the criteria for measuring success of the action. The criteria are derived from the requirements and refined into measurable form in light of the Preservation Technology Specifications; i.e., the criteria enable measurement of how well a requirement has been satisfied using the applicable preservation method. The Strategy also includes the terms and conditions for transfer of records accepted by the creator, the appraiser, and the person responsible for preservation.

A1.4, Evaluate Execution of Preservation
Consistent with Institutional Requirements for management analysis, and guided by applicable Preservation Strategies and by both Synthesized Requirements for Preservation, analyze Management Information about Preservation in order to determine how well goals and objectives are being achieved. The analysis addresses whether records are being preserved successfully. If not, it identifies the causes of problems and determines how they might be resolved or avoided. Identify cases where rules or objectives conflict, are unclear, insufficient, inappropriate or misdirected, as well as possible improvements in operation. Assess whether the set of Preservation Strategies in effect are coherent and not conflicting. Determine if Preservation Methods are operating as intended and whether the Technological Infrastructure is adequate and appropriate. Produce evaluations of both performance strategies and the overall preservation framework. Review Management Information about the Output of Electronic Records to determine if requirements for authentic copies are being satisfied and to characterize customer satisfaction. Determine if advances in the State of the Art of Information Technology indicate that the selection of preservation technologies should be revisited. Produce an Evaluation of Execution appropriate to guide other management processes. Output Information about Preservation for external stakeholders. Produce a Report on Authenticity of Records to respond to any challenge to the adequacy and efficacy of the preservation process. This process may be triggered by a Request for Strategy Decision.

A2, Bring in Electronic Records
Following direction established in the preservation strategy for a given body of records selected for preservation, the “Bring In” or ingest function applies preservation method(s) targeted to that body of records to implement the preservation action plan for those records by processing each transfer of electronic records into accessioned electronic records. The ingest process also produces information about each transfer of electronic records, which is used in the “Manage” process to confirm or revise the preservation strategy and action plan(s) applicable to those records, and also management information which is used to evaluate execution of the ingest function. The process is carried out by persons responsible for preservation, using infrastructure technology.

A2.1, Register Transfer
Following the Registration Procedure defined by the Preserver’s Accessioning Policy, register the transfer of electronic records by capturing information about the transfer, such as submitter’s name, record creator’s name, and the date of receipt of the transfer in a Record of the Transfer, and establish basic control over the materials transferred by identifying what has been transferred and where it is located. The registration record forms the basis for identifying and tracking materials in the transfer and information about it in all other preservation processes. The registration process should also inspect what was received in order to ensure that the physical transfer has been accomplished correctly. This inspection provides quality assurance of the physical transfer, but does not address any questions related to the specific records reportedly contained in the transfer. Registration produces a Notification of Receipt, which is sent to the submitter. This
notification should describe any problem identified in receiving the transfer, such as network errors in transmission, missing media volumes, or obvious damage, and request the submitter to send a new transfer or otherwise correct the problem. The Registered Transfer is forwarded for verification that it is authorized.

**A2.2, Verify that the Transfer Is Authorized**

Acting under the Preserver’s Accessioning Policy, determine if the transfer is authorized; i.e., it comprises the records that have been selected for preservation, and those records have been submitted either by the records creator or an agent acting for the creator. Verification that a transfer is authorized is based on comparing the terms and conditions for transfer, established as part of the Preservation Strategy during appraisal, with information accompanying the Registered Transfer. This information is reviewed to determine if it indicates that the transfer was sent by an authorized person; it comprises records specified for transfer; it includes necessary information about the records, their digital components, and the basis for asserting the authenticity of the records as received; and that the materials transferred are of the correct types and in the specified formats. This verification may include steps that are specific to the records reportedly contained in the transfer whenever such specific tests are set out in a Preservation Action Plan related to the records. If the terms and conditions of transfer have been satisfied, the Conforming Transfer is passed to the next step, where its contents are examined. Otherwise either the transfer is rejected outright or the submitter is asked to address any problems identified. If a transfer is determined to be authorized, a Request for Information about Authenticity of the records is sent to activity A3.1 This request is for the information about the basis for presuming the authenticity of the records as maintained by the records creator. That information was developed in determining the requirements for preservation of the body of records. The retrieved information will be factored into the final decision on whether to accession the records.

**A2.3, Examine Electronic Records**

Acting under the preserver’s Accessioning Policy and in accordance with the Preservation Strategy established for the records reportedly included in the transfer, examine the digital files and digital components of records in the transfer, along with accompanying information to: (1.) determine if the transfer actually includes all records and aggregates of records specified in the terms and conditions of transfer and that these records and aggregates are adequately and accurately described in the accompanying information to enable their preservation, reproduction in authentic form, and interpretation; (2.) identify any actions required to preserve both the individual records transferred and the archival sets in which these records belong; (3.) initiate technical or other preservation actions that should be taken immediately and schedule preservation actions that should be taken at a later date. This examination will include any specific tests included in a Preservation Action Plan for the records. If on examination any record, digital component, or aggregate is found to have unexpected properties that would make it difficult or impossible to follow the established Preservation Strategy, or would make it questionable whether following that strategy would satisfy archival requirements, the situation should be reported to the “Manage” function for evaluation. If the evaluation results in a change in the applicable Preservation Strategy, the “Manage” function will issue a Preservation Action Plan specifying how to resolve the problem(s) identified in the examination. If the results of the examination are positive, the records in the transfer are deemed to be suitable for preservation, and the Preservable Records are forwarded to be accessioned. If the examination reveals unresolvable critical problems, the transfer is rejected and returned to the Submitter. If it might be possible for the Submitter to correct problems that the Preserver could not resolve, the Rejected Transfer includes a request for corrective action and resubmission.

**A2.3.1, Map Records and Digital Components within Transferred Materials**

In accordance with an institution’s Accessioning Policy and the applicable Preservation Strategy, using the Technological Infrastructure in place, determine how the records, their
digital components, and the information about them included in a Conforming Transfer are identified, and where they are located in the digital files and other materials received. Identify any records or components that should be in the transfer but are not found and determine if there are any records or components in the transfer that should not have been transferred. If critical deficiencies are found, terminate processing and output the Rejected Transfers. If no critical problems are encountered, output the Mapped Records and Digital Components, along with related information.

A2.3.2, Verify that the Records in the Transfer Can Be Preserved and Reproduced
Ensure that transferred records can be preserved and reproduced in accordance with the applicable preservation strategy. Determine all the records that should be in the transfer can be reconstituted and presented; if all digital components of these have been received and are in the formats stipulated in the terms and conditions of transfer; if archival aggregates established by the records creator can be re-established; if the archival bonds among records can be expressed or instantiated, and also what basis exists for asserting the authenticity of the records as transferred. Records that can be preserved and reproduced are output as Preservable Records. In the case of a record that cannot be preserved in accordance with the preservation strategy, identify the digital component(s) that prevent or impede implementation of the strategy and forward them for actions that will enable preservation of the record. Digital components modified as a result of such actions are returned for verification of the success of the modification. If it was necessary to modify the applicable preservation strategy to enable preservation of the records, the secondary verification will be on the basis of the revised strategy. The verification process will take into account any cases where digital components were not successfully modified, along with other problems discovered with the transfer, to determine whether, under the institution’s accessioning policy, the preservable records—including records deemed to have only minor problems—should be accessioned or the transfer should be rejected.

A2.3.3, Take Action Needed to Preserve the Record
When it has been determined that a record in a transfer cannot be preserved in accordance with the applicable preservation strategy, take the steps indicated in the relevant preservation action plan with respect to any digital component of the record that needs to be modified to conform with the strategy. Effect such modifications by invoking process A3.3, “Update Digital Components.” After the update, output the Conforming Digital Components. If the action plan cannot be implemented, or if it will not produce the desired result, refer the action to the “Manage Preservation” process as a request for a strategy decision. This request should trigger a revision in either the preservation strategy or action plan. When a revised strategy and/or plan is received, effect the necessary modification(s) as indicated above. If the result of a request for a strategy decision is not to modify either the strategy or plan, output the problem components as Non-Conforming Digital Components.

A2.4, Accession Electronic Records
Acting in accordance with the institution’s Accessioning Policy, formally accept responsibility for preserving a transferred body of records that have been determined to be preservable. Create a Record of the Accession and forward the Accessioned Electronic Records to the “Maintain” process. The decision to accession records takes into consideration the basis for presuming that the records, as received from the records creator, are authentic. This information uses Retrieved Information about the Presumption of Authenticity developed in determining the preservation requirements for the body of records to which they belong. The Retrieved Information is supplemented and, if appropriate, updated by any additional information either received with the transfer of the records or developed in examining the transferred records. Other elements of Accessioning Policy that should be taken into account in the decision whether to accession the records include resource constraints, where it might be determined that the
institution cannot afford to preserve the records, or anticipated demands for retrieval or restrictions on access that the institution cannot satisfy.

A3, Maintain Electronic Records
Following direction established in the preservation strategy for a given body of records selected for preservation, apply preservation method(s) targeted to that body of records to implement the preservation action plan for those records by maintaining the digital components of accessioned electronic records, along with related information necessary to reproduce the records, certify their authenticity, and enable correct interpretation of the records. This maintenance activity enables the output, in response to a retrieval request, of the digital components of a record, along with information about that record, or, if the request is only for information, the requested information. The “Maintain” process also produces management information which is used to evaluate execution of the ingest function. The process is carried out by persons responsible for preservation, using infrastructure technology.

A3.1, Manage Information about Records
The “Manage Information about Records” process collects and maintains information necessary to carry out the Preservation Strategy for a body of electronic records being preserved, including information about their digital components, the archival aggregates they comprise, their authenticity, their interpretation, and the preservation activities performed on them. Carrying out actions specified in the Preservation Action Plan, information about Accessioned Electronic Records is collected when they are accessioned. It is combined with Storage Information identifying the files, locations, and other relevant data about the digital components of the Accessioned Electronic Records when they are placed in storage and subsequently when storage parameters are changed. When a Preservation Action Plan entails any modifications to digital components, Information about those Digital Components is provided to ensure that all affected components are updated appropriately and, after the modification, Information about the Updated Digital Components is also updated. In response to a Retrieval Request for information, Retrieved Information about a Preserved Record is provided. In response to a Retrieval Request for a record, information identifying the digital components of the record and their storage location(s) is retrieved to produce a Request for Digital Components, which is used to retrieve those components from storage; Information about those Digital Components and Retrieved Information about the Preserved Record is output to support reproduction of the record and, if needed, certification of its authenticity.

A3.1.1, Maintain Information about Records
As dictated by the Preservation Strategy, Information about Accessioned Records and information providing the basis for asserting the authenticity of the records as transferred is collected and maintained over time. This information is kept up to date, by the input of Updated Storage Information, to reflect changes in the storage of the digital components of the records, e.g., when storage media are replaced or defective files are recovered. Similarly, when changes in the applicable Preservation Strategy lead to updating of any digital components, the related information is updated through input of Information about the Updated Digital Components. This activity outputs Maintained Information about a Record and Maintained Information about Digital Components.

A3.1.2, Retrieve Information about Records
Following the applicable Preservation Strategy, receive Maintained Information about Records and respond to any Retrieval Request for information about a record by outputting Retrieved Information About a Preserved Record. When the Retrieval Request is for the digital components of a record, output Information Identifying the Digital Components of a Record. Output the same information when a Preservation Strategy requires updating of the digital components of a record. A special type of retrieval request occurs when a transfer of electronic records has been determined as authorized. That determination generates a Request for Information about Authenticity of the Records. In
response to this type of request, the activity outputs Retrieved Information about the Presumption of Authenticity of the records as maintained by their creator. This information will be taken into account in the decision whether to accession the records.

A3.1.3, Retrieve Information about Digital Components
In accordance with the applicable Preservation Strategy, upon receipt of Information Identifying the Digital Components of a Requested Record, retrieve Maintained Information about Digital Components and output the requested Information about Digital Components along with a Request for Digital Components to be retrieved from storage. Follow the same procedure when triggered by a Plan for Updating Digital Components.

A3.2, Manage Storage of Digital Components of Records
In accordance with the Preservation Strategy established for a body of records, and applying the Storage Method selected to implement that strategy, place the digital components of Accessioned Electronic Records into storage, taking the specific steps defined in the Preservation Action Plan for these records, and maintain them. In response to a Request for Digital Components, retrieve the requested components and output them. When digital components are output for updating in accordance with a Preservation Action Plan, place the Updated Digital Components in storage and, as provided by the Action Plan, either maintain or delete the older versions of these components. Provide to the “Manage Information” process Updated Storage Information about the identities, locations and other relevant parameters of stored digital components whenever components are updated or other changes, such as media refreshment, are made in storage.

A3.2.1, Place Record Components in Storage
When electronic records are accessioned, place the Digital Components of the Accessioned Electronic Records into one or more Stored Digital Files in the storage system prescribed by the preserver’s Storage Method, and provide Updated Storage Information about those components. When any of the stored components are updated in response to a Preservation Action Plan, place the Updated Digital Components into storage and either replace or retain prior versions of those components, as dictated by the applicable Preservation Strategy (tunneled to this diagram). Provide Updated Storage Information about the updated components and about any stored components that have been deleted or superseded. When a Stored Digital File is copied to new storage media, place the Refreshed File into storage, delete the older copy of the file, and provide Updated Storage Information about the current location of the digital components stored in that file. When a problem discovered in a Stored Digital File is corrected, place the resultant Recovered File in Storage, delete the problem file, and provide Updated Storage Information about the current location of the digital components stored in the file and about any data loss or other residual problems with any of those components.

A3.2.2, Refresh Storage
When triggered by a Plan for Updating Storage, use the prescribed Storage Update Method to copy a Stored Digital File to new storage media. When indicated by the plan, replace one or more components of the storage subsystem, ensuring that any Stored Digital File involved in, or affected by, such replacement is carried forward without any inappropriate alteration. Document the process as part of the preservation history of all records whose digital components are contained in these digital files. Send any Refreshed File to be placed in storage and provide Updated Storage Information about any digital components affected by the process.

A3.2.3, Monitor Storage
In accordance with the Preservation Strategy (tunneled to this diagram) for the storage system, apply the prescribed Monitoring Method to monitor the operation of the storage system, the media on which Stored Digital Files are recorded, the files themselves, and
the facilities where the system and files are located. Provide Updated Storage Information about the problems identified and the stored digital components they affect.

**A3.2.4, Correct Storage Problems**
Upon notification of a Storage Problem, apply the Problem Correction Method prescribed by the Preservation Strategy (tunnelled to this diagram) to take the actions indicated by the Plan for Problem Handling to eliminate the problem. If the Storage Problem affects any Stored Digital File, take action to copy the data stored in that file to a Recovered File and generate Updated Storage Information on each digital component affected by the process, including the identity of the Recovered File where the component is written, the success of the copy process, and any data loss or residual, uncorrected problem.

**A3.2.5, Retrieve Components from Storage**
In response to a Request for Digital Components, apply the Retrieval Method specified in the Preservation Strategy (tunnelled to this diagram) to retrieve the Stored Digital File(s) in which the requested components are written from storage and output copies of the Retrieved Digital Components. If the Preservation Strategy applicable to a given record or archival aggregate is modified, and the modification requires updating of the digital components, the new strategy triggers retrieval of the components that need to be updated.

**A3.3, Update Digital Components**
As indicated by the Preservation Strategy established for a given body of electronic records (tunnelled to this diagram), take the steps indicated in the applicable Preservation Action Plan, applying the Method(s) for Updating Components prescribed by the strategy to update Digital Components of a Record that Cannot Be Preserved because of technological obsolescence, changes in Preservation Strategy, or similar factors. Examples of processes for update include migration, standardization, and transformation to persistent form. Return the Updated Digital Components to Storage, providing Information about the Updated Digital Components to the “Manage Information” process. If the Updated Digital Components belong to a record that is the subject of a Retrieval Request, also send the components, along with related information, to the Output Electronic Record process. However, if the updating was done only to satisfy conditions of a Retrieval Request and was not required to conform to Preservation Strategy, the Updated Digital Components are sent, along with related information, to the Output Electronic Record process, but they are not sent to Storage. This process may be invoked directly when records in a transfer are being examined and it is determined that there is a need to take action to preserve a record, before the components are sent to storage.

**A4, Output Electronic Record**
Following direction established in the preservation strategy for a given body of records selected for preservation, apply preservation method(s) targeted to that body of records to implement the Preservation Action Plan for producing an authentic copy of a records in response to a request for it. If specified in the request for a copy of the record, produce a certificate attesting to the authenticity of the copy. Alternatively, if requested, produce a reproducible electronic record; i.e., the digital component(s) of the record along with instructions for producing an authentic copy of the record and information necessary to interpret the record. In the case of a request only for Information about a record, deliver the response. To produce all of these outputs, translates an external request for a record or for information about a record into a Retrieval Request to the “Maintain” function. Also produce management information which is used to evaluate execution of the “Output” function. The process is carried out by persons responsible for preservation, using infrastructure technology. The Output Electronic Record process may also be invoked when records in a transfer are being examined in order to verify that the records in the Transfer can be preserved and reproduced.
A4.1, Manage the Request
Following provisions in the applicable Preservation Strategy, register an incoming Request for a Record and/or Information about a Record. Translate the request into terms that can be executed in the preservation system, and send the Retrieval Request to the Maintain Records process. Define Request Controls to ensure that the request is fulfilled and accounted for. These controls govern the process of responding to a request. If any problem is encountered in fulfilling the request, a Report of Problem with Retrieval Response is sent as feedback to this process. If the request cannot be satisfied, produce an Accounting for Unsatisfied Request and send it to the requester.

A4.2, Review Retrieved Components and Information
Under the control of the applicable Preservation Strategy and Request Control, receive Retrieved Digital Components and/or Retrieved Information about a Preserved Record and determine whether all components and information necessary to satisfy the request for records have been received and can be processed for output. If the request entails producing a copy of a record, send the Requested Digital Components forward to Reconstitute and Present the Record. If a request for a record does not require reproduction of the record within the preservation system, send the Requested Digital Components forward to be packaged with related Information and delivered to the requester. If the request is only for information, output the Requested Information about a Preserved Record. If the request cannot be satisfied in accordance with the Request Control, produce a Report of Problem with Retrieval Response.

A4.3, Reconstitute Record
Under the control of the applicable Preservation Strategy and Request Control, apply the appropriate Targeted Preservation Method to Retrieved Digital Components to link or assemble the components as necessary to reproduce the record and output the Requested Reconstituted Record. If the record cannot be reconstituted, produce a Report of Problem with Retrieval Response.

A4.4, Present Record
Under the control of the applicable Preservation Strategy and Request Control, apply the appropriate Targeted Preservation Method to Retrieved Digital Components to the Requested Reconstituted Record to present the record with the appropriate extrinsic form. If requested, produce a Certificate of Authenticity for the Reproduced Electronic Record. If the process of reproducing the record is unsuccessful, produce a Report of Problem with Retrieval Response.

A4.5, Package Output
Under the control of the applicable Preservation Strategy and Request Control, apply the appropriate Targeted Preservation Method to Retrieved Digital Components to combine Requested Digital Components with Information, including instructions on how to reproduce the record, into a package suitable for reproducing the record on an external system designated by the Requester. If the Request for a Record entailed reconstituting the record within the preservation system, package the Requested Reconstituted Record suitably for presenting the record on an external system designated by the requester. If the process is unsuccessful, produce a Report of Problem with Retrieval Response.
iv) Arrow Definitions

Accessioned Electronic Records
A body of electronic records selected for preservation, transferred to the preserver and accepted by the preserver for preservation.

Accessioning Policy
The policy of the institution or person responsible for preservation with respect to accepting responsibility for records transferred for preservation. Includes standards and specifications for acceptable and unacceptable deviations from standards, such as when records that should be in a transfer are missing or when information that should accompany the transfer is missing, inappropriate, or unclear.

Accounting for Unsatisfied Request
An explanation of why a Request for a Record and/or Information about a Record could not be satisfied in whole or in part.

Archival Aggregate Requirements
Stipulations as to how the original order of records is to be respected in the physical or logical structuring of sets or archival aggregates of records, and how they are to be presented for use.

Archival Requirements
Requirements derived from archival science, diplomatics, best practices, and prevalent standards within the archival community.

Basis of Authenticity of Records
Information that indicates whether records can be considered as authentic on the basis of how the records creator addressed the Benchmark Requirements for Authenticity up through the time when the records were transferred to the preserver or, alternatively, whether their authenticity needs to be verified through corroborating evidence.

Certificate of Authenticity
An attestation by the person responsible for preservation that one or more records are authentic.

Classes of Records
A class of records is a set of records with common attributes and methods determined on the basis of their documentary form.

Conforming Digital Components
A conforming digital component is a digital component that can be processed using current preservation methods in order to preserve and reproduce an electronic record.

Conforming Transfer
A transfer of electronic records where information accompanying the transfer indicates that it satisfies the terms and conditions stipulated for the transfer.

Determination that Records Cannot Be Preserved
A judgement that, for technical and/or resource reasons, a body of records deemed to have archival value cannot be preserved in electronic form, either with any preservation methods currently in use or with any available technological options.

Digital Components of a Record that Cannot Be Preserved
A Digital Component of a Record that Cannot Be Preserved is a digital component of an electronic record that cannot be reconstituted or presented, or whose archival bonds
cannot be expressed, or whose arrangement in archival aggregates cannot be re-established using the Preservation Methods specified in the applicable Preservation Strategy.

**Digital Components of Accessioned Electronic Records**
The digital components of the electronic records included in a transfer and accepted by the preserver for preservation.

**Digital Components that Need Updating**
A Digital Component of a Record that cannot be reconstituted or presented in accordance with current Preservation Strategy applicable to that record, or a Digital Component of a Record belonging to an archival aggregate that cannot be re-established, reconstituted, or presented using the Preservation Methods specified in the applicable Preservation Strategy.

**Evaluation of Execution**
The result of evaluating management information about the execution of preservation process to determine whether requirements were satisfied, preservation strategies are effective, and action plans meet objectives and performance targets.

**Facilities**
Locations where digital preservation technologies are installed and operate, and locations where electronic records components are stored.

**Information about Accessioned Records**
Information identifying records that have been accessioned, their digital components, and the preservation strategies that will be applied to them.

**Information about Digital Components**
Metadata or other information retrieved in response to a request for a record or for information about a record, or in response to a plan for updating the digital components and used to satisfy the request or to carry out the plan.

**Information about Digital Components of an Electronic Record**
Technical information concerning a digital component or a class of digital components of electronic records necessary to store and retrieve the digital components of an electronic record and to reconstitute the record from the components and to present it in authentic form, properly ordered with respect to related records.

**Information about Electronic Records Selected for Preservation**
Information output from the appraisal process identifying and characterizing records that are to be preserved, including what information about the records should accompany the transfer, the basis for asserting the authenticity of the records as maintained by their creator, and the terms and conditions of transfer.

**Information about Preservation**
Reports on, data about, or accounting for the exercise of the preservation function, the preserved records, and the authenticity of copies of those records. The coverage of such information may range from the preservation function as a whole, to one or more processes within the function.

**Information about Presumption of Authenticity of Appraised Records**
Information indicating the extent to which the creator of the records that have been appraised satisfied the criteria set out in the Benchmark Requirements for Authenticity.

**Information about Presumption of Authenticity of Transferred Records**
Information indicating the extent to which the records creator satisfied the criteria set out in the Benchmark Requirements for Authenticity through the point where the records were transferred to the preserver.
Information about Transferred and Accessioned Records
Information about electronic records, their arrangement, and their digital components, which accompanies the transfer of electronic records and which is developed in the process of bringing them under the preserver's control.

Information about Updated Digital Components
Information about changes that have been made to a digital component of an electronic record in the process of updating it, about any problems that occurred in the process, and about the identity and location of the component on storage media and in a storage system or facility.

Information and Communications Technology
Digital hardware, software, and storage and communications media.

Information and Communications Technology Infrastructure
Hardware and general-purpose software, such as operating systems, file systems, communications utilities, device drivers, and database management systems.

Information Identifying Digital Components of a Requested Record
Information that specifies all the digital components necessary to reproduce a record, and their unique identifiers.

Institutional Requirements
External legal, regulatory, societal, and cultural constraints imposed on the institution responsible for preserving records, together with the preserver's internal standards, policies, procedures, goals, objectives, and criteria applicable to records in general or of electronic records specifically. An external legal, regulatory, societal, and/or cultural constraint imposed on the institution responsible for preserving records, and/or an internal policy, procedure, goal, or objective, applicable to the preservation of records in general or of electronic records specifically.

Maintained Information about Digital Components
Information identifying a stored digital component, and the record(s) that comprise that component, and information enabling the reproduction of the record(s).

Maintained Information about Records
Information about records, archival aggregates, and the authenticity of records maintained by the preserver to support preservation and reproduction of authentic copies of the records.

Management Information about Preservation
Information about an electronic record that is transferred for preservation and about the processes of bringing in, maintaining, and reproducing the record. This information enables management to evaluate those activities and the preservation strategies and plans they implement, and to determine what changes may be needed in requirements or strategies.

Mapped Records and Digital Components
A mapping of information about electronic records reportedly transferred for preservation to the digital file(s) that were transferred, identifying and locating the digital components included in the file(s), linking them to the records that they constitute, and determining whether the information about the records and the digital components in the transfer is appropriate and sufficient for preservation and reproduction of the records.

Method for Updating Components
A Targeted Preservation Method used to migrate, transform, or otherwise modify digital components of electronic records in accordance with the preservation strategy applicable to those records.
Monitoring Method
A preservation method for determining whether a storage system is properly maintained and functioning or whether storage media are intact and free from problems that would interfere with reading the data written on the media.

Non-Conforming Digital Components
A non-conforming digital component is a digital component on an electronic record in a format that prevents or impedes the reproduction of the record in accordance with the applicable preservation strategy.

Notification of Receipt
A record sent to the submitter acknowledging that the preserver has received the transfer and, if needed, asking the submitter to address any problems identified in registering the transfer.

Packaging Method
A Targeted Preservation Method used to wrap or encapsulate either the Digital Components of an electronic record or a Reconstituted Electronic Record in a form that will enable reproduction of the record on a designated target system, and to combine the resultant digital object with instructions for reproducing the record on the target system.

Persons Responsible for Preservation
Persons authorized and charged with carrying out the preservation function or processes within that function.

Presentation Method
A Targeted Preservation Method used to present an electronic record in appropriate documentary form.

Preservable Records
An electronic record comprised of digital components that all conform to the applicable preservation strategy.

Preservation Strategy
A coherent and comprehensive approach for preserving a body of records selected for preservation, derived from archival and institutional requirements, taking into account Evaluation of Execution of current and prior Preservation Strategies and reflecting the State of the Art of Information Technology. A Preservation Strategy includes objectives for maintaining components of electronic records and related metadata and information over time, and for reproducing the records in authentic form, in the order imposed by the records creator, and criteria for evaluating execution of the Preservation Strategy. The strategy includes specifications for handling exceptions to its standards, and identifies the targeted preservation methods to be used. Each Preservation Strategy has a specified scope which may range from all records transferred for preservation through a specific body of records selected for preservation to an individual transfer. The Preservation Strategy also includes one or more preservation action plans. A preservation action plan specifies preservation actions to be taken for the classes of digital objects covered by the Preservation Strategy, and indicates the time or conditions when such actions should be taken. Preservation actions should be planned for the transfer of records to the archives, in accessioning, during maintenance, and even in output.

Preservation Technology Specifications
Information about a Targeted Preservation Method.

Problem Correction Method
The method stipulated in a preservation strategy for correcting problems of a specified type.

Record of Accession
A record documenting the preserver's acceptance of responsibility for preserving a set of electronic records.
Record Preservation Requirements
Specifications for reconstituting a record from its digital components and for presenting the reconstituted record for use.

Record Reconstitution Method
A Targeted Preservation Method used to assemble or link the Digital Components of an electronic record to enable its presentation in appropriate documentary form.

Recovered File
A physical or logical file that has been successfully read, or reconstituted, and placed back into appropriate storage after one or more storage problems affecting the file were found.

Refreshed File
A physical or logical file that has been copied from an older storage medium or system to a newer when the older medium or system has been replaced or updated.

Registered Transfer
A transfer is determined as authorized if and only if it comprises a record that has been selected for preservation and the record has been submitted either by the record's creator or an agent for the creator.

Registration Procedure
The procedures and recordation required by the preserver to identify and track a transfer of electronic records.

Rejected Accession
A set of electronic records selected for preservation and transferred to the person responsible for preservation, but not accepted for preservation by the preserver.

Rejected Transfer
A transfer of electronic records which does not satisfy requirements for being accessioned or preserved.

Report of Problem with Retrieval Response
Information describing why a request for a record or for information about the record cannot be satisfied in whole or in part.

Report on Authenticity of Records
An account of preservation activities related to a given record or aggregate of records to support the assertion or certification that reproductions of the record(s) are authentic.

Reproduced Electronic Record
An authentic representation or other version of a record reconstituted from its digital components, along with information supporting the interpretation of the record.

Reproducible Electronic Record
Digital components of an electronic record, the technical information necessary to reproduce the record from the digital components, information about the authenticity of the record, criteria for certifying the authenticity of the reproduced record, and information that supports interpretation of the record.

Request Control
A specification of the deliverable(s) to be produced in response to a request for a record or for information about a record, and the criteria to determine whether the request has been satisfied.

Request for Digital Components
An instruction to retrieve the digital components of a record.

Request for Information about Authenticity
A request for information already maintained by the preserver concerning the authenticity of a record or a body of records for use in examining records being brought into the
preservation system. This information includes data produced in the appraisal of the records and data developed in examining any prior transfers of records from the same body of records.

Request for Record and/or Information about Record
A request to output a preserved record, or to provide information about a record that cannot be found in or derived from archival description or finding aids, but requires either retrieval and processing of the preserved record or of information that is created or specifically maintained in the preservation system.

Request for Strategy Decision
A request for a management decision formulated when it is determined that an electronic record cannot be preserved in accordance with the applicable preservation strategy because one or more digital components of the electronic record do not conform with the specifications of that strategy. The request asks for a determination of whether the preservation strategy should be changed or, alternatively, the component(s) should simply be determined to be non-conforming digital components.

Requested Digital Components
All of the digital components necessary to satisfy a request for a record.

Requested Information about a Preserved Record
The information provided in response to a request for information about a preserved record.

Requested Reconstituted Record
The digital components of a requested electronic record linked or reassembled to enable reproduction of the record.

Requirements for Physical and Logical Files
Requirements for physical files include standards and specifications of what media are to be used for what purpose, how individual volumes are to be labelled, and how physical files are to be inscribed on the media. Requirements for digital files include, at a minimum, how both physical and logical files are to be identified, how logical files are mapped to physical files, and how integrity of a file is ensured.

Retrieval Method
The preservation method to be used to retrieve the digital components of an electronic record from storage.

Retrieval Request
A request for retrieval of a record and/or information about the record.

Retrieved Digital Components
The digital components of an electronic record retrieved from storage in response to a request.

Retrieved Information about a Preserved Record
Information retrieved from storage in response to a retrieval request.

Retrieved Information about Presumption of Authenticity
Information about the presumption of authenticity of a body of records transferred to the archives used in deciding whether to accession the records.

State of the Art of Information Technology
The state of the art of the technology with respect to its ability to satisfy archival preservation requirements, the state of the underlying computer science with respect to its ability to develop relevant capabilities not within the state of the technology, and the existence and prevalence of applicable standards.
Storage Method
A Targeted Preservation Method used to place and maintain digital components in storage, to retrieve them from storage, or to remove them from storage.

Storage Problem
A problem with storage media, storage formats, a storage system or facility that could impact on the continued preservation of records.

Storage Update Method
A method used to ensure that stored digital components are completely and correctly brought forward when any component of a storage subsystem is changed or when digital files are moved or migrated to newer storage media.

Stored Digital File
A digital file placed in a storage system on digital media.

Synthesized Requirements for Preservation
A coherent set of requirements for preserving electronic records. Each set of requirements applies to a specified domain of digital objects or records. The digital object domains include the storage media to be used for transfer and maintenance of digital files, and the digital files themselves. Requirements for media include standards and specifications of what media are to be used for what purpose, how individual volumes are to be labelled, and how physical files are to be inscribed on the media. Requirements for digital files include, at a minimum, how both physical and logical files are to be identified, how logical files are mapped to physical files, and how the integrity of a file is ensured. A record domains is either a body of records selected for preservation, one or more archival aggregates within that body, or the records included in a transfer from the selected body of records.

Targeted Preservation Method
Software used to implement a preservation strategy or strategies. A preservation method is targeted by specifying the bodies of records, types of electronic records, and/or classes of digital components to which it will apply, along with conditions for its application to each target group.

Technological Infrastructure
Hardware and common services software—such as operating systems, communications software, database management systems, electronic mail applications—necessary for the implementation of preservation methods and the execution of the preservation processes defined in the “Preserve Electronic Records” model.

Terms and Conditions of Transfer
The specifications governing the transfer of a body of electronic records selected for preservation to the preserver. These specifications include, at a minimum, when the records should be transferred, whether the entire body of records should be transferred at one time or incrementally, the types of records or selection criteria for including records in a transfer, what information should accompany the transfer, label and format standards for transfer files, and what physical medium or telecommunications channel should be used to effect the transfer.

Transfer of Electronic Records Selected for Preservation
The digital components, the accompanying information related to preservation and reproduction of the related record, and transmittal information.

Types of Record Aggregates
A type of archival aggregate expresses the general logical or physical method for ordering of members of an archival aggregate of records, from aggregates of individual records through series to the entire archival fonds.
**Updated Digital Components**
An updated digital component is a component that has been modified under a preservation action plan.

**Updated Storage Information**
Information indicating a change in the location of a digital component in storage, the occurrence of a storage problem, the action taken to correct a storage problem, the results of such action, or the copying of digital files from older to new storage media.
Appendix 6

How to Preserve Authentic Electronic Records

Preservation Task Force

8 October 2001
Introduction

This report communicates the results of the work of the InterPARES Preservation Task Force. The task force was chartered to identify and develop the procedures and resources required for the implementation of the conceptual requirements articulated in the InterPARES research on authenticity and appraisal of electronic records.\(^1\) To achieve this goal the Preservation Task Force formulated and analyzed the problem of preserving authentic electronic records in order to articulate a detailed, in-depth understanding of this problem. The task force followed two principal paths in articulating this understanding: (1) an international survey of current practices and plans in the preservation of electronic records and (2) a formal modelling of the function of preserving electronic records. InterPARES' United States Research Team independently contributed to the task force's work by developing a bibliography on digital preservation\(^2\) and a report on digital storage media.\(^3\)

The analysis has produced a functional model of the processes necessary to preserve electronic records selected for preservation, a model of the information needed to support the preservation function, a glossary defining the terms in the models, and a report on a survey of current practices and plans in the preservation of electronic records.

Electronic Records

Any approach to the preservation of electronic records has to start with clear recognition of the basic characteristics of such records, to what extent these basics are the same as those of traditional records, and how they differ.

Recording Information

Recording information enables it to be transmitted across time and space and between or among persons or organizations. Any recording of information requires some way to represent that information on a physical medium. Textual information, for example, is represented through the use of alphabets or character codes, punctuation marks, abbreviations, fonts or handwriting styles, page layout, etc. When this information is recorded on paper, ordinarily there is absolutely no difference between the way the message is represented on the medium and the way it is presented to humans for interpretation and use. But when the information is recorded digitally, there is an inevitable difference between the way it is represented on a medium and the way it is presented for use. In part, this difference derives from the fact that different types of media are used for storage and presentation; for example, digital information is typically stored on magnetic or optical media, but displayed on cathode ray tube (CRT) or liquid crystal display (LCD) screens. More important, however, is the basic difference between the digital encoding of information in binary values for storage, transmission, and processing by computers and the translation of that encoding into a form that can be used by humans.

Whether a textual document is stored digitally as a scanned image of a paper document or in a character-based representation, such as ASCII or Unicode, it is necessary to transform that representation into a very different one to make the document readable. For example, in character mode, every single character must be presented in the chosen font, but in storage, the font may be indicated only by a special code that precedes an entire block of text to which it applies.

It is not possible to store an electronic record in the form of a record. The “form of a record” is the

---

\(^1\) <http://www.interpares.org/researchplan.htm>

\(^2\) <http://is.gseis.ucla.edu/us-interpares/bib_pres.htm>

\(^3\) P. C. Hariharan. Media, A Presentation for the InterPARES Panel at UCLA. 2 December 1999.
documentary form that enables it to achieve its intended purpose in the first place and to communicate the same information over time. Electronic records are stored in forms that differ substantially from those in which they can serve their intended purpose as records.

This report was created using word-processing software on a PC. Essential to its form as a record is the visual presentation of natural language text in lines, paragraphs, and sections. But this report cannot be stored, in the computer’s memory, on its hard drive, or on any digital medium, in this form. It can only be stored as one or more sequences or strings of bits. Each textual character (letter, space, and punctuation mark) in this report is stored as a sequence of eight bits. Special sequences of bits are used to indicate breaks in the flow of text, such as the separation of paragraphs and section headings, and the indentation of this paragraph, and other presentation features, such as different type sizes, bold, and italics.

The difference between the way digital information is represented in storage and the way it is presented for use occurs even when the use is within or between computer systems and does not involve humans.

When an individual uses a chequing account debit card to make a purchase, the store’s computer system reads the magnetically encoded data on the card to identify the customer’s bank and account. The store’s computer then sends a message to the bank’s computer asking for transfer of the amount of the sale. The bank’s computer then checks its chequing account database to see if the requested funds are available and, if so, records the transaction, deducting the sale total from the account balance. Then it notifies the store’s computer, which creates its own record of the transaction. In this single transaction, there are several transformations of the way the information involved is represented digitally. Transformations occur, for example, when the computer reads the magnetic strip on the debit card; when the store sends the message to the bank over a communications channel; when the bank’s computer receives the message and stores it in its memory; when the computer invokes the chequing account database; when the database management system invokes its optimizer; when the optimizer interacts with the storage subsystem to request and retrieve the account data, etc. There are additional transformations when the sales information is displayed on the cash register and printed as receipts for the store and customer.

Differences between the storage and use versions of electronic records are not limited to the way the data are inscribed on physical media. There may also be substantial differences between the units in which digital information is stored, that is digital files, and the units in which that information is organized for use. There is no necessary relationship between digital files and the archival units of record, file of records, record series, etc. A single record could be stored in one or more digital files. A single digital file could contain one or more records, and a single digital file, such as one containing the specifications for the layout of a form or report, could be used in thousands of records. In business applications that rely on database systems, the database often consists of several thousand logical files, each of which contains part of the contents of potentially millions of records.

Often the digital file in which an electronic record is “stored” does not in fact contain the entire record. Ordinarily, if a document was created using, say, word-processing software and stored as a single digital file in the native format used by the software, we consider that this file is or contains the record. However, the word-processing file is insufficient to reproduce the record as it was meant to appear. Data that are essential to correct rendering of the presentation features of an electronic record are usually stored outside of the file, or files, that contain the content of the record.

In order to present this report, or any word-processing file, on a video display device, it is necessary to pull in data from one or more other files. While the word-processing file contains bytes representing all of the text characters in the document, the word-processing application needs to use data that the Windows operating system stores in separate files called dynamic link libraries, or dll files, in order to display these characters in the font or fonts used by the writer. Such font files are used with all user-created files where the pertinent fonts are selected. From an information technology perspective, the font files are extensions of the Windows operating system, rather than parts of the user-created files. Nonetheless, from an archival perspective, a font file must be treated as a digital component of any record that relies on the font for proper presentation. While some font files may contain bitmapped images of each character, most often the dll file contains data about the characters that the software uses dynamically to synthesize or form the characters for display. Similar processes occur for other types of content, such as graphics, and in other types of applications that contain such data types.
In sum, the relationship between archival units and digital files may be one-to-one, one-to-many, many-to-one, or many-to-many.

Differences between storage representations and use presentations do not only occur among different records and digital files. The same presentation of a record can be produced from a variety of storage representations. Conversely, a single storage representation may be processed to output a variety of presentations.

One technique that is used to ensure the fixity of both the content and the appearance of digital documents is to convert them from word-processing files to formats, such as portable document format (PDF) files, which include the data necessary to form font characters properly, eliminating dependence on dynamic link libraries.

While the data content of a record may be stored in many different tables in a relational database, the data must be brought together in order to present the record in its proper documentary form. Once the record is reconstituted from its different data components, it may be possible, and in fact simplest, to save it as a single digital file, for example as a word-processing file, without in any way impacting its identity or integrity.

Because there is no necessary mapping between digital files and records, in many cases it is possible to change the way the record is stored in digital files without changing the record itself.

The inevitable, and often repeated, changes between the storage representation and the presentation for use of digital information create paradoxical elements in the preservation of electronic records.

**Keeping Electronic Records**

When a recording of information is intended to serve as a record of an action or state of affairs, it is essential that the message it transmits be fixed. An authentic record is one that is what it purports to be and that is free from tampering or corruption. Determining that it is what it purports to be means confirming its identity. Determining that it is free from tampering or corruption means demonstrating that its integrity remains intact through space and time. In the case of records on paper and other "hard" media, the authenticity of a record over time rests on the assumption that the physical object that embodies the record has not changed in any way that would affect the message it was intended to communicate. Thus, in traditional archival practice, preference was given to the original record and to an unbroken chain of custody. The original is preferable because any process of copying it introduces an opportunity for alteration. The principle of the unbroken chain of custody stipulates that, throughout their life cycles, records should be in the custody of known parties who can be trusted to preserve them intact. Continuous custody is important because any break in control over the record also creates risk of deletion, alteration or substitution; furthermore, any discontinuity in custody may make it impossible to demonstrate that a record has not been altered.

Probably the most basic aspect of preserving records on hard media is placing and keeping them in storage. Ideally, the environment in which the records are stored should not include any elements that would damage the records or cause them to deteriorate; furthermore, if needed, the environment should be designed to reduce or retard any deterioration that is intrinsic to the physical media on which the records are stored or to the physical means used to inscribe the records on the media, for example, by controlling temperature and humidity. Active conservation measures are taken to prevent or recover from any damage or deterioration. The goal is to justify belief that a record retrieved from storage is the same in all essential respects as the record previously placed in storage. Traditionally, the preservation of records has focused on ensuring their fixity through the processes of conservation and maintenance, where maintenance refers to keeping records in places and under conditions that protect them from harm and minimize or reduce any innate tendencies towards deterioration, and conservation refers to interventions to repair damage and to prevent deterioration which has a high risk of occurring.

Keeping electronic records is more complex and difficult. Like all records, an electronic record must transmit the message intended by its creator; however, the fixity of the message carried by an electronic record is at risk because of the changes between the way it is represented in storage and the way it is...
presented for use. Both placing an electronic record in storage and retrieving it for use entail transforming the way the content, structure, and appearance of the record are inscribed on a physical medium. Both storage and retrieval transformations create risks that the record may be altered. These risks are compounded by the software, hardware, and the physical media used, including both the storage media and the media on which the records are presented in record form. While maintenance and conservation of the stored information remain essential, the integrity of an electronic record depends on guaranteeing that none of the changes between storage representation and presentation for use, in either direction, has altered the message the record was intended to convey from the time it was first filed as a record to the point of any subsequent use. The necessity of ensuring that transformations between storage and use do not corrupt the records adds a new focus to the preservation of electronic records. An electronic record cannot be said to have been preserved unless it can be delivered in authentic form.

The process of preserving electronic records extends over the entire life cycle of the records from creation to disposition and, in the case of records that are preserved for posterity, to the reproduction of those records. The overall process of preservation must be continuous. If there is ever a point where we cannot reasonably assert that the record continues to carry its original message intact, we can never thereafter assert that it is authentic. It is important to recognize that while the process must be continuous over time, the activities that constitute the process are discrete steps. Each instance where the way the information is represented changes—whether moving between storage and use or between storage media or subsystems—is a potential point of failure, a weak link where the entire chain could be broken. The process of preserving electronic records extends to and includes interactions between computer systems and human users and interoperations between computer systems, subsystems and applications.

Foundation Concepts

This review of differences between traditional and electronic records leads to the articulation of several basic concepts about the preservation of electronic records which distinguish it from preservation of traditional records.

Digital Components of Electronic Records

The most basic concept demanding attention in preserving electronic records is that, in addition to all the intrinsic and extrinsic elements of form that make up any record, an electronic record also comprises one or more digital components. A digital component of an electronic record is a digital object that is, or is part of, an electronic record, or that contains one or more parts of one or more electronic records, and that has specific methods for storage and reproduction. The complexity of this definition derives from the unlimited cardinality in the relationships between electronic records and their digital components. While complex to define, the concept of digital component is relatively easy to describe and to grasp intuitively.

Every electronic record has at least one digital component: a stream of bits representing information contained in the record. Each digital component has one or more associated methods, or programs, for decoding the bitstream and/or presenting it for use. In the simple case of a record with a single component, the record and the digital component are congruent: the component is the record. However, the contents of a record may be stored in several bitstreams; for example, a single record may be stored as a compound document, with different parts of the record stored in different digital files. In such cases, the record has as many digital components as it has bitstreams. When a record has more than one component, each digital component is part of the record. Digital components may be distinguished from one another based on the fact that they are stored separately. If a physical file contains all of the components of an electronic record, it contains the record. If a physical file contains some, but not all of the components, it contains parts of the record.

Physical storage as a criterion for identifying digital components is a containment relationship: a physical file contains all or part of one or more records. In fact, the containment relationship may be either physical or logical, or both. An electronic record may contain one or more distinct digital components stored entirely within a file (physical), or a file may only contain links pointing to distinct components stored elsewhere (logical). For example, in a relational database, a record—in the archival sense—is likely to
comprise data "stored" in several different database tables. Each table that contains part of the record is a
digital component of that record. Relational database management systems typically store all of the
tables that constitute a database in a single physical file. Nonetheless, reconstructing the record with all
of its data content requires locating the row or rows in each table that contains such data and retrieving
the record content from it. Thus the logical tables are distinct digital components.

Digital components may be identified based on other criteria besides separate storage or containment
relationships. A record may consist of many digital components, which may be stored together in one
physical file. A textual record may contain non-textual content, such as a spreadsheet, an image or even
a voice annotation. In such cases, it does not matter whether the different types of content are stored
together or separately. Whether the units of non-textual content constitute separate digital components
depends on the format or data-type used to represent the contents. In a word-processing file, a picture
would constitute a distinct digital component because the word-processing content is character-type data
while the picture is binary or raster data. However, if the textual record was a scanned document image,
both the picture and the text would be binary image data and constitute only a single digital component.
Thus, data-type appears as a second criterion for distinguishing digital components. However, this
criterion needs to be refined. Specifically, following the definition of digital component given above, a
digital component has one or more methods associated with it. A word-processing document and a
spreadsheet both use character data, but each requires different software for proper processing and
presentation. Conversely, alphanumeric characters and linear graphics are two different types of data, but
most word-processing applications are capable of vertical and horizontal lines; therefore, a word-
processing document containing only text and such graphic lines has only one digital component
determined by data-type. Data-type is a criterion for distinguishing digital components of records if and
only if different data-types have different methods associated with them.

An electronic record may also have one or more distinct digital components necessary to render the
record correctly. A bitstream may contain data that indicate how the information content is to be
presented for use; for example, codes indicating fonts, type sizes, line endings, paragraph indentations,
etc. However, codes indicating presentation features may also be stored in separate bitstreams, such as
dynamic link libraries, report templates, or style sheets. Each bitstream indicating presentation features is
an additional digital component of the record. In such cases, the components containing presentation
specifications may be processed by the same software used to process the bits representing the content
of the record. Even though the same link library or other such file may be used in many different
documents, each such file required to present a record correctly should be considered as a digital
component of the record. These components are distinguished on the basis of separate storage.

Both containment and data-type criteria may apply to the same record. A word-processing file containing
two digital photographs has only two data-types, but three digital components.

In sum, each digital component is a logical or physical object that the system processes as an unit. In
other words, each component has a specified method or methods. In fact, when several components are
stored together in a physical file, there must be specified methods for locating and extracting the
components. Thus the physical file is itself a component, in addition to the components it contains.

As this discussion shows, there is no necessary relationship between the elements of form of an
electronic record and its digital components. In fact, in some cases the relationships between the record
and its digital components can be changed without significant impact on the record as such. For example,
a textual record originally created as a word-processing file could be changed to a binary image or
portable document format without impacting any of its essential characteristics as a record. However,
changes in the digital components of a record could corrupt the record. Therefore any such changes need
to be under preservation control.

**Preservation Control**

A technological boundary exists between any two states of a system or of interoperating systems when
the transition from one state to another does, or can, entail significant changes in the attributes or
methods of a digital object. For records, significant changes are those that affect identity or integrity.
Technological boundaries exist at macro and micro levels. Macro level boundaries occur at the interfaces
between systems, subsystems or applications, such as during system, media, or data format migrations
or in transfers between the "live" systems in which the records are created, and other applications in
which they are transmitted over space or stored over time. Micro level boundaries occur when a record is
decomposed into separate digital components or is reconstituted from its components, and when different
methods are invoked to process distinct components. Transitions from storage representation to
presentation for use can involve both macro and micro boundaries.

Preservation control is critical in transitions across technological boundaries. Preservation control consists
of actions, conditions, and constraints designed to ensure the preservation of records and their continued
authenticity. While preservation controls during maintenance of the records in storage must be adequate
and effective, the risks of corruption or loss of records are more frequent and complex during transitions
across technological boundaries. Thus preservation controls can be divided into two types: systemic
controls are those that ensure records remain unchanged over time within a given system or subsystem;
dynamic controls are those that ensure records remain authentic across technological boundaries.

Preservation control, in most instances, will be accomplished through technical means, but it must be
determined according to archival principles and criteria.

Archival Requirements for Preservation

Naively, preservation may be seen as a process that keeps records free from change. However, it can be
easily shown that it is practically impossible for any record to remain absolutely unaltered or immutable
over time. More important, many changes that occur naturally or accidentally do not impact the
authenticity of the record. Paper darkens. Ink fades. Microfilm scratches. Although such changes may
indicate threats to the continued existence of the records, they do not necessarily make the records
inauthentic. One of the best-known records in the world is the original of the Constitution of the United
States. The paper has yellowed and the ink has faded considerably over the years. Facsimiles of this
document have been produced that arguably look more like the document originally did than the original
currently does; however, no facsimile or other copy can ever approach the value persistently attributed to
the original. Hundreds of thousands of people regularly visit the National Archives Building in Washington
to view the original record. In fact, the physical changes in the original evince its authenticity because
they result from the operation of the laws of nature: the original record should not look as it did two
hundred years ago.

The requirement for records to remain unchanged, thus, needs to be qualified. Even in the case of hard-
copy records, this requirement effectively means that the record should not be changed in any way that
relates to its essential record nature, rather than to its existence as a physical object. For electronic
records, this qualification has been aptly stated in the InterPARES report Requirements for Assessing
and Maintaining the Authenticity of Electronic Records: “When we refer to an electronic record, we
consider it essentially intact and uncorrupted if the message that it is meant to communicate in order to
achieve its purpose is unaltered.”

The archival requirement for integrity, and therefore for authenticity, depends on the message intended
by the record creator. The interpretation of a record depends on the reader as much as the creator;
therefore it is beyond the control of the record preserver. Whatever interpretation is made, it must be
consistent with the creator’s intended communication, and that undeniably involves the information
content of the record, and it may involve the way the content is presented. Here again, neither
requirement is absolute. If an accident, such as water damage, resulted in the loss or blurring (content or
presentation) of a few words in a document, we would say that the record is damaged, but not that it is
inauthentic. The strength of the requirement for unaltered content and presentation depends on the
intended use of the record. The requirement is greatest when the user wants to see the “original,”
regardless of interpretation. But for many valid uses, it is sufficient if the content—of all or part of the
record—is substantially intact and unaltered and its presentation is basically consistent with the original.

If this were not true, it would be impossible for historians, political scientists, and others to cite original

4 Requirements for Assessing and Maintaining the Authenticity of Electronic Records, Appendix 2.
5 In this case, “original” refers to the state of the record at the moment of its creation, rather than to a durable physical
object.
sources in analytic works.

As stated in the Requirements, the requirements for authenticity of copies of records depend on the purposes for which the copies are made. In the case of electronic records, all access after the records have been stored is to reproductions of the records.

"Original" Electronic Records

The transformations entailed by storage, retrieval, and presentation of electronic records make the concept of original record of uncertain applicability in the domain of electronic records. The traditional concept of an original record is tightly coupled with its inscription on a specific physical medium. The same record inscribed on any other unit of physical medium is not considered to be an original, but a copy. Given that electronic records are not stored even on the same type of medium used to present them to humans, and that the physical inscription of the record on a storage medium is fundamentally different from the inscription on a display device, a strict application of the traditional concept of original record would mean that the original ceases to exist at the moment it is committed to storage and deleted from the video device the writer used to create it. Therefore, the closest we can come to an electronic original, once the record has been set aside, is a copy in the form of the original. Given that electronic records are not stored in their original form, to produce a copy in the form of the original we need to maintain information about that form and also about the methods that are needed to translate between the storage representation and the presentation for use.

In the digital environment, there is an important distinction to consider in the concept of the form of the original. With records in hard copy, the form of the original is effectively the form in which the record was inscribed on a medium, because in inscribing the record the writer fixed the information content in a determined form. The inscription expresses the writer's intent. With electronic records, concern is often expressed about preserving the "look and feel," that is the presentation features, of the record; however, there are elements of presentation that the writer cannot fix in an immutable form, but can be changed at whim by any user.

Depending on the software used, simply changing the size of the window in which a document is viewed can change properties such as character size or line length. Similarly, changing the magnification or "zoom" ratio, or switching between draft and print image modes in word processors alters the appearance of a document on a screen.

Unless there is evidence that the writer intended to fix the presentation of the record, changes such as window size or magnification would not be regarded as producing a different record, especially when the software that permits such changes also makes it easy to reverse them. Such variations do not alter a record, as such, any more than viewing it under a magnifying glass changes a record on paper.

The Need to Reproduce Electronic Records

The transformations entailed by storing, retrieving, and presenting an electronic record led to the recognition that, in literal terms, you cannot preserve an electronic record, you can only preserve the ability to reproduce the record. A logical corollary to this assertion is that the only real way to prove that an electronic record has been preserved is to reproduce it. While the production of copies is usually seen as part of the archival reference or communication function, in the case of electronic records it is also within the scope of preservation. These functions overlap at the point of reproduction. The emphasis of the reference function is that the copies produced respond to the interests and requests of users, while the preservation function emphasizes the production of certifiably authentic copies.

Reproducing the record involves both its intrinsic and extrinsic elements of form. With respect to the bitstreams that are maintained in storage over time, an electronic record contains one or more digital components. The first step in reproducing an electronic record is to reconstruct it by assembling all of its digital components in the proper arrangement. The second step is to present or render each of the components individually and all of the components collectively in the proper documentary form. The final

---

6 E.g., magnetic and optical storage media are drastically different from CRT or LCD displays.
process in reproducing the record is to re-establish its immediate context. The immediate context of a record is its archival bond: the relationship between a record and other records. This is a two-step process. The first step consists of re-establishing the structure of the set of records in which the record belongs. The second step is to populate that structure with the relevant records.

The "Chain of Preservation"

An electronic record "in storage" is simply not the same as it was either before being stored or after retrieval. To justify belief that an electronic record retrieved from storage is the same in all essential respects as the record previously placed in storage, the rationale that is applied to hard-copy records—namely, that a physical object has been under continuous control that prevented it from being altered—is not sufficient. Given that the storage and retrieval processes for electronic records inevitably entail physical and representational transformations, the traditional concept of an unbroken chain of custody needs to be expanded to encompass the processes that are necessary to ensure that an electronic record is transmitted over time without inappropriate alteration. This expanded concept can be called the unbroken chain of preservation: the entire process of committing an electronic record to storage, maintaining it in storage, retrieving it, and presenting it must adequately preserve all its essential attributes in order to support a credible claim that the retrieved electronic record is authentic. In addition to what is entailed in the chain of custody, the chain of preservation will include information about the records creator's practices to support a presumption of authenticity, in accordance with the benchmark requirements for authenticity, information about the processes of bringing the records into the archives and maintaining them over time, and information about the reproduction of records, in accordance with the Baseline Requirements Supporting the Production of Authentic copies of electronic Records. In the digital environment, where records are not affixed in stable fashion to durable media in the forms in which they are presented for use, it is necessary to invoke an additional principle: the unbroken chain of preservation. It is not sufficient to hold on to the records. We must also ensure that any action that affects the way the records are presented protects their integrity.

Preservation, Conservation, and Maintenance

These concepts reflect a substantial departure from prevailing thinking about digital preservation. Most attention in this area has been focused on overcoming technological problems of obsolescence and media fragility. The focus of the Preservation Task Force is not on dealing with technological problems, but on achieving the positive objective of transmitting authentic electronic records over time and generations of technology. It is this objective, and the archival requirements attendant to it, that define the parameters and criteria for selecting among technological alternatives and evaluating the success of approaches and actions for preserving the records. While it is possible to compare the merits of different approaches to obsolescence and media fragility from a purely technological perspective, technology alone cannot determine what is the best choice simply because what is best depends primarily on the purpose for which an action is taken, not on the method of acting.

Within this view, steps taken to counteract obsolescence and media fragility may be termed conservation actions and measures taken to avoid or minimize the effects of obsolescence and media fragility may be regarded as maintenance activities. Conservation and maintenance are part of preservation. However, preservation activities are not limited to solving or avoiding technological problems. Preservation actions—such as media migration, storage system updating or replacement, use of different software, and even changing the data formats in which the digital components of the records are stored—may be taken not only to solve problems but also simply because new and better alternatives have been developed in information technology.
The Process of Preserving Electronic Records

The “Preserve Electronic Records” Model

The InterPARES Preservation Task Force has focused its efforts on the articulation of a formal model of the process of preserving electronic records. The process model was articulated using the Integration Definition (IDEF) methodology adopted by the International Team. Specifically, the model was articulated in accordance with the IDEF(0) standard for function modelling.7

In IDEF(0), “A function model is a structured representation of the functions, activities or processes within the modeled system or subject area.” The “Preserve Electronic Records” model is intentionally generic. It identifies and describes the processes necessary to preserve electronic records, articulates the inputs needed by each process, the controls under which it operates, the mechanisms necessary to accomplish the process, and the output(s) produced by each process. The model defines the relationships among these entities and processes. It should be interpreted as describing a subject area, rather than a specific system. That is, while the model is systematic, it does not prescribe an implementation. Rather than defining a preservation system, the "Preserve Electronic Records" model provides a comprehensive, precise, and coherent road map which institutions and persons concerned with the preservation of electronic records can use in designing, developing, and evaluating systems that address their specific requirements, objectives, and constraints.

The basis for the content of the preservation process model is the Open Archival Information System (OAIS) Reference Model, which is an ISO standard.8 “An OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community.” The “Preserve Electronic Records” model is built on the basic assumptions of the OAIS that the records are produced outside of the archival system, that they are to be available to a user community that is also outside of the archival system, and that the archival system is thus a mediator which takes information from producers and delivers it to users over long periods of time. Thus the OAIS model has a much broader scope than the "Preserve Electronic Records" model. The reference model is intended to apply to any type of information, not just records. For example, the information preserved in an OAIS might be scientific data, or it might be information about physical objects in a museum. At a high level, it may be said that the "Preserve Electronic Records" model is a specification of an OAIS for the specific classes of information objects comprising electronic records and archival aggregates of such records.

---

7 Draft Federal Information Processing Standards Publication 183, Integration Definition For Function Modeling (IDEF0), 21 December 1993.
Figure 1. Open Archival Information System

Here again it is necessary to distinguish between the function described by the "Preserve Electronic Records" model and a system that would implement the model. The preservation function might be carried out by a system that provides only the functionality described in the model. But it might equally well be implemented in a system that includes additional functionality, including the appraisal of records, the management of current and temporary records, and reference and dissemination functions.

This reveals another aspect in which the "Preserve Electronic Records" model is narrower than the OAIS: the preservation model does not include all activities related to making records available, but only those that are inextricable from the preservation function. The preservation function extends to the production of copies of records, because that is necessary to guarantee their authenticity, but it does not include order agreements as described in the OAIS model or any "value-added" dissemination or access services. Similarly, the preservation model does not include processes that inform potential users what records are being preserved or what conditions govern access to the records.

The boundaries of the preservation function model derive from the viewpoint according to which the model is constructed. IDEF(0) models “functions (actions, processes, operations), functional relationships, and the data and objects.” The relationships between functions are logical, and not necessarily chronological. IDEF(0) does not explicitly model temporal sequences. Moreover, in IDEF(0), the viewpoint determines what can be ‘seen’ within the model context, and from what perspective or ‘slant’. Depending on the audience, different statements of viewpoint may be adopted that emphasize different aspects of the subject. Things that are important in one viewpoint may not even appear in a model presented from
The horizon for the viewpoint of the preservation model is determined by the scope of the InterPARES Project as whole. The project is concerned with the preservation of electronic records that have been selected for preservation when they are no longer needed for the practical purposes for which they were originally created. The scope of the InterPARES Project corresponds to that of the OAIS box in the centre of Figure 1. Therefore, the process described in the “Preserve Electronic Records” model begins with the transfer of the records from their creator, or from an agent acting for the creator, to a person whose primary responsibility is that of preserving authentic records; that is, the preserver. However, the preserver, as defined by the InterPARES Project, has responsibilities that are broader than the preservation process itself. For example, the preserver is presumed to be responsible for selecting the records that are to be preserved. In the “Preserve Electronic Records” model, the viewpoint is literally and strictly that of “the person responsible for preservation.” The model’s viewpoint includes only those entities and processes connected with someone, or some organization, carrying out the role of preserving the records. The same person or organization may have other roles or other, coincidental responsibilities, such as appraisal or reference, but coincidental responsibilities are excluded from the “Preserve Electronic Records” model. The role of preserving records includes all and only those activities necessary to ensure the transmission of authentic electronic records over time, according to the concept of preservation as described earlier in this report.

In contrast to the OAIS model, the viewpoint of the “Preserve Electronic Records” model only includes those aspects of submission and preservation that relate directly to transforming Submission Information Packages into Archival Information Packages, and it only includes those aspects of dissemination that relate to reproducing electronic records or providing requesters with the wherewithal to reproduce the records themselves. Although the OAIS model includes determining what will be submitted to the OAIS and who is the designated customer community, these activities are beyond the scope of the “Preserve Electronic Records” model.

The viewpoint largely determines the relationships between the appraisal and preservation models. Naively, one may assume that preservation follows appraisal because records must be selected for preservation before they are preserved; however, the relationship between the appraisal and preservation models is not that of a simple sequence, but rather reflects two different viewpoints on the same overall archival process. Each of the two models includes activities that do not appear in the other, but such activities may be related through their inputs, outputs, or controls. For example, the selection of what records are to be preserved is not itself a preservation activity; therefore, selection does not appear as a process in the preservation model. However, the records selected for preservation are a major input to the preservation model. Conversely, maintaining the digital components of the selected records in storage is not a selection activity and does not appear in the appraisal model. The preservation function selects methods for preserving records comprising different types of digital components. This selection determines feasibility of preserving different types of components; therefore, it acts as a control on the appraisal function.

There are activities that appear in both appraisal and preservation models. For example, disposition of records is modelled as an appraisal function. Although the term disposition does not appear explicitly in the preservation model, the transfer of records to the preservation system is a disposition action that is included in the preservation model. Similarly, establishing the terms and conditions for transfer of records is a feedback loop between the appraisal and preservation models.

The preservation model is intended to articulate the procedures and resources required for preserving authentic electronic records. However, the requirements for preserving authentic electronic records were being developed by the Authenticity and Appraisal Task Forces, working in parallel to the Preservation Task Force. Because these task forces’ results were not available, the preservation function model has been articulated up to this time so as to be neutral with respect to requirements for authenticity. The requirements for ensuring that the preserved records remain authentic can, and should be, incorporated in the model. This is work that needs to be done in the future.

Preservation Overview

Figure 2 is the context diagram for the "Preserve Electronic Records" model. It does not show any detail about the process itself. Instead it shows what goes into and comes out of the process. Three different kinds of things go into the process: controls, which govern how the process is carried out; mechanisms, which enable the process to happen; and inputs, which are the things acted on in the process. Following IDEF(0) convention, in all diagrams, controls are shown going into the top of a process box; mechanisms are shown going into the bottom of the process box; and inputs enter the process from the left. What comes out of any process are its outputs. In IDEF(0) outputs always come out of the right side of a process box.

The context diagram shows that three factors control the process of preserving electronic records. In order to preserve records, and especially to preserve them as authentic, we need to know what the requirements are for doing so. These requirements derive from archival science and principles and related standards and best practices for managing records. These requirements are labelled as "Archival Requirements" in the diagram. Preserving electronic records entails using digital information technology. The possibilities for doing so are limited by the state of the art of information technology, which constitutes the second type of control on the preservation process. The state of the art of technology includes products and services available in the marketplace, and also the feasibility of developing systems or applications for preserving records from these products and services, and also standards. Finally, the exercise of the preservation function will also be governed by requirements of the institution in which this function is carried out.

The diagram shows three mechanisms that are necessary to perform the preservation process. They are...
an information and communications technology infrastructure, facilities where the electronic records will be stored and processed, and persons responsible for the process. Although the state of the art of technology determines what is possible and impossible to do, the technology infrastructure comprises the hardware, software, and physical media used to store and process the digital components of electronic records. The brackets surrounding the points on the three mechanism arrows indicate that these mechanisms are used in all preservation activities; therefore, they are not shown in the more detailed diagrams that follow below.

There are two primary inputs to the process of preserving electronic records. The first, and most obvious, are transfers of electronic records selected for preservation. In simple terms, the records are what the process is all about. Records are preserved because they have been determined to have enduring value. That value is realized in use. So the second primary input consists of requests for the records, or for information about them. The preservation process also needs a third input, information about the records that have been selected for preservation. This information is necessary to determine what information technology, facilities, and staff will be needed to preserve the records and to organize the process to guarantee that the records can be preserved as authentic.

The Main Preservation Processes
Preserving electronic records involves four processes: managing the preservation function, bringing records into the preservation system, maintaining them over time, and outputting them. These processes are depicted in Figure 3. This diagram is rather complex, but can be easily understand by tracing the basic path that each of the inputs follows.

Figure 3. The Main Preservation Processes
The first input, in chronological sequence, is Information about Electronic Records Selected for Preservation. It is input to the process named Manage the Preservation Function. This management process is unique in that it controls the other three basic processes. The management process takes what must be done and determines how it should be done and what the results should be. The other three processes carry out preservation activities according to the parameters established by management. The management process has the same basic relationship to all three execution processes, as depicted in Figure 4.

Figure 4. Management and Execution of Preservation Processes

Managing the Preservation Function:
The management process synthesizes the external controls to determine how preservation should be accomplished and what the results should be. In each case of records selected for preservation, it articulates the archival and technical requirements for preserving the records taking into account their documentary form as records, their archival bonds, their digital components, and the requirements for producing authentic copies of the records; specifies procedures; selects and acquires technology appropriate for satisfying these requirements; and establishes criteria for documenting the preservation process and for determining if the process has been carried out successfully. The ensemble of specific requirements, and the means of meeting the requirements and objectives in doing so constitute a preservation strategy for the records selected for preservation. The preservation strategy is thus an output of the management process which functions as a control on each of the execution processes. The preservation strategy for a body of records will specify software to be used for all processing of the records. This software constitutes preservation methods and also controls each execution process. The preservation methods require information technology infrastructure which includes hardware, media, and general-purpose software, such as operating systems, database management systems, storage
subsystems, and communications protocols. The management function selects and acquires the information technology infrastructure that is used in carrying out the three execution processes.

The management function also sets up preservation action plans for actions that must be done in each execution process, either at specified times or under specified conditions. For example, preservation action plans should specify how to determine, in each case, if the terms and conditions of transfer—established when the records were appraised—have been satisfied and what to do if not. The preservation action plans are input to each of the execution processes.

Each execution process should output management information about the process. This information is sent as feedback to the management function to enable it to evaluate the execution of preservation processes and determine if preservation strategies or methods, or infrastructure should be changed. The evaluation will also determine the feasibility of preserving different types of electronic records. This determination will be used during the appraisal process.

This description of the relationships between the management and execution processes applies to all three execution processes. It will not be repeated in the description of those processes, thus simplifying their descriptions.

**Bringing Records into the Preservation System:**

The second input to the preservation process comprises actual Transfers of Electronic Records Selected for Preservation. The transfers are input to the Bring in Electronic Records process, as shown in Figure 5. This process determines whether the transferred records are accessioned or rejected. If they are accepted, they are sent to the Maintain Electronic Records process. If rejected, they are returned to whomever submitted them. In either case, information about the transfer is sent to the management function where it is combined with information received from appraisal about records selected for preservation in developing or modifying preservation strategies.

![Figure 5. Bringing Records into the Preservation Process](image-url)
Reproducing Electronic Records: The third input to the “Preserve Electronic Records” process consists of Requests for Records and/or for Information about Records. Responding to such requests is the ultimate objective of preserving records; moreover, as explained earlier, producing copies is the final step in the process of preserving electronic records. Given that the model is constructed from the viewpoint of the person responsible for preserving the records, this final step is included in the model.

However, the model does not include all facets of responding to requests for records or information. It is assumed that certain preliminary steps, and possible subsequent ones, related to the requests are taken by persons responsible for access to the records. For example, the persons responsible for the access function will help requesters to identify the records or information in which they are interested, and will determine if they have a right to receive the records or information they request. As with the appraisal function, there will be some overlap between the preservation and access functions. If a model of the reference or access function were articulated, it would include activities that coincide with or overlap some of the activities within the Bring In process because the examination of the records and related information during that process would provide the opportunity to develop the information needed for producing finding aids. Similarly, the reference function would probably determine the computer interface used to provide requesters with access to electronic records, but preservation would determine how the records need to be presented in that interface in order to guarantee that the reproduced records are authentic.

Figure 6. Reproducing Electronic Records

Figure 6 depicts the Output Electronic Records process of the preservation model. The requests that are input to this process can be for records or for information about records, or both. Although the reference function is responsible for helping requesters to identify the records of interest to them, descriptions and
other finding aids will not contain all the detailed information about the records maintained by the preservation function. For example, a requester who needs a copy of the record that is certified as authentic may inquire whether such certification is possible before requesting the copy. Information about the records being preserved is maintained by the Maintain Electronic Records process. When a request for information is received, the Output process formulates the request in a way that the Maintain process can respond to. For example, most of the information about the records is probably maintained in a database. The Output process ensures that the request is in a format that can be executed as a query against the database. The Maintain process retrieves the information and sends it to the Output function. The Output process determines if the retrieved information is complete and delivers it to the requester. The Output process may also need to provide an explanation of the response.

For example, a requester asks if it is possible to produce a copy of a record that can be certified as authentic. If the creator had kept the record in a system that had limited export capability, what was transferred to the archives may have been a plain-text version of the record, losing the fonts, italics, and other aspects of presentation the record had in the creator’s system. In such cases, the preserver could certify that a copy contained the complete and correct contents of the records, but that it could not certify that the appearance of the record was identical to the original.

When the request is for records, the same sequence of steps between the Output and Maintain processes is followed to return the components of the requested records to the Output process. Then the records are reconstituted from their components. What happens after that point, however, depends on what the request specifies should be delivered. If the request is for an electronic copy of a record, in most cases that copy can only be produced on a system that is under the control of the preservation function. The final step in reproducing an electronic record is to render it with the appropriate presentation features. If the system on which the copy is presented is not under preservation control, there is no guarantee that it is properly presented.

In many cases, requesters will want to access electronic records on their own systems, either over the Internet or on digital media that can be read by their systems. In such cases, the Output process most often will not actually produce copies of the requested records. Rather it will deliver the digital components of the records along with instructions on how to reproduce the records from these components. The character of these instructions will vary depending on the delivery specifications in the request. For example, for records that will be rendered in a Web browser, the instructions for reconstituting and rendering the records will be packaged together with the components and executed automatically by the requester’s system. Even with such automatic reproduction, the preserver cannot guarantee that a copy on the requester’s system is authentic. For example, the requester’s system may not have all of the software needed to render the records properly. In cases where the requester wants to bring electronic records into its own application, there may be a need to invoke middleware that mediates between the preservation system and the target application. The preserver should choose software mediators that protect end-to-end the integrity of the digital components that are transmitted between the two systems, but even here the preserver cannot certify the authenticity of copies produced in requesters’ systems. In some cases, instructions for reproducing records from digital components will have to be in human-readable form.

These variant scenarios for the reproduction of copies of electronic records explain why the Output process has the two distinct outputs: Reproduced Electronic Records and Reproducible Electronic Records. The first of these outputs is produced when the records are presented on a system under preservation control. The second, consisting of digital components and instructions for reproducing the records, is produced for delivery to a system outside of preservation control.

The "Preserve Electronic Records" process is designed to enable the production of authentic copies of electronic records. In some cases, requesters may want the preserver to attest to the authenticity of the copies. If so, the Output process checks the preservation history kept by the Maintain function to ensure that the chain of preservation for the requested records is intact; that is, that there was an adequate basis for presuming the records were authentic when transferred to the preserver’s custody and that the records have been properly preserved since that time, up to the point of reproduction. If these conditions have been satisfied, the Output process issues a Certification of Authenticity. In cases where digital components are delivered to external systems, the instructions for reproducing the records should aim at
the production of authentic copies. Accordingly, the Output process should also provide the requesters with criteria they can use to determine if the copies are authentic.

**Figure 7. Maintaining Electronic Records**

Maintaining Electronic Records over Time:

One of the four basic processes remains to be described, Maintain Electronic Records. This is an internal process: none of the inputs that come into the preservation process from the outside go directly to this sub-process, and the outputs of the maintenance process go to other sub-processes within the function. Nonetheless, the Maintain process is the core preservation process with respect to transmitting electronic records over time. It is depicted in Figure 7. The Maintain process is connected to both the Bring In and Output processes. The digital components of the records—along with information needed to reproduce them, re-establish their archival bonds, and certify their authenticity— are received from the Bring In process. Requests for records and information are received from the Output process and the results of executing such requests are returned to Output.

The diagrams of the four preservation processes (Figures 4 to 7) depict all of the information entities that are in the comprehensive diagram of the four processes (Figure 3), except for two outputs of the Manage process: Information about Preservation and Report on Authenticity of Records. The first of these, Information about Preservation, reflects the fact that the preserver will undoubtedly be responsible and accountable to others. In the case of an institutional archives, such as that of a corporation or university, the preserver will be responsible to the institution itself. In the case of government archives, not only will the preserver be responsible to higher levels of government, but it will also be accountable to the people. In any case, the preserver will need, and will want, to communicate information about its activities. The preserver will need to be able to produce a Report on Authenticity of Records to justify its methods and procedures.

To this point we have examined the four preservation processes depicted in Figure 3 as they relate to one
another. In the following sections, we consider the activities that take place within each of the four preservation processes.

**Preservation Sub-processes**

**Manage the Preservation Function**

Managing the Preservation Function determines how all other preservation processes will be carried out, and with what results; selects and acquires the necessary technology; and evaluates the execution of the function.

The management process consists of four sub-processes: Determine Preservation Requirements, Select Preservation Methods, Specify Outputs and Outcomes, and Evaluate Execution of Preservation.

Preservation requirements are determined by generic archival principles (such as provenance and authenticity of records) and specific institutional requirements (such as accessioning and dissemination policies and information technology standards), as well as information about the records to be preserved and knowledge derived from evaluation of the preservation of records already transferred. This process integrates and synthesizes external controls to determine the specific archival requirements for preserving and reproducing the records. This requires the specification of the attributes and characteristics of the records that must be preserved; a knowledge of how the records are composed from their digital components; an understanding of how records are to be grouped in the proper order; and a determination of how the authenticity of different classes of records will be certified.

Given these archival requirements for preservation, and working within the limits imposed by the state of the art of computer science and information technology, specific technological methods are identified, evaluated, and selected to be used in all aspects of preserving the different classes of records and archival aggregates selected for preservation: bringing them into the preservation system, maintaining them over time, and reproducing them in authentic form. The selection of these methods will be influenced by the institution’s IT architecture, data standards and related procedures, security requirements, and access restrictions. Selecting preservation methods entails the identification and evaluation of available preservation methods; the selection of the method(s) that meet the archival requirements for preservation of each class of records; and the acquisition and adaptation, configuration or enhancement of the technology and other resources necessary to apply the selected method. Each preservation method will have a specified domain of application. Some methods, such as physical media and storage systems, may be used for all records and digital components. Others will be specific to given classes of records or types of digital components.

A preservation method and its scope of application form the basis for a preservation strategy. The strategy is completed by specifying what outcomes will be produced by its application to the relevant domain. “Specifying Outcomes” sets the objectives and performance targets for the operation of the method and the means for identifying, measuring, and reporting on the achievement of the objectives and targets. The outcomes should also encompass the results of risk assessments and indicate how to handle exceptional and problem cases.

In order to ensure that the objectives of the preservation function are being realized, and to maximize its performance, its execution has to be evaluated. Evaluating Execution of the Preservation Function uses feedback from the processes of bringing in records, maintaining them, and reproducing them. Each of these processes is required to output Management Information about Preservation. The evaluation sub-process measures performance against objectives, and identifies areas for improvement, at the micro level of objectives and targets and/or at the macro level of strategies and methods. Moreover, its output Information about Preservation acts as an input to the Appraisal function, transmitting information that will influence the Determine the Feasibility of Preservation sub-process of that function.

Taken as a whole, Manage governs the preservation function as a dynamic process in which the changing nature and volume of records to be preserved; shifts in institutional requirements; the evolving capacity of IT; and feedback from the function itself combine to produce preservation strategies, methods,
and objectives that evolve over time.

**Bringing Electronic Records under Preservation Control**

In accordance with the preservation strategies established in the preservation management process, electronic records are brought under preservation control. Bringing electronic records under preservation control includes four activities: registering the record transfer, verifying the authority for transfer, examining the records, and accessioning the records.

The process starts with the transfer of records selected for preservation from the submitter. The person responsible for preservation will first determine that there is no evidence of problems occurring in the process of transfer. (If problems are identified, the submitter should be asked to resend the materials.) Registering the transfer captures information about the transfer such as submitter’s name, record creator’s name, and current date which is contained in documentation accompanying transfer. Registration establishes basic control over the materials transferred by assigning a unique identifier to them. The next step involves verifying the authority for transfer, also based on information about the transfer. A transfer is determined as authorized if and only if it comprises records that have been selected for preservation and those records have been submitted either by the records creator or an agent for the creator.

Examining the electronic records that have been transferred is the principal means of bringing the records under preservation control. It serves four purposes: to determine whether the archives will preserve the records; to identify preservation strategies to be used; to determine when preservation interventions should occur; and to identify, produce, and capture information necessary to assert the authenticity of the records. A transfer includes digital files and information identifying what records and digital components are contained in those files. The first step in examination is to determine where the records and components are located in the transferred files and how they are identified. This is achieved by reading the digital files to verify the information accompanying the records.

Once the records and components have been identified in the transferred material, they are examined to determine if the terms and conditions of transfer are satisfied. This includes determining that all components necessary to compose all the records that should be included in the transfer are present and intact; that the formats of the components and the methods necessary to reconstruct and render the records are known; and that there are sufficient data to reconstruct the files, series, and other archival aggregates in which the records are organized. The examination should also determine if any preservation intervention, such as migration or transformation to persistent form, is necessary to enable the records to be preserved using applicable preservation strategies. If so, the necessary interventions should be identified. They will be carried out in the process of maintaining the records.

The examination should also include a final review of information that provides a basis for presuming that the records were maintained as authentic. That information then becomes part of the preservation history of the records. It will be retained, and augmented in the course of maintaining the records over time in order to document the "chain of preservation" and enable copies of records to be certified as authentic.

These activities enable the archives to accession the records or reject the transfer.

**Maintaining Electronic Records over Time**

Electronic records are stored as digital components, which may be separate digital files, or contained in a single digital file. The preservation function aims to deliver records that convey the intellectual content and intent of the creator. To do this, it is necessary to reassemble the components to reconstitute the records and present them in their original documentary form. This has to be done every time the record is accessed simply because an electronic record cannot be stored in the same form in which it is presented to humans. Strictly speaking, maintaining electronic records over time means maintaining the ability to output them. Maintaining electronic records, thus, requires storing their digital components and

---

10 This information was reviewed previously as part of the appraisal process.
maintaining information about the records, such as what digital components they contain, how those components are related to each other and the records, and how the records should be presented. In order to be able to output authentic electronic records, it is also necessary to maintain information that justifies an assertion of authenticity.

The activities needed for maintenance of electronic records over time include putting record components into storage, managing the storage of the record components and information about records, and maintaining the ability to retrieve components and reproduce the records. These functions and operations are governed by preservation strategies.

Maintaining Electronic Records requires managing information about the records and record components which, in the first instance, can be viewed as belonging to one of three classes: intellectual, technical, and administrative. The intellectual information includes: provenance; documentary context; and description of contents, structure, and form. The technical information includes metadata about the components, their relation to the records, and the methods required to reconstitute the records and render them on output media, which may be digital, such as a computer screen, or analog, such as paper or microform. Maintainability also requires administrative information such as storage information (identification and location storage media, digital files, and digital components; type of media, health of the media), and the history of actions taken to maintain the records and to prevent inappropriate alteration. The three types of information are used both in other processes related to maintaining electronic records and in responding to requests for records or information about records.

The second process necessary to maintain electronic records is to Manage Storage, which involves putting the digital files containing the components into storage; monitoring storage both to identify any damage or deterioration that may occur and to determine when it is necessary to refresh or migrate storage media or storage systems; and correcting any storage problems occurring, including disaster recovery and retrieving components in need of maintenance or for reproduction. The basic objective in all processes involved in managing storage is to keep the bitstreams that comprise the digital files and digital components intact and retrievable. Managing storage does not involve any preservation intervention that would change the digital components, such as by migrating them to new data formats. When corrective actions are taken to address problems that could alter the contents of one or more digital components, the affected records should be reproduced in order to prove that the corrective action was successful or to document the impact on the records being preserved. Managing storage entails reporting information about changes in the media or the location of the files and components, and about any actions taken to prevent or recover from storage problems. As appropriate, such information is used in managing information about the records or in managing the preservation function.

The third type of process covers actions taken to maintain the ability to output electronic records. Such actions will be taken either when the "Manage Preservation" process determines to change a preservation strategy affecting the reproduction of any records being preserved or when records are examined during the process of bringing them into the preservation system and it is determined that some technical intervention is needed to enable the records to be reproduced under applicable preservation strategies. The types of actions will depend on the methods selected in applicable preservation strategies. Under some strategies, such as migration and persistent object strategies, the actions may change the digital components themselves. Other strategies—such as maintaining the original computer systems and emulation—will leave the digital components unaltered, but change the hardware or software used to reconstitute and render records. In all cases, however, preservation strategies determine both the formats in which digital components are stored and the methods applied to them to reproduce records, files, series, etc.

Thus, whenever a preservation strategy that affects the reproduction of electronic records is changed, the new strategy should be evaluated to ensure that the records can still be reproduced and to document any impact of the change in strategy. Similarly, whenever digital components are altered to conform to preservation strategy, the result should be evaluated. The evaluations should be performed at the level of individual records and, when sets of records, such as databases or case files, are impacted, at the aggregate levels as well.

Evaluating the success of such technical interventions requires reconstituting and rendering the record(s) and/or archival set(s) based on the reproduction strategies and authenticity requirements. In both cases,
any impact on the records should be documented. The results are needed both for managing information about the records and for managing preservation overall.

Reproducing Electronic Records
The ultimate objective of preserving electronic records is to transmit them over time to users need them or have an interest in them. The process of preserving an electronic record is “complete” only at the point where the record has been reproduced in authentic form. Providing an electronic record entails a process of reproducing or reconstituting it from its digital component(s). The person responsible for preservation needs to ensure that this process can be executed to output authentic records.

The process starts with the receipt of a request for a record, or records. It is presumed that, prior to a request coming into the preservation process, the access or reference function has determined that the requested records are available and that the requester has a right to them.

Reproducing an authentic electronic record requires retrieving the digital components of the record and using the methods required to reconstitute the structure and content of the record from its digital components to presenting the record in appropriate form. Certifying the authenticity of the reproduced record requires information (i) supporting the presumption that the creator preserved the records as authentic up to the time the records were transferred to the archives and (ii) the audit trail of their preservation from the time of transfer to the process of reproduction.

The first step is to retrieve the digital components and related information. Then, the methods dictated by the applicable reproduction strategy can be applied to the components to reproduce the record. Depending on the request, this can be done by the person responsible for preservation, by a person responsible for access or dissemination, or by the requester.

When someone else will carry out the reproduction of records, the person responsible for preservation has to provide the components and the information needed to reproduce the records and to support an assertion that the output record is authentic. In any case, the preserver should conform to the Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records set out in Requirements for Assessing and Maintaining the Authenticity of Electronic Records. If another person carries out the reproduction process, the preserver should provide that person with either assurance or evidence that all requirements up to the reproduction process have been satisfied, along with specifications that apply to the reproduction itself. These specifications will articulate the baseline requirements as they apply specifically to making authentic copies of the records in question.

Conclusion and Recommendations
Much attention to the preservation of electronic records has focused on the twin problems of the relatively short life expectancy of digital media and the rapid obsolescence of hardware and software. The InterPARES Project started with recognition of these problems and cast the preservation issue in terms of evaluating practical methods for solving them. The research plan called for the Preservation Task Force “to identify and develop the procedures and resources required for the implementation of the conceptual requirements [for preserving authentic electronic records] and criteria [for appraising electronic records] identified in the first two domains.”11 This formulation of the problem of preserving electronic records clearly situates it not in technology, but in the interface between the goal of preserving electronic records and the technology on which they depend. Technology itself is not a problem. If we did not need to preserve records beyond the life expectancies of hardware, software, and digital media, we would not have any preservation problem. Similarly, technology cannot determine the solution. It is archival and records management requirements that define the problem. It must be archival and records management criteria that determine the appropriateness and adequacy of any technical “solution.” The question “What

11 <http://www.interpares.org/researchplan.htm>
is the best technological method for preserving electronic records?” is as meaningless as the question “What is the best medicine for making people healthy?” Neither can be answered without specifying the conditions it is meant to address. The InterPARES Project defined these conditions as the archival requirements for authenticity and the archival criteria for selecting records to be preserved.

As previously stated, because the InterPARES task forces on authenticity, appraisal, and preservation worked in parallel, the Preservation Task Force could not formulate solutions based on specific conceptual requirements and criteria. Nonetheless, through communications and cross-fertilization among the task forces during the entire course of the research, the Preservation Task Force has been able to produce a model of the process of preserving electronic records that does in fact identify the procedures and resources needed to implement the requirements and criteria. The procedures are the processes defined in the "Preserve Electronic Records” model, and the resources include the mechanisms needed to carry out these processes as well as the information about both the processes and the records that needs to flow across processes. This model does not describe a computer system, and it does not itself reach conclusions about what technological systems, tools, or methods are best suited for preserving electronic records. Rather it provides an extensive, detailed, and highly coherent framework for identifying and analyzing the specific challenges faced in implementing appraisal decisions that select specific bodies of electronic records to be preserved. This framework guides the evaluation of technological options and the articulation of specific preservation strategies addressing both the archival and technological characteristics of the records to ensure the continuing availability of authentic copies of the records across time and generations of technology.

Thus the "Preserve Electronic Records" model can be a guide to implementation, but it does not prescribe an implementation. There is greater value in this model than there would be in one that described how to design a particular preservation system. It would be simplistic, and erroneous, to assume that a single technical solution would be optimal in all circumstances. The "Preserve Electronic Records" model can be used to develop solutions that address varying circumstances, including not only diversity in the characteristics of the records to be preserved, but also variety in the external requirements imposed on the preserver, and in the goals and objectives to be achieved in preserving the records.

Recommendation 1.
The primary recommendation that comes out of this work, then, is for analysts and institutions to use the "Preserve Electronic Records" model as a framework for developing solutions to the challenges of preserving electronic records.

Recommendation 2.
Use of the "Preserve Electronic Records" model should be based on an understanding of the particular characteristics of electronic records and what those characteristics entail for preserving these records, as summarized in the foundation concepts that were set out earlier:

- Digital Components of Electronic Records
- Preservation Control
- Archival Requirements for Preservation
- "Original" Electronic Records
- The Need to Reproduce Electronic Records
- The Chain of Preservation.

The key to all of these concepts is the recognition that the chain of preservation for electronic records must extend over their entire life and that the process of preserving electronic records extends to and includes reproducing the records.

Recommendation 3.
Solutions to the preservation of specific bodies of electronic records should be inherently dynamic. The reason is twofold. First, most archives and other preservers will accumulate electronic records over time. Over time, the specific properties of the records brought into the archives will change. The preservation
system must be capable of being expanded, adapted, and modified to accommodate new and different types of electronic records, and new ways of organizing, accessing, and presenting such records. Second, the goal of preserving electronic records is not to keep them, in archives or elsewhere, but to make them available to persons who have a need for, or an interest in, them. While the preserver has a fundamental responsibility for providing access to authentic records, their availability will be impacted by the continuing evolution of information technology. Preservers should assume that future users will want to use the best available technology for access to the records. Preservation solutions should be designed to be able to interface with evolving technologies for information discovery, retrieval, communication, and presentation.

Recommendation 4.
The InterPARES Project has been so fruitful that it has not only provided valuable products in response to the research questions that it originally posed, but it has also raised the threshold of research by articulating issues that are entailed in the original questions, although not explicit in them, by identifying new questions, and by opening up lines of research that should provide grounds for valuable results for years to come. For example, the project has moved beyond its foundation in the science of diplomatics to recognize that, in the digital environment, many of the concepts and methods that traditionally were applied to individual documents need to be applied to sets of records. This insight needs to be explored more fully. The work of the Preservation Task Force has focused on defining a comprehensive framework for preserving authentic electronic records. More work is needed to analyze the data and information requirements for executing the processes defined in the preservation model. The model should also be applied to test cases both to validate and enrich it. The model should also be extended to address the application of specific technologies for overcoming technological obsolescence. The accomplishments of the InterPARES Project should be applied to related areas of concern, such as the process of archival description. In sum, this work should not stop when the current project ends. The archival profession, our collaborators, and our stakeholders, have an interest and responsibility to see that further progress is made.
Appendix 7

Walkthrough Applying The “Preserve Electronic Records” Model
Version 5.1

Preservation Task Force

Walkthrough Team:
Kevin Glick
Hans Hofman
Rich Lysakowski
Ken Thibodeau
Bill Underwood
Bruce Walton

8 October 2001
Introduction

A walkthrough of the “Preserve Electronic Records” model has been conducted. The primary purpose was to test and improve the model using data from one of the InterPARES case studies.

The Preservation Task Force used the IDEF(0) methodology to analyze the problem of preserving authentic electronic records, and the lowest-level activities of the resulting model are believed to be a solution to the problem. The task force wants this model to be of the highest quality, to be understandable to archivists and computer system developers, and apply to a broad variety of electronic records.

Walkthroughs are an effective way to improve the quality of documents that describe the analysis, design, code and user interface of a system. A walkthrough is a peer-group review of any information system product. There are a number of types of walkthroughs—activity model, data flow model, and user requirement walkthroughs that are concerned with the functionality of the system; design walkthroughs that are concerned that a system or program design meets functional requirements; code walkthroughs that are concerned that program code satisfies the program design; and test walkthroughs to ensure the adequacy of test data for a system.

The invention of walkthroughs should probably be attributed to both Gerald Weinberg and Michael Fagan. Weinberg used the concept of a walkthrough in a Programmer Team approach to developing software. Fagan developed a process called Formal Inspections to deal with the problem of reducing errors in the development of large software systems. Yourdon embraced Fagan’s ideas and developed the concept of Structured Walkthroughs.

The objectives of the walkthrough of the Preservation Task Force preservation model were:

- To test the IDEF(0) model for the preservation of electronic records (version 5.1) with data from a specific case—in other words, to demonstrate that there is a real-world interpretation of the model.
- To create examples of a preservation strategy, preservation action plan(s), and targeted preservation methods that are linked to the body of records in an actual case.
- To more precisely specify the definition of an activity as a transformation of inputs to outputs.
- To identify the data elements of the inputs and outputs to activities of the model.

The walkthrough was conducted using version 5.1 of the “Preserve Electronic Records” model. Lessons learned in this walkthrough were applied to develop version 6, the final version of the first InterPARES Project.

---

1 The walkthrough was conducted at a Preservation Task Force meeting during InterPARES Workshop No. 9, 15–17 October and 20 October 2001, Rome.
The Walkthrough

The walkthrough organization was relatively informal. The roles were:

- The presenter, who "put on the table" the preservation model that was being reviewed.
- The reviewers, who had a good understanding of the preservation model and raised issues and suggested solutions to problems.
- The case study expert, who answered questions posed by the reviewers about the data from the case study.
- The secretary, who recorded the discussed facts and issues and took and distributed the minutes.

The method used in the walkthrough was to iteratively step through each of the lowest-level activities in the preservation model:

1) Reviewing the activity definition and the input, output, and control definitions.
2) Identifying data elements of labels on input and output arrows.
3) Defining the transformation of inputs to outputs.
4) Determining values of the data elements that are related to the specific body of records.
5) Recording the results and any problems or issues that arise and suggesting possible solutions.

The Case Study

Case Study 26, the New York State Workers' Compensation Board (WCB) Electronic Case Folder System, was used in the walkthrough of the "Preserve Electronic Records" model. The following items collected or prepared during the case study were used in the walkthrough:

a) WCB Electronic Case Folder Workflow
b) Questions about Business Context
c) Case Study Overview
d) Case Study Interview Protocol (CSIP)
e) Template Element Data Gathering Instrument (TEDGI)
f) Diplomatic Analysis of Case Study
g) Interview transcript
h) Round 3 and 4 Pre-interview
i) Description of System (NY Workers' Compensation Web site)\(^5\)
j) ERWIN data model of Electronic Case Folder

There was no Appraisal Report or Terms and Conditions of Transfer for this case.

The discussion in this section is structured around the activities of the preservation model. For each activity, the definition and inputs and outputs to an activity are reviewed. The reader will

\(^5\) [http://www.wcb.state.ny.us/]
notice that for some activities there is a partial discussion, or no discussion at all. Those activities
were only partly reviewed or not reviewed due to time constraints.

In the notes of the walkthrough, sources of information from Case Study 26 are indicated in
parentheses after data elements. The endnotes in this paper indicate possible modifications of
the IDEF(0) diagrams based on the walkthrough. They correct some errors and inconsistencies
and suggest some simplifications.

A1 Manage the Preservation Function

Activity Definition:
"... preservation is managed by producing a comprehensive preservation framework consisting of
sets of preservation strategies and preservation actions plans, each linked to a specific body of
electronic records selected for preservation, along with the technological infrastructure and
preservation methods needed to implement the action plans. ..."

The reviewers identified the following data elements for the three inputs to this activity.

Inputs:
1. Information about Electronic Records Selected for Preservation
   a) Record Creator’s Name: New York State Workers’ Compensation Board (Source: TEDGI
      3.1)
   b) Transfer Agent’s Name: Same as above
   c) Identification of Records
      Title: New York State Electronic Case Folder (Source: CSIP)
      Description: Series of case files for adjudicating benefits of disabled workers (Source:
      TEDGI).

      Model or Description of the Business Processes that generated these records: "Workers’
      compensation insurance provides weekly cash payments and the cost of full medical
treatment, including rehabilitation, for covered employees who become disabled as a
result of a disease or injury connected with their employment. It also provides payments
for qualified dependents of a worker who dies from a compensable injury or illness. In
administering this program, the Workers’ Compensation Board receives and processes
workers’ claims for benefits, employers' reports of injury, and medical reports from
physicians and other health care providers. The board adjudicates and resolves all issues
and makes awards and findings as rapidly as possible to ensure that an entitled claimant
receives benefits and medical treatment promptly. Hearings are conducted before law
judges, or, on review or appeal, before panels of three board members." (Source: NYS
WCB Web-site Description of the Mission of the Organization)

      A business process analysis of the business processes was used to determine
improvements in the business processes, and needs for changes in regulations and in
designing the NYS Electronic Case Folder system. It was not collected as a part of the
case study. (There may be different types of cases, depending on this business process
model.)

      Document types:
      Claims for benefits
      Employer’s reports of accidents and illness
      Correspondence
      Medical reports from physicians and other health care providers
      Insurance carrier’s reports
      (Source: supporting documents)
Volume:
There are over 300,000 open cases (Source: Supporting documentation)

File or Data Structure:
Relational Database describes the records and their relationships.

d) Disposition (Scheduled date of transfer): (Source: Supporting documents)
e) Media on which current records are stored: Net HT Series 30GB 12” OSCAR optical WORM disks
f) Information requirements to support a presumption of authenticity

2. Information about transferred and accessioned records
   a) Record Creator’s name
   b) Transfer Agent’s name
c) Date of transfer
d) Identification of records
   Title
   Description
   Volume
   File or Date Structure
   Technical Information (e.g., XML SQL, file formats, encoding)

e) Information supporting presumption of authenticity of records transferred by the records creator. This consists of evidence for each of the following requirements:
   A.1.a Identity of the Record
   A.1.b Integrity of the Record
   Requirement A.2 Access Privileges
   Requirement A.3 Protective Procedures: Loss and Corruption of Records
   Requirement A.4 Protective Procedures: Media and Technology
   Requirement A.5 Establishment of Documentary Forms
   Requirement A.6 Authentication of Records
   Requirement A.7 Identification of Authoritative Record
   Requirement A.8 Removal and Transfer of Relevant Documentation

   In lieu of a presumption of authenticity, a verification of authenticity should be provided.

3. Management Information about Preservation: This will only be available after the first transfer has been processed because it’s feedback information. However, an archive would probably use feedback it had about series with similar characteristics in developing the strategy and plan.

A1.1 Determine Preservation Requirements

A1.1.1 Identify Types of Archival Properties that must be Preserved

Input:
Information about Electronic Records Selected for Preservation

The reviewers identified case study data corresponding to the three outputs of this activity.

Outputs:
   a) Classes of Records:
      Claims for Benefits
      Employers’ reports of accidents and illness
      Correspondence
      Medical reports from physicians and other Health Care Providers
Insurance carrier’s reports
(Source: supporting documents)

b) Types of Record Arrangement: Indexed on case file identifier. Secondary indexes on other attributes of the case file. (Source: ERWIN Data Model of Electronic Case Folder System) (To articulate the preservation strategy for these records, the preserver would need to determine: Are these attributes data elements in the database? Do they have any relationship to specific business processes, such as audits?)

c) Types of Archival Bonds: Contents of case file ordered by document number. Document numbers are created when a document is imported into the electronic case folder system. The File Net High-Performance Image Import (HPII) system creates this number.

A1.1.2 Determine how Records are composed from Digital Components
Case study data were identified for one of the outputs of this activity.

Output:
Record Composition Requirements:
1. Documents in NYS Workers’ Compensation Board Electronic Case File system are page images represented as TIFF 6 files. Page images must be reproducible (Source: TEDGI, Supporting Documentation).
2. Metadata about documents and cases are represented in a relational database managed using Sybase (Source: Supporting Documentation).
3. The TIFF 6 files are stored as files external to the database.

A1.1.3 Determine How Records are Arranged
Output:
Requirements for arranging records: Must be able to present cases in order of case file identifying number.

A1.1.4 Determine How Archival Bonds are Expressed
Output:
Requirements for instantiating archival bonds: must be able to order documents in a folder by document number.

A1.1.5 Synthesize Requirements for Preservation
Inputs:
Classes of Records
Record Composition Requirements
Types of record arrangement
Requirements for arranging records
Requirements for instantiating archival bonds

Case study data were identified for the output of this activity.

Outputs:
Specified Requirements for Preservation

a) Classes of Records:
All classes of records are stored as document images.
Page layout and appearance must be preserved.
Each TIFF file contains multiple pages.

b) Types of Record Arrangement:
Records must be arranged into case files according to case file ID and document number.
Case files must be arranged by case file ID and by an index on case file ID.
c) Types of archival bonds:
archival bonds are instantiated on the basis of case file and document numbers.

A1.1.6 Determine Basis for Certifying Authenticity

Activity Description:
Guided by Archival Science and Institutional Requirements (both tunnelled to this diagram) and the Specified Requirements for Preservation applicable to a body of records selected for preservation, determine the basis for asserting the authenticity of the records. This basis will have two parts: information supporting the presumption of the authenticity of the records as transferred from the creator and information about how the preserver satisfies the applicable Specified Requirements for Preservation after the records are transferred to the preserver. The information requirements to support a presumption of authenticity of records up to their transfer is received as part of Information about Electronic Records Selected for Preservation and Information about Transferred and Accessed Records, and will be designated for retention along with other information about the records. The types of information required to support an assertion that the preserver has preserved and reproduced authentic records become a control on subsequent preservation activities.

Controls:
Archival Science

Benchmark Requirements:
Information supporting the presumption of authenticity of the records as transferred from the creator. This consists of evidence as to how the creator addressed each of the following:

A.1.a Identity of the record
A.1.b Integrity of the record

Requirement A.2 Access Privileges
Requirement A.3 Protective Procedures: Loss and Corruption of Records
Requirement A.4 Protective Procedures: Media and Technology
Requirement A.5 Establishment of Documentary Forms
Requirement A.6 Authentication of Records
Requirement A.7 Identification of Authoritative Record
Requirement A.8 Removal and Transfer of Relevant Documentation

If the evidence does not support a presumption of authenticity of the records as retained and transferred by the creator, their authenticity will have to be verified on a case-by-case basis.

Baseline Requirements: How the preserver satisfies the applicable requirements for preservation after the records are transferred to the preserver. This consists of demonstration that each of the following requirements is met.

B.1 Controls over Records Transfer, Maintenance, and Reproduction
    B.1.a Unbroken custody of the records is maintained.
    B.1.b Security and control procedures are implemented and monitored.
    B.1.c The content of the record remains unchanged after reproduction.

B.2 Documentation of Reproduction Process and its Effects
    B.2.a The date of the records' reproduction and the name of the responsible person.
    B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver.

6 Authenticity Task Force, Requirements for Assessing and Maintaining the Authenticity of Electronic Records, Appendix 2.
7 Ibid.
B.2.c The impact of the reproduction process on their form, content, accessibility and use.
B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user.
B.3 The Archival Description of the fonds includes ... information about changes the electronic records of the creator have undergone since they were first created.

Inputs:
Specified Requirements for Preservation
   Information about Electronic Records Selected for Preservation
   Information about Transferred and Accessioned Records

Outputs:
   Basis for Certifying Authenticity of Transferred Records
   Authenticity Requirements for Preservation

A1.2 Select Preservation Technologies
For any aggregate of electronic records selected for preservation, preservation technologies must be used which are appropriate and adequate for reproducing the records, maintaining the archival bonds among the records, and satisfying the benchmark requirements for authenticity. Given continuing change in information technology, the selection is not likely to be a one-time decision. Rather A1.2 is a dynamic process dependent on changes in types of digital components of electronic records, the state of the art of IT (computer storage, operating system, database technology) in which new preservation options must be identified and evaluated and new preservation methods and technological infrastructure acquired.

A1.2.1 Identify Preservation Options
Case study data were identified for each of the inputs and outputs.

Inputs:
Information about Digital Components of Electronic Records
   a) Information about current media on which digital components are stored: WORM
   b) Relationship between TIFF files and relational database: Presuming the logical model of the database cannot be reduced to a flat file, is it essential that this structure be maintained in relational form? If the structure is essentially that of a graph, could it be replicated in other methods, e.g., XML DTD, data warehouse, or OO database? Could the TIFF files be integrated into the database?

Output:
Preservation Options
   The institution might require keeping document images as TIFF files and storing digital components on DLT tapes. That would dictate a decision in this case to keep the digital components as TIFF files and to transfer the files to different media, e.g., from WORM to DLT tapes.

A1.2.2 Evaluate Preservation Options
Choice of storage media will be based on per unit cost, market penetration, transfer rate, longevity, and storage device costs which are determined by the State of Information Technology and budget, which is an institutional constraint.

Choice of archival storage system will be based on business needs and budget, e.g., frequent access to some bodies of records and infrequent access to other bodies of records might dictate the choice of a hierarchical storage system.
A1.2.3 Select Preservation Methods

Input:
Evaluated Preservation Options
Case study data were identified for the output of this activity

Output:
Selected Preservation Methods: For example:
  a) Format for document images: preserve in TIFF format with ability to reproduce documents with TIFF viewers.
  b) Format for textual documents: preserve in PDF format with ability to reproduce documents with PDF viewers.
  c) Method for reconstituting case files and, if needed, records.
  d) Metadata about records or digital components should be stored in a relational database.
  e) The schema for relational databases should be represented in SQL.

A1.2.4 Acquire Capability to Apply Selected Preservation Methods

Examples of the Technology Infrastructure needed for this case was identified.

Outputs:
1. Technology Infrastructure: Includes computers, media, storage systems, and generic software such as Operating Systems, Database Management System, Packaging Software.
   a) Viewer software: Acquire viewer for TIFF that operates on current computer hardware and operating systems.
   b) Format conversion software and hardware: Acquire hardware and software to convert WORM to DLT. Decide whether the creator—in which case it becomes a condition of transfer—does that conversion.
   c) Acquire software to migrate obsolete format to current format.

The definitions of activities that used preservation methods were used to identify the kinds of preservation methods that would be needed.

2. Targeted Preservation Methods

   Definition:
   Software used to implement a preservation strategy that is related (targeted) to a body of records, types of electronic records, and classes of digital components. The source of information about the body of records, types of electronic records, and classes of digital components is the Specified Requirements for Preservation. Note: see information about the method and the body of records, types of electronic records, and classes of digital components to be added to the output of A.1.2.4 to produce Targeted Preservation Methods.

Targeted Preservation Methods (from Definition of Activity 1.2)

- Method for checking the integrity of the transfer process.
- Methods for checking the completeness of the transfer: Are all required case files included? Are all records in each case file present? Are case ID and document numbers uniquely assigned? Do data in the database match corresponding attributes of the TIFF files?
- Methods for managing and storing of digital components in digital files and on physical media.
- Methods to reconstitute and reproduce records.
- Methods to enable others to reproduce the records.
Targeted Preservation Methods (from A3, A3.2, and A3.3)

**Storage Methods:**
- Method for placing record components in storage, e.g., packaging and preserving integrity of record components
- Storage update method
- Monitoring method
- Problem correction method
- Retrieval method
- Methods for Updating Components.

Targeted Preservation Methods (from A4)
- Record reconstitution method
- Presentation method
- Packaging Method (for dissemination).

**A1.3 Specify Preservation Strategies and Actions**

**A1.3.1 Articulate Preservation Strategy**

**Activity Definition:**
Controlled by the Specified and Authenticity Requirements for Preservation applicable to a body of electronic records selected for preservation, and limited by the State of the Art of Information Technology. Use Information about this body of Electronic Records Selected for Preservation and Information about Transferred and Accessioned Records from the same body of records, along with Preservation Technology Specifications, which describe the Targeted Preservation Methods applicable to these records, to specify and output a comprehensive Preservation Strategy for preserving the body of records.

**Controls:**
- a) Specified Requirements for Preservation: Are the criteria for evaluating execution of the preservation strategy in the Specified Requirements for Preservation?
- b) Authenticity Requirements for Preservation
- c) Evaluation of Execution (Feedback from A1.4)

**Inputs:**
- a) Information about Electronic Records Selected for Preservation
- b) Preservation Technology Specifications, i.e., Information about Targeted Preservation Methods
- c) Information about Transferred and Accessioned Records

**Output:**
Preservation Strategy

**Definition:**
A coherent and comprehensive approach for preserving a body of records selected for preservation, derived from archival and institutional requirements, taking into account Evaluation of Execution of current and prior Preservation Strategies and reflecting the state of the Art of Information Technology. A preservation strategy includes objectives for maintaining components of electronic records and related metadata and information over time and for reproducing the records in authentic form, in the order imposed by the records creator, and criteria for evaluating execution of the preservation strategy. The strategy includes specifications for handling exceptions to its standards, and identifies the targeted preservation methods to be used.

Activities A2, A3, and A4 provide a Preservation Framework from which a Preservation Strategy specific to a body of records is constructed. A Preservation Framework is a generic preservation...
strategy for achieving preservation goals such as: authentic records, retrievable records, and reproducible records.

To determine the elements of a preservation strategy, we considered how the preservation strategy is used as a control on all remaining processes in the preservation model. In parallel with this examination, we sought to determine the elements of a preservation action plan (output of A1.3.2), as they were inputs to many of the remaining processes.

Preservation action plans are not transformed by the activities to which they are input. The actions of an activity are being triggered by the execution of the preservation action plan. Preservation action plans are a sequence of preservation actions or triggers.

When a Preservation Action Plan is executed it sends control (preservation) actions to other activities of the model. It might help to have a separate activity called "Execute Preservation Action Plan" that has Preservation Action Plans as an input and Preservation Actions as an output that triggered the rules and methods of preservation activities.

The reviewers concluded that having a preservation strategy as a control and preservation actions plans as inputs to subactivities of A2, A3, and A4 may be redundant. Rather the preservation strategy may only be needed as a control on A1.3.2 and as an input to A1.3.3.

The reviewers were unable to characterize a preservation strategy for a specific body of records apart from the Generic Preservation Framework and the preservation actions for a specific body of records. This leads us to suggest a change in terminology.

<table>
<thead>
<tr>
<th>Preservation Model Version 5.1</th>
<th>Suggested Terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preservation Framework</td>
<td>Preservation Strategy</td>
</tr>
<tr>
<td>Preservation Strategy</td>
<td>Preservation Plan</td>
</tr>
<tr>
<td>Preservation Action Plan</td>
<td>Preservation Action</td>
</tr>
</tbody>
</table>

Activity 1.3.1 is a very complex activity and is probably worthy of decomposition.

**A1.3.2 Plan for Implementing Preservation Strategy**

The reviewers identified data elements making up the Terms and Conditions for Transfer and some sample data from the case study.

**Outputs:**

(a) Terms and Conditions for Transfer:

1. Record Creator's Name: NY State Workers' Compensation Board
2. Transfer Agent's Name and Authority: John Doe, Records Manager
3. Identification of Records
   a) Identification Number (of Record Series)
   b) Title: Electronic Case Folder System
   c) Description: Series of case files for adjudicating benefits of disabled workers
   d) Document Types: Claims for Benefits, Employer's reports of accidents and illness, Correspondence, Medical Reports, Insurance Carrier's reports
   e) File or Data Structure: Relational Schema
   f) Types of files or selection criteria for including/excluding files in a transfer: (e.g., closed files, history files)
   g) Volume (number of files, number of media, size of files): 300,000 cases, one million files, nine million document images.
4. Expected transfer date or period: Position of the transfer in the records life cycle: (e.g., one year after closeout of a file)
5. Medium or means that will be used to effect the transfer: DLT Tape

6. Technical information that should accompany the transfer
   a) Metadata for the digital components should comprise: document type, digital component, the format, preservation history, and the originating technological environment.
   b) All digital components that are document images that are not in TIFF format should be converted to TIFF format.
   c) The schema describing the metadata for the records in the Electronic Case Folder System should be represented in SQL.
   d) Metadata for the records should comprise: the provenance (including a description of the business process procedures); the documentary context; the benchmark requirements for authenticity; information about how the records relate to the particular digital components that constitute the records; information about the different types of documents, the archival description of the records, and their arrangement.

7. Information supporting presumption of authenticity of records transferred by the records creator. This consists of evidence for each of the following requirements:

A.1.a Identity of the record
   A.1.a.i Name of author, Name of addressee: The ECFS data model permits the association of author's name, addressee, name of action or matter, and chronological date with each document.
   A.1.a.ii Name of action or matter.
   A.1.a.iii Chronological date.
   A.1.a.iv Expression of Archival Bond: When documents are imported by FileNet, a case file is ordered by document number.
   A.1.a.v Indication of attachments: Document preparation and mail transmittal preparation rules address how attachments are kept in the case folder.

A.1.b Integrity of the record
   A.1.b.i Name of Handling Office: NY WCB.
   A.1.b.ii Name of Office of Primary Responsibility: Same as above.
   A.1.b.iii Indications of types of annotations: FileNet supports annotations, but they are not used.
   A.1.b.iv Indication of technical modifications: Paper documents are scanned into document images in TIFF 6 format and maintained on WORM disks.

A.2 Access Privileges:
   Access to ECFS is controlled via passwords, job titles, workgroups, geographic location and business need.

A.3 Protective Procedures:
   Loss and Corruption of Records: There are back-up copies of the WORM disks and transaction logs.

A.4 Protective Procedures:
   Media and Technology: WORM disks are guaranteed for more than one hundred years.

A.5 Establishment of Documentary Forms:
   Each form is described in a procedure manual that is managed in Lotus Notes.

A.6 Authentication of Records:
   Authentication of document images in a case file is occasionally required in the adjudication process. They are presumed authentic because they are scanned images of paper documents and they are used in the normal course of business.
A.7 Identification of Authoritative Record:
The document images are the authoritative record unless the paper file is still available.

A.8 Removal and Transfer of Relevant Documentation:
There has not yet been a transition of active records to semi-active or inactive status, which involves a removal of records from the electronic system.

In lieu of a presumption of authenticity, a verification of authenticity should be provided.

b) Preservation Action Plans
The reviewers created some examples of preservation action plans for the specific body of records in the case study. The following are examples of preservation action plans that a creator might need to perform before transferring records to an archives.

1. When records are transferred from the record creator to the preserver, there must be an integrity check.
2. When data from a relational database are to be transferred, they will be transferred as flat ASCII files.
3. When a schema for a relational database is to be transferred, it should be represented in SQL.

The following is a possible Preservation Action Plan Associated with a Body of Records Selected for Preservation that can be used with Transferred and Accessioned Records.

Preservation Action Plan
1. Retrieve digital components for Claims for Benefits in Electronic Case Folder System that are ASCII text files.
2. Convert the ASCII Text files to TIFF multipage format using preservation method (TiffMaker).
3. Store the digital components converted to TIFF multipage format back to archival storage.
4. Store in the database the information that on this date the digital components for Claims for Benefits in the Electronic Case Folder System that were in ASCII text format have been converted to TIFF multipage format.

A1.3.3 Assess Strategy and Plan

Activity Definition:
Using Information about (the application of the Preservation Strategy and implementation of the Preservation Action Plan to) Transferred and Accessioned Records, determine whether and to what extent the Preservation Strategy and Preservation Action Plan(s) applicable to a body of electronic records selected for preservation have succeeded in satisfying Specified and Authenticity Requirements for Preservation. Use this assessment to produce an Updated Strategy and/or Updated Action Plans(s). If monitoring of electronic records selected for preservation prior to their transfer produces significant Updated Information about Electronic Records Selected for Preservation, revise the strategy or plan accordingly. Similarly, if applicable preservation methods have been changed, revise the strategy and/or plan to reflect the Updated Preservation Technology Specifications.

A1.4 Evaluate Execution of Preservation

Activity Definition:
“Review Management Information about Output of Electronic Records to determine if requirements for authentic copies are being satisfied and to characterize customer satisfaction.”
“Produce a Report on Authenticity of Records to respond to any challenge to the adequacy and efficacy of the preservation process.”

This is where the Preservation Framework (strategy) is demonstrated to meet the Authenticity Task Force’s Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records.8

**Baseline Requirements**
The preserver should be able to demonstrate that:

**Requirement B.1: Controls over Records Transfer, Maintenance, and Reproduction**
The procedures and systems(s) used to transfer records to the archival institution or program, maintain them, and reproduce them embody adequate and effective controls to guarantee the records' identity and integrity.

This requirement is satisfied by: (1) activity A1.3.2 for creating Terms and Conditions for Transfer; (2) activity A2.2, which compares the transfer with the Terms and Conditions for Transfer; (3) activity A2.2.3, which carries out Preservation Action Plans that use Preservation Methods to bring digital components into compliance with the preservation strategy; and (4) activity A4, which reproduces the record from maintained digital components.

and specifically that:

- **B.1.a Unbroken Custody of the records is maintained**
  This requirement is satisfied by the activities for Selection of Creator’s Records for Preservation, Terms and Conditions for Transfer, and Transfer and Preservation of Records.

- **B.1.b Security and control procedures are implemented and monitored**
  This requirement is satisfied in part by Access Control and Access Privileges of a DBMS (database management systems).

- **B.1.c The content of the record remains unchanged after reproduction.**
  This requirement is satisfied by Preservation Methods that preserve the content of electronic records.

**Requirement B.2: Documentation of [Update/] Reproduction Process and its Effects**
The activity of reproduction has been documented and that this documentation includes:

- **B.2.a The date of the records' reproduction and the name of the responsible person;**
- **B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver;**
- **B.2.c The impact of the reproduction process on their form, content, accessibility, and use;**
- and

  **B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user.**

  These requirements are satisfied by activity A1.2.3, Selecting a method to apply to a class of preservation objects, and by Preservation Action Plans that store a record of updates to digital components and by assessment of the effect of a Preservation Action Plan on the reproduction of form and content.

**Requirement B.3: Archival Description**
The archival description of the fonds containing the electronic records includes—in addition to information about the records’ juridical-administrative, provenancial, procedural, and documentary

---

8 Appendix 2.
contexts—information about changes the electronic records of the creator have undergone since they were first created.

This requirement is satisfied by: (1) requiring the transfer contain evidence supporting the presumption of authenticity of the transferred records, (2) maintaining records of refreshment of digital media and updates to digital components, and (3) activity A3.3, Update Digital Components, and specifically by preservation action plans that document updates to digital components.

Each of the Requirements for Supporting the Production of Authentic Copies of Electronic Records is satisfied by some set of activities of the Preservation Framework. This leads us to conclude that the Preservation Task Force's "Preserve Electronic Records" model provides a framework for preserving authentic electronic records. Within that framework, a variety of preservation strategies can be developed by archival institutions that are dependent on the characteristics of the selected, transferred and accessioned records; institutional requirements; and the current and changing state of information technology. The framework guides the development of strategies that can satisfy the Authenticity Task Force's baseline requirements.

**A2 Bring in Electronic Records**

**Activity Definition:**
"Following the direction established in the preservation strategy for a given body of records selected for preservation, the 'bring in' or ingest function applies preservation methods(s) targeted to that body of records to implement the preservation action plan for those records by processing each transfer of electronic records into accessioned electronic records. …"

**A2.2 Verify that the Transfer Is Authorized**

**Inputs:**
Registered Transfer

**Outputs:**
Rejected Transfer

**Activity Definition:**
Verification that a transfer is authorized is based on comparing the terms and conditions for transfer established as part of the Preservation Strategy during Appraisal, with the information accompanying the Registered Transfer.

The following is a more precise specification of the transform of inputs to outputs.

Compare the terms and conditions for transfer with the information accompanying the Registered Transfer to determine whether:

a) The transfer was sent by an authorized person,
b) It comprises the records specified for transfer,
c) It includes required information about the records and their digital components, and
d) It includes required information about the basis for asserting the authenticity of the records as received.

If the terms and conditions of transfer are satisfied, the conforming transfer is passed to the next step where its contents are examined. Otherwise, the transfer is rejected or the submitter is asked to address any problems identified.
A2.3 Examine Electronic Records

Input:
Preservation Action Plans\textsuperscript{xi} \textsuperscript{xii}

Conforming Transfer

2.3.1 Map Records and Digital Components within Transferred Material

Output:
Mapped Records and Digital Components\textsuperscript{xiii}

Rejected Transfer

Activity Definition:
In accordance with an institution's Accessioning Policy and the applicable Preservation Strategy, using the Technological Infrastructure in place, determine how the records, their digital components, and the information about them included in a Conforming Transfer are identified, and where they are located in the digital files and other materials received. Identify any records or components that should be in the transfer but are not found and determine if there are any records or components in the transfer that should not have been transferred. If critical deficiencies are found, terminate processing and output the Rejected Transfers. If no critical problems are encountered, output the Mapped Records and Digital Components, along with related information.

The following is a more precise specification of the transformation of inputs to outputs.

Examine the digital files and digital components of records in the transfer, along with accompanying information to:

a) determine how the records, their digital components, and the information about them in the conforming transfer are identified.

b) determine where they are located in the digital files and other materials received.

c) identify any records, aggregates of records or components that should be in the transfer but are not found.

d) identify any records, aggregates of records or components in the transfer that should not have been transferred.

Produce a mapping of information about electronic records transferred to the digital files that were transferred. Identify and locate the digital components included in the files(s), and link them to the records that they constitute.

A2.3.2 Verify that the Records in the Transfer can be Preserved and Reproduced.

Activity Definition:
Ensure that transferred records can be preserved and reproduced in accordance with the applicable preservation strategy. Determine if all of the records that should be in the transfer can be reconstituted and presented; if all digital components of these have been received and are in the formats stipulated in the terms and conditions of transfer; if archival aggregates established by the records creator can be re-established; if the archival bonds among records can be expressed or instantiated, and also what basis exists for asserting the authenticity of the records as transferred. Records that can be preserved and reproduced are output as Preservable Records. In the case of a record that cannot be preserved in accordance with the preservation strategy, identify the digital component(s) that prevent or impede implementation of the strategy and forward them for actions that will enable preservation of the record. Digital components modified as a result of such actions are returned for verification of the success of the modification. If it was necessary to modify the applicable preservation strategy to enable preservation of the records, the secondary verification will be on the basis of the revised strategy. The verification process will take into account any cases where digital components were not successfully
modified, along with other problems discovered with the transfer, to determine whether, under the institution's accessioning policy, the preservable records—including records deemed to have only minor problems—should be accessioned or the transfer should be rejected.

Input:
Mapped Records and Digital Components

The following is a more precise specification of the transformation of inputs to outputs.

The Mapped Records and Digital Components are used to examine the digital files and digital components of records in the transfer, along with accompanying information.

If:

a) all digital components are in the formats stipulated in the terms and conditions of transfer and there are methods for reconstituting and reproducing records from these digital components;
b) the archival aggregates established by the records creator can be re-established;
c) the archival bonds among records can be expressed or instantiated; and
d) there is adequate evidence for a presumption of authenticity of the records as transferred, then output the Preservable Records.

Otherwise, if actions are required to preserve some of the individual records transferred and the archival sets in which these records belong, then output these Digital Components of a Record that cannot be preserved.

When Conforming Digital Components are returned as a result of taking actions needed to preserve the Record, combine them with the rest of the preservable records and output the preservable records.

When Non-Conforming Digital Components are returned as a result of taking actions needed to preserve the Record, reject that portion of the transfer.

A2.3.3 Take Action Needed to Preserve the Record

When there are Digital Components of a Record that cannot be preserved, and the conditions of the Preservation Action Plan for this Transfer apply to these components, take the preservation actions and output the Conforming Digital components, or output the Non-Conforming Digital Components.

A3 Maintain Electronic Records

Activity Definition:
“Following direction established in the preservation strategy for a given body of records selected for preservation, apply preservation methods(s) targeted to that body of records to implement the preservation action plan for those records …”

A3.1 Manage Information about Records

The reviewers identified data elements for the first input to this activity.

Inputs:

1. Information about Accessioned Electronic Records
   a) Record Creator’s name
   b) Transfer Agent’s name
   c) Date of transfer
   d) Identification of records
      Title
      Description
      Volume
File or Date Structure
Technical Information (e.g., XML SQL, file formats, encoding)
e) Information supporting presumption of authenticity of records transferred by the records creator. This consists of evidence for each of the following requirements:

A.1.a Identity of the record
A.1.b Integrity of the record
A.2 Access Privileges
A.3 Protective Procedures: Loss and Corruption of Records
A.4 Protective Procedures: Media and Technology
A.5 Establishment of Documentary Forms
A.6 Authentication of Records
A.7 Identification of Authoritative Record
A.8 Removal and Transfer of Relevant Documentation

In lieu of a presumption of authenticity, a verification of authenticity should be provided.

f) Information about the media on which the records were transferred.

2. Basis of Authenticity of Transferred Records
3. Information about Updated Digital Components
4. Updated Storage Information: Location of files in Accessioned Electronic Records.
5. Retrieval Request

This activity requires the following functions:

1. The capability to store, retrieve, and update data.
2. A data definition language (DDL) for defining the conceptual schema (data items, records, primary and secondary keys, record relationships and integrity constraints), user views, and a data dictionary.
3. [Access Control and Access Privileges] Mechanisms for controlling access to data and for defining what actions may be taken by processes or users.
4. [Transaction integrity] The capability to define transaction boundaries, i.e., the logical beginning and end of transactions.
5. [Control of concurrent transactions] Safeguards to prevent erroneous results that can occur when two or more processes attempt to access a data record concurrently.
6. [Recovery Services] The capability to restore the database in the event of some system failure.

These functions can be provided with a DBMS.

A3.1.1 Maintain Information about Records
A3.1.2 Retrieve Information about a Requested Record
A3.1.3 Retrieve Information about Digital Components
A3.2 Manage Storage of Digital Components of Records
A3.2.1 Place Record Components in Storage
[File size constraint] When digital components are placed in archival storage, they should be stored as a file no larger than the capacity of the chosen media.

A3.2.2 Refresh Storage
When the prescribed date for updating a storage medium occurs, use the prescribed storage update method to transfer files to the new storage medium.
A3.2.3 Monitor Storage
When an environmental condition for archival storage (e.g., humidity or temperature) is out of range, signal the type of archival storage problem.

A3.2.4 Correct Storage Problems

A3.2.5 Retrieve Components from Storage

A3.3 Update Digital Components

A3.3.1 Migrate Digital Components to Current Formats
[Migration Strategy Rule] When format of digital component is obsolete, use a Targeted Preservation Method to migrate the digital component in obsolete format to the chosen current format, and associate with this group of digital components documentation of the migration of digital components to the new format.

A3.3.2 Convert Digital Components to Standardized Formats
[Conversion to Standard Format Rule] When format of digital component is proprietary, use a Targeted Preservation Method to convert the digital component in proprietary format to the chosen standard format.

A3.3.3 Transform Digital Components to Persistent Format
When the preservation action for the digital components of an archival set is to transform them to persistent format, apply the method for transforming the digital components to a software and hardware independent format, e.g., XSL-FO.

A4 Output Electronic Records
Activity Definition:
“Following direction established in the preservation strategy for a given body of records selected for preservation, apply preservation method(s) targeted to that body of records to implement the preservation action plan for producing an authentic copy of a record in response to a request for it. …”

A4.1 Manage the Request

A4.2 Review Retrieved Components and Information

A4.3 Reconstitute Record

A4.4 Package Output
When requested records are to be disseminated, wrap or encapsulate digital components of the electronic record, or the Reconstituted Electronic Record, with the associated metadata using the chosen Packaging Method.

Summary and Conclusions
A primary result of the walkthrough was to clarify the concepts of a Preservation Strategy, Terms and Conditions for Transfer, Preservation Action Plans, and Targeted Preservation Methods. Examples were also created for preservation plans and preservation methods. Another result was the demonstration that the preservation model provides a framework for satisfying the Authenticity Task Force’s Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records.
Conducting a walkthrough of the preservation model without having a data model was very time-consuming. However, we now have a clearer idea of the data elements making up the inputs and outputs to activities. During InterPARES 2, the data elements identified can be used in constructing a preservation data model.

Because the definitions of activities in the preservation model were descriptive, rather than defining transforms of inputs to outputs, the product (preservation model version 5.1) was not ready for a test of the model against case data. The walkthrough of the model amounted to a review of the model—raising issues and problems and discussing them, providing feedback on the model mixed with continued problem solving.

The case data were not adequate for testing the product. The case studies in the InterPARES Project were not designed with a walkthrough of the preservation model in mind. However, the attempt to walk through the model with case study data provided us with insight into the kinds of information that will need to be collected in order to perform a walkthrough. During InterPARES 2, a case study can be designed to collect the data that are needed to support a more complete walkthrough and validation of the preservation model.
Notes

i "Information about Transferred and Accessioned Records" should not include the term "accessioned" because the records are not yet accessioned. Also information to be included in "The Information about Transferred Records" is set out in the "Terms and Conditions for Transfer."

ii May not need requirements for preservation as a control.

iii Can "Targeted Preservation Method" that is an output of A1.2 and a control on A2, A3, and A4 be eliminated because it is an input to A1.3 (even though named "Preservation Technology Specifications") where it is incorporated into Preservation Strategy, which controls A2, A3, and A4? A strategy is a statement of direction: where you want to go, how you will get there, and how you will know you arrived. The method is software. You can't carry out the strategy without it, but it is distinct from the strategy. A method may be invoked by several strategies, any one of which may change. For reasons of economy and effectiveness, an archive will tend to limit the number of methods it supports. As long as one strategy still requires a method, the software remains in use.

iv That a transfer conforms to the Terms and Conditions of Transfer is checked in activity A2.2, Verify that the Transfer is Authorized. The output of that activity is Conforming Transfer, which has the definition "A transfer of electronic records that satisfies the terms and conditions stipulated for the transfer." If the transfer doesn't conform, it is rejected. It appears that activities A1.3.1 and A1.3.2 only need to be performed for Information about Electronic Records Selected for Preservation in order to produce a Preservation Strategy, Terms and Conditions for Transfer, and Preservation Action Plans. The Assessment activity (A1.3.3) could assess the information about transferred and accessioned records and update the preservation strategy and actions plans without having to re-perform activities A1.3.1 and A1.3.2. However, when we got around to figuring out how the preservation strategy and preservation actions plans would be used in A3.3, Update Digital Components, it was realized that some activity must be monitoring changes in the State of Information Technology (new computers, new operating systems, new formats, new standards) and this activity (ies) would need to revise strategies, acquire targeted preservation methods (viewers, format converters) and then some activity would need to trigger (or plan) retrieval of types of digital components needing preservation actions and trigger the preservation acts themselves. In other words, a walkthrough needs to show the preservation planning and execution process (sequence of activities for selected records, for transferred records, and for maintained records in the face of changing technology. It may be that the activities of A1.2 and A1.3 are involved in all three processes. The key question is: Which activity monitors changes in Information Technology? It will trigger new options, selections, acquisitions, and preservation actions.

v For instance, develop general preservation framework, develop preservation strategies for specific bodies of records, and then one to push the triggers for preservation actions.

vi Should there also be a feedback from A1.3.3 to A1.3.1? If so, to be consistent, A1.3.3 should have one output, assessment, which feeds back to both A1.3.1 and A1.3.2. The updates are then carried on the existing strategy and plan outputs.

vii There is an output of A1.4, Evaluation of Execution, which loops back to control A1.3. In converting the IDEF(0) diagrams to PDF, that feedback loop was superimposed on the Target Preservation Method that is output from A1.2. The overlap should be fixed. A1.4 produces as an output Evaluation of Execution that is fed back to A1.3, but it is not shown as a control on A1.3.3 where it would be used.

viii Preservation Action Plans in A2.2 are used to trigger planning actions for verification.

ix Activity A2.2 only addresses the information about the transferred records. Preservation Action Plans should not be input to A2.2, but only used in A2.3 where the objects of the transfer are examined.

x This implies that the Terms and Conditions for Transfer must be an input to A2.2, or that it is a part of the Preservation Strategy, which is a control on A2.2. It also implies that activity A1.3.2 should take much of the information this in the "Information about Records Selected for Preservation" and put it into the "Terms and Conditions for Transfer." It is suggested that the Preservation Strategy that is used as a control on this
activity, and most of the other activities in A2, A3, and A4, be just a (Generic) Preservation Strategy (or Preservation Framework) and that Terms and Conditions for Transfer be a specific input for each Transfer.

xi The reason that Preservation Action Plans are needed in A2.3 is that the creator, as part of the Terms and Conditions of Transfer, may not have been required to perform all preservation actions that are required by the preserver's Preservation Strategy for that particular body of records.

xii The arrows in A2.3.2 and A2.3.3 directed from the lower edge of the activity box to the bottom of the page represent links to activities included in other diagrams. Why is the arrow labeled A4, Output Records, directed out of A2.3.2 needed? Digital Components of Records that cannot be Preserved are forwarded to A2.3.3. Why is the arrow labelled A3.3, Update Digital Components, needed? A2.3.3 has all the resources that are needed to update the Digital Components, without going to A3.3. The resources consist of the Preservation Action Plan and Targeted Preservation Methods. This may entail making Targeted Preservation Methods a control on A2.3.3, unless Targeted Preservations Methods are part of the Preservation [Control] Strategy.

xiii In the definition of the output "Mapped Records and Digital Components," the clause "and determining whether the information about the records and the digital components in the transfer is appropriate and sufficient for preservation and reproduction of the records" should be deleted because that function is performed in the subsequent activity A2.3.2, Verify that the Records in the Transfer Can Be Preserved and Reproduced.

xiv It may be necessary to have the Conforming Transfer as an input, because the Mapped Records and Digital Components are just a map, not the records and digital components themselves. Furthermore, the definition of A2.3.2 states that the basis for asserting the authenticity of the records as transferred will be examined, and this requires the Conforming Transfer, not just the map.

xv The Basis of Authenticity of Transferred Records does not need to be a separate input to A3.1 because it is a part of the Information about Accessioned Electronic Records (see description of inputs for A1 in this document).

xvi Plan for Updating Digital Components is an input to A3.1.3 so that components can be retrieved for updating.

xvii Which activity triggers the activities in A3.3, Update Digital Components? Is it activity A1.3.2, Plan for Implementing Preservation Strategies? Instead of Preservation Action Plans being inputs, perhaps these are the Intervention Triggers that occurred in prior versions of the preservation model.
Glossary

action
The conscious exercise of will by an officer of the records creator or by an external person aimed to create, maintain, modify or extinguish situations.

active record
A record needed by its creator for the purpose of carrying out actions.

affix
To store on a medium.

annotation
An addition made to a record after it has been created. See also: execution annotation; business matter annotation; record management annotation.

archival bond
The relationship that links each record, incrementally, to the previous and subsequent ones and to all those which participate in the same activity. It is originary (i.e., it comes into existence when a record is made or received and set aside), necessary (i.e., it exists for every record), and determined (i.e., it is characterized by the purpose of the record).

archival date
The date on which a record is officially incorporated into the creator’s records.

archival fonds
The whole of the records of a creator.

archives
The whole of the records of a creator.

An agency or institution responsible for the preservation and communication of records selected for permanent preservation.

A place in which records selected for permanent preservation are kept.

attestation
The written validation of a record by those who took part in the issuing of it (author, writer, countersigner) and by witnesses to the action or to the signing of the record.

audit trail
A means of tracking all the interactions with records within an electronic system so that any access to the system can be documented as it occurs for the purpose of preventing unauthorized actions in relation to the records.

authentic record
A record that is what it purports to be and that is free from tampering or corruption.

authentication
A declaration of a record’s authenticity at a specific point in time by a juridical person entrusted with the authority to make such a declaration.

authentication certificate of trusted third party
An attestation issued by a trusted third party for the purpose of authenticating the ownership and characteristics of a public key. It appears in conjunction with the digital signature of the author of a record, and is itself digitally signed by the trusted third party.

authenticity
The quality of being authentic, or entitled to acceptance. As being authoritative or duly authorized, as being what it professes in origin or authorship, as being genuine.
Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records
The minimum conditions necessary to enable the preserver to attest to the authenticity of copies of electronic records.

Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records
The conditions that serve as a basis for the preserver’s assessment of the authenticity of electronic records.

business matter annotation
An addition made to a record in the course of handling the business matter in which the record participates, that reflect actions taken subsequent to the creation of the record for the purpose of handling the activity or the matter in which the record participates.

certificate of authenticity
A declaration by the preserver that one or more reproduced or reproducible electronic records is authentic.

chronological date
The date (and, possibly, the time of day) of a record, included in the record by its author, or by the electronic system on the author’s behalf, in the course of its compilation.

class
A subdivision of a classification scheme and the concrete body of records corresponding to it.

class code
The component of the classification code that uniquely identifies the class to which a record belongs.

classification code
A series of alphabetical, numerical, or alphanumeric symbols used to identify the record in its documentary context.

classification scheme
Diagram, table, or other representation categorizing the creator's records by hierarchical classes and according to a coding system expressed in alphabetical, numerical, or alphanumeric symbols.

classify records
To assign records to the pre-established classes in the classification scheme and, within each class, to the dossiers to which they belong (if applicable), by attaching to each record a classification code.

competence
A sphere of functional responsibility entrusted to an office.

competent person
The office that is given a competence and has, therefore, the authority and capacity to act within it.

completeness
The characteristic of a record that refers to the presence within it of all the elements required by the creator and the juridical system for it to be capable of generating consequences.

context
The framework in which the action in which the record participates takes place. The types of context include juridical-administrative context, provenancial context, procedural context, documentary context, and technological context.

copy
An entity that is generated from and resembles a record.
corroboration
The explicit mention of the means used to validate the record.

create records
To make and set aside or receive and set aside records.

creation medium
The physical carrier on which a record, made or received, is set aside, for further action or reference.

custody
The physical and legal control over the existence, authenticity, location, and accessibility of records.

data
The smallest meaningful facts.

description of action
A presentation of the ideal motivation (the preamble) and the concrete reason (the exposition) for an action, as well as the action or matter itself (the disposition).

destroy inactive records
To obliterate inactive records from existence without reasonable hope of recovery.

digital component
A digital object that is part of an electronic record, or of a reproduced electronic record, or that contains one or more electronic records, or reproduced electronic records, and that requires specific methods for preservation.

digital time-stamp issued by a trusted third party
An attestation by a trusted third party that a record was received at a certain point in time.

diplomatics
The study of the genesis, inner constitution, and transmission of archival documents, and of their relationship with the facts represented in them and with their creator.

document
Recorded information.

documentary context
The archival fonds to which a record belongs, and its internal structure.

documentary form
The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated. The two types of documentary form are extrinsic elements and intrinsic elements.

draft
A record made for purposes of correction.

draft number
The unique identifier assigned to sequential draft versions of the same record. The draft number is added to the record when it is stored.

electronic record
A record that is created (made or received and set aside) in electronic form.

electronic seal
A specific means of authenticating a record or ensuring that is opened only by the intended addressee. A type of electronic signature. An example of an electronic seal is a digital signature, which is an electronic signature based on public key cryptography.
**electronic signature**
A digital mark that has the function of a signature in, is attached to, or is logically associated with a record, and is used by a signatory to take responsibility for, or to give consent to, the content of the record.

**execution annotation**
An addition made to a record, after its creation, as part of the document execution phase of an administrative procedure.

**extrinsic elements**
An element of a record that constitutes its external appearance. The types of extrinsic elements include presentation features, electronic signatures, electronic seals, digital time-stamps issued by a trusted third party, and special signs.

**file**
An integrated aggregation of records, constituted according to the person, action, or matter to which it refers.

**file identifier**
The means used to identify uniquely the file in which the record belongs.

**folder**
A cover in which non-electronic records, belonging in the same dossier, are loosely kept, usually in chronological order. A dossier may be distributed across a number of folders.

**function**
All of the activities aimed to accomplish one purpose, considered abstractly.

**graphic**
A representation of an object or outline of a figure, plan, or sketch by means of lines; a representation of an object formed by drawing.

**image**
An artificial imitation or representation of the external form of any object, or an optical appearance or counterpart of an object, such as is produced by rays of light, refracted as through a lens, or falling on a surface after passing through a small aperture. A subset of image is moving images, which are visual images, with or without sound, that, when viewed, present the illusion of motion.

**inactive records**
Records which are no longer needed for ongoing agency business.

**indication of action**
The subject line(s) and/or the title at the top of a record.

**indication of attachment**
The mention of autonomous items that have been linked inextricably to the record before transmission (i.e., added during its execution) in order for it to accomplish its purpose.

**information**
An assemblage of data in a comprehensible form capable of communication.

**institutional requirement**
An external, legal, regulatory, societal, and/or cultural constraint imposed on the institution responsible for preserving records, and/or an internal policy or procedure for the preservation of electronic and/or non-electronic records.
**intrinsic elements**
The elements of a record that constitute its internal composition. The types of intrinsic elements include name of author, name of originator, chronological date, name of place of origin of record, name of addressee(s), name of receiver(s), indication of action (matter), name of writer, corroboration, attestation, and qualification of signature.

The elements of a record that convey the action in which the record participates and its immediate context.

**juridical-administrative context**
The legal and organizational system in which the creating body belongs.

**juridical person**
An entity having the capacity or the potential to act legally and constituted either by a succession or collection of physical persons or a collection of properties.

**medium**
A physical carrier.

**mode of transmission**
The method by which a record is communicated over space or time.

**name of addressee**
The name of the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

**name of author**
The name of the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

**name of creator**
The name of the physical or juridical person in whose archival fonds the record exists.

**name of handling office**
The office with the authority and capacity for treating an action or matter.

**name of originator**
The name of the person assigned the electronic address in which the record has been generated and/or sent.

**name of place of origin**
The name of the geographic place where the record was generated, included in the content of the record by the author or in the electronic system on the author’s behalf.

**name of receiver**
The name of the physical or juridical person(s) to whom the record is copied for information purposes.

**name of writer**
The name of the physical or juridical person having the authority and capacity to articulate the content of the record.

**needed source records**
A source record after conversion that, instead of being destroyed, is returned to storage as a semi-active record.

**original**
The first complete and effective record.

**overall presentation**
The record’s overall information configuration, i.e., the manner in which content is presented to the senses, using text, image or sound, either alone or in combination.
presentation features  
A set of perceivable features (graphic, aural, visual), generated by means of encoding and program instructions, and capable, when used individually or in combination, to present a message to our senses. Presentation features include overall presentation and specific presentation features.

preservation framework  
The whole of the principles, policies, and strategies for maintaining digital components and related information over time, and for reproducing the related authentic records and/or archival aggregations of records, that is produced by interpreting external controls and applying them to the records selected for preservation.

preservation strategy  
Coherent set of objectives and methods for maintaining digital components and related information over time, and for reproducing the related authentic records and/or archival aggregations.

preserver  
The juridical person whose primary responsibility is the long-term preservation of authentic records.

presumption of authenticity  
An inference as to the fact of a record’s authenticity that is drawn from known facts about the manner in which that record has been created and maintained.

priority of transmission  
Indication of the priority in which a record is to be transmitted.

procedural context  
The business procedure in the course of which the record is created.

procedure  
In general, the body of written and unwritten rules governing the conduct of a transaction, or the formal steps undertaken in carrying out a transaction. In particular, the legislative machinery set up to carry out a given transaction.

provenancial context  
The creating body, its mandate, structure, and functions.

qualification of signature  
The mention of the title, capacity, and/or address of the persons signing a record.

record  
A document made or received and set aside in the course of a practical activity.

record attribute  
A defining characteristic of a record or of a record element.

record element  
A constituent part of a record’s documentary form.

record identity  
The distinguishing character of a record, i.e., the attributes of a record that uniquely characterize it and distinguish it from other records.

record integrity  
The quality or state of being a complete and uncorrupted record.

record item identifier  
The component of the classification code that corresponds to the progressive number of the record within the dossier or, in the absence of dossiers, within the specific class.
**record management annotation**
An addition to a record made in the course of handling the record itself. It reflects actions taken subsequent to the creation of the record for the purpose of managing it as part of an agency's records.

**reproduce electronic record**
To process the digital components of an electronic record in such a way as to produce an authentic copy of that record.

**reproduction**
An assemblage of data in a comprehensible form capable of communication.

**sound**
An aural representation of words, music, or any other manifestation of sound.

**special sign**
A symbol, such as a digital watermark, or a logo or crest of an organization, that identifies one or more persons involved in the compilation, receipt, or execution of a record.

**specific presentation features**
The specific aspects of the record's formal presentation that are necessary for it to achieve the purpose for which it was created.

**storage medium**
The physical carrier on which a record is affixed, when it is different from the creation medium.

**technological context**
The characteristics of the technical components of an electronic computing system in which records are created.

**terms and conditions of transfer**
A document that identifies in archival and technological terms electronic records to be transferred, together with relevant documentation, and that identifies the medium and format of transfer, when the transfer will occur, and the parties to the transfer.

**text**
A word, number, or symbol.

**transmission, form of**
The form that the record has when it is received.

**validate**
To grant official sanction by marking.

**verification of authenticity**
The act or process of establishing a correspondence of known facts about the record itself and the various contexts in which it has been created and maintained with the proposed fact of the record's authenticity.

**version**
One of two or more forms in which a record is issued.