Digital Records Pathways: Topics in Digital Preservation

Module 1: Introduction – A Framework for Digital Preservation

InterPARES / ICA
DRAFT July 2012

Table of Contents

Tables and Figures

Acknowledgements

Many people have contributed to the creation of these modules. In particular, students in the PhD program at the University of British Columbia – Elizabeth Shaffer, Corinne Rogers, Donald Force, and Elaine Goh – have drafted the contents, based on the work of InterPARES 1 and 2, and case studies conducted in InterPARES 3. Acknowledgment should also be made of the many Graduate Research Assistants who conducted the case studies, and therefore supported the development of these modules, InterPARES Team Canada, the many international researchers involved with InterPARES, and of course the Director of InterPARES, Luciana Duranti. Finally, thanks go to all who reviewed and commented on these modules, with special mention of InterPARES researchers: John McDonald, Information Management Consultant (modules 1, 2, 7, and 8), Jim Suderman, Director, Information Access from the Toronto's City Clerk Office (module 3), Evelyn McLellan, Systems Archivist, Artefactual Systems Inc., and Paul Hebbard, Records Management Archivist, Simon Fraser University (module 6).

1 Preface

Digital Records Pathways: Topics in Digital Preservation is an educational initiative developed jointly by the International Council on Archives (ICA) and International Research on Permanent Authentic Records in Electronic Systems Project (InterPARES). It offers training to archivists and records professionals in the creation, management and preservation of authentic, reliable and usable digital records. The program assumes that the user has a solid grounding in basic concepts of records management and archival theory, and builds on that knowledge.

Consisting of eight independent modules, and supported by the ICA International Terminology Database, *Digital Records Pathways* addresses theoretical and practical knowledge needed to establish the framework, governance structure and systems required to manage and preserve digital records throughout the records' lifecycle. Each module addresses a specific topic of relevance to the management and preservation of digital records. Each module is intended to be capable of standing alone or studied in conjunction with other modules.

2 About the ICA and InterPARES

The ICA and InterPARES are committed to establishing educational materials for the continuing education of archivists and records managers, to build upon foundational knowledge, disseminate new findings, and to equip archivists and records professionals with the necessary specialized knowledge and competencies to manage and preserve digital records.

The International Council on Archives (ICA) (www.ica.org) is dedicated to the effective management of records and the preservation, care and use of the world's archival heritage through its representation of records and archive professionals across the globe. Archives are an incredible resource. They are the documentary by-product of human activity and as such are an irreplaceable witness to past events, underpinning democracy, the identity of individuals and communities, and human rights. But they are also fragile and vulnerable. The ICA strives to protect and ensure access to archives through advocacy, setting standards, professional development, and enabling dialogue between archivists, policy makers, creators and users of archives.

The ICA is a neutral, non-governmental organization, funded by its membership, which operates through the activities of that diverse membership. For over sixty years ICA has united archival institutions and practitioners across the globe to advocate for good archival management and the physical protection of recorded heritage, to produce reputable standards and best practices, and to encourage dialogue, exchange, and transmission of this knowledge and expertise across national borders. With approximately 1500 members in 195 countries and territories the Council's ethos is to harness the cultural diversity of its membership to deliver effective solutions and a flexible, imaginative profession.

The International Research on Permanent Authentic Records in Electronic Systems (InterPARES) (www.interpares.org) aims to develop the knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form and provide the basis for standards, policies, strategies and plans of action capable of ensuring the longevity of such material and the ability of its users to trust its authenticity. InterPARES has developed in three phases:

InterPARES 1 (1999-2001) focused on the development of theory and methods ensuring the preservation of the authenticity of records created and/or maintained in databases and document management systems in the course of administrative activities. Its findings present the perspective of the records preserver.

InterPARES 2 (2002-2007) continued to research issues of authenticity, and examined the issues of reliability and accuracy during the entire lifecycle of records, from creation to permanent preservation. It focused on records produced in dynamic and interactive digital environments in the course of artistic, scientific and governmental activities.

InterPARES 3 (2007-2012) built upon the findings of InterPARES 1 and 2, as well as other digital preservation projects worldwide. It put theory into practice, working with archives and archival / records units within organizations of limited financial and / or human resources to implement sound records management and preservation programs.

3 Audience

The audience for this program includes archivists and records and information professionals interested in expanding their competencies in the management of digital records. Taken as a whole, the modules form a suite of resource materials for continuing professional education with particular focus on issues influencing the preservation of reliable, accurate and authentic digital records.

3.1 Assessment Tools – Individual Readiness

In order to gain the full benefit from these modules, you should have a basic working knowledge of archival and records management theories, principles and practices, including core concepts such as selection and appraisal, arrangement and description, retention and disposition. Many excellent resources exist that provide this fundamental knowledge, and it is not the intention of these modules to reproduce that knowledge base. Two such resources, available on line at no cost, are produced by *The International Records Management Trust* (IRMT). These training and education programs, published ten years apart, provide comprehensive coverage of archival and records management issues with a particular focus on public sector records. The current modules are intended to complement the IRMT training materials.

1. The IRMT's *Management of Public Sector Records Study Programme* (1999) is a comprehensive set of training resources for archivists and records managers providing basic theoretical knowledge. The modules stress the importance of good record keeping, particularly within the public sector, and discuss the need to manage information as a

strategic resource. They present a rationale for developing an integrated records management program that restructures existing information and records systems and outlines the key activities undertaken in records and archives management. The information presented in this and the other modules can be used in government, corporate, organizational or personal settings. The principles apply equally whether the agency is public or private.

2. *Training in Electronic Records Management* or *TERM* (2009) was developed by the IRMT as part of a wider project to investigate issues associated with establishing integrity in public sector information systems. The focus of the study was payroll and personnel records, since payroll control and procurement are the two major areas of government expenditure most vulnerable to misappropriation, and payroll control is, therefore, a highly significant issue for all governments. Providing route maps for moving from a paper-based system to an electronic environment, the project examined the degree to which the controls and authorizations that operated in paper-based systems in the past have been translated into the electronic working environment.

Both are available at:
http://www.irmt.org/educationTrainMaterials.php

3.2 Assessment Tools – Institutional Readiness

As well as assessing individual readiness, you are encouraged to determine the level of sophistication of your organization's recordkeeping strategies. Many excellent maturity models for measuring recordkeeping capacity are available. These will help you to assess your organization's ability to preserve its digital records. This in turn will help you determine which of the modules in this series will be the most beneficial for you and your organization.

The records management models listed below reflect practice from the UK, Australia, and the United States. They all offer tools to help you determine where your organization sits on a scale of readiness or maturity. Five levels are identified, from *ad hoc*, or no formal records management, to "Optimized" or "Transformational", indicating that the organization has "integrated information governance into its overall corporate infrastructure and business processes to such an extent that compliance with the program requirements is routine."[1]

Records management maturity models:

[1] ARMA (2010) GARP Maturity Model.

- JISC InfoNet (2009). Records and Information Management – Maturity Model. Available at http://www.jiscinfonet.ac.uk/records-management/measuring-impact/maturity-model/index_html.
- ECM3 (2010). ECM Maturity Model, v. 2. Available at http://ecmmaturity.files.wordpress.com/2009/02/ecm3-v2_0.pdf.
- Queensland State Archives (2010). Recordkeeping maturity model and road map: Improving recordkeeping in Queensland public authorities. Available at http://www.archives.qld.gov.au/downloads/maturity_model_road_map.pdf.
- ARMA International (2010). GARP: Generally Accepted Recordkeeping Principles – Information Governance Maturity Model. Available at http://www.arma.org/garp/metrics.cfm.

4 How to Use *Digital Records Pathways*

Each module in *Digital Records Pathways* consists of theoretical and methodological knowledge and its practical application, illustrated through case studies and model scenarios. While the modules have been developed by InterPARES Team Canada, and so are illustrated with examples from the Canadian context, each module is customizable to a specific domain or juridical context. For wider applicability, they will be translated into the languages of the ICA partners.

The modules can be studied as a set, or individually according to need and interest, covering the range of competencies required. They can be self-administered by individuals, or offered through professional associations or workplace training. Several of the modules also contain templates that may be adapted by universities and professional associations to develop specific course curricula, or on-site training materials for students and professionals on digital recordkeeping and preservation issues. Universities and professional associations are free to adapt the materials and develop their own context-specific course curricula and training kits.

Throughout the modules you will find links to additional resources, indicated by a "help" graphic:

And links to more information within the modules themselves, indicated by a "file" graphic:

5 Objectives

The modules have the following objectives:

- To provide educational resources based on current research in digital records issues to professional archival and records management associations for the benefit of their members;
- To provide archivists and records professionals with the necessary theoretical knowledge and procedural and strategic skills to develop, implement and monitor a digital recordkeeping and/or preservation system;
- To illuminate theoretical concepts with practical applications through real life examples drawn from case studies, anchored in specific administrative and technological contexts;
- To provide university programs with content and structure for courses on digital records management and preservation.

5.1 Architecture of the set

<table>
<tr><td colspan="4">Architecture of the set</td></tr>
<tr><td colspan="3">1. A Framework for Digital Preservation
2. Developing Policy and Procedures for Digital Preservation</td><td>Foundational</td></tr>
<tr><td>3. Organizational Culture</td><td>4. An Overview of Metadata</td><td>5. Appraisal Strategies</td><td>General purpose</td></tr>
<tr><td>6. E-mail</td><td>7. Websites</td><td>8. Cloud Computing</td><td>Specific purpose</td></tr>
<tr><td colspan="3">International Terminology Database</td><td>Foundational</td></tr>
</table>

The first two modules offer information fundamental to any program of digital records preservation. They provide a foundation upon which the subsequent modules are built. The next three modules offer general information on topics common to digital preservation – the role of organizational culture, an overview of metadata, and an overview of appraisal in the context of managing records outside an electronic recordkeeping system (ERMS). The final three modules address specific topics of concern – the management of e-mail, preservation of records in web environments, and the issues arising from the increasing reliance on cloud computing.

Each module consists of some or all of the following components:

- **Overview** of the topic and scope of the module;
- **Learning objectives** and expected level of knowledge upon completion;
- **Methodology** or the procedures to follow to apply the module;
- **Templates (where appropriate)** to facilitate the implementation of the module;

- **Examples and Case Study(ies)/Scenarios (where appropriate)** that provide real-world examples of module topic[2]
- **Exercises** covering key learning points
- **Review questions** to enhance comprehension and understanding of the topic
- **Additional Resources** for each module
    - Readings, standards and other templates for reference

Where appropriate, distinctions are drawn between the management and preservation activities involving active records and responsibilities for records that are no longer required for business purposes, whether they are preserved by their creator or by a trusted third party.

6 Scope

Digital Records Pathways: Topics in Digital Preservation consists of the following eight modules:

6.1 Module 1: Introduction – A Framework for Digital Preservation

This module introduces the set of modules and explains how to use them, outlines objectives, and summarizes the contents of each module. It includes resources for institutional readiness and self-assessment tools to assist individuals and organizations in assessing their readiness or capacity for digital preservation. It introduces two complementary models for digital preservation: the InterPARES Chain of Preservation model, and the Open Archival Information System Reference Model (OAIS), and an annotated bibliography of digital preservation research and learning resources.

6.2 Module 2: Developing Policy and Procedures for Digital Preservation

This module explains the purpose and benefits, and provides the knowledge and tools necessary to create a digital preservation policy. It guides the user in developing, writing and implementing an effective digital preservation policy within his/her organization and includes methodology for the development of policies, practical tools to aid in policy development, examples of existing policies and further resources to assist users in policy and procedure development.

[2] The examples and case studies cited in the modules are taken from real case studies in InterPARES 3. They are intended to support the learning experience of the modules. While they reflect the research findings of InterPARES, they are not necessarily intended to be viewed as templates of best practice applicable in all cases. Every organization (creator or preserver) is different and preservation of their records must embrace best practice from a pragmatic perspective of the feasibility of implementation.

6.3 Module 3: Organizational Culture and its Effects on Records Management

Records professions often overlook issues of organizational culture when developing and implementing a recordkeeping system. The objective of this module is to highlight issues arising from organizational culture that act as enabling or constraining factors in the adoption of a recordkeeping system. The module outlines the different types of organizational culture and their salient characteristics. The module provides tools to enable stakeholders within an organization to identify the type of culture of a business unit as well as to assess the overall culture of their organization based on an organizational culture assessment checklist and an accompanying list of indicators. The module includes strategies on how to promote records management within each type of culture. It helps users understand the factors that shape an organization's culture and how these factors facilitate or hinder people's adoption of a recordkeeping and preservation system.

6.4 Module 4: An Overview of Metadata

Metadata is integral to digital records management and preservation. This module provides an overview of the roles of metadata in digital records management and preservation. It outlines the different kinds of metadata, dependent on functional requirements – descriptive metadata for identification and access; administrative metadata, including technical, rights, and preservation metadata; and structural metadata, that documents the structural relationships between or within digital resources. The module functions as a metadata primer, and a compendium of the more common metadata standards currently in use. It also presents the InterPARES General Study on an application profile for authenticity metadata.

6.5 Module 5: From *Ad Hoc* to Governed – Appraisal Strategies for Gaining Control of Digital Records in Network Drives

Appraisal consists of four distinct activities: compiling information; assessing value; determining feasibility of preservation; and making the appraisal decision. Assessment of authenticity in the context of assessing value is an integral part of records' appraisal. Appraisal must rest on a foundation of solid research, which will be of particular assistance in assessing record value and authenticity, and identifying digital components that must be preserved. The purpose of this module is to introduce the user to the recommended process of appraisal as guided by the Chain of Preservation Model (InterPARES 2) and measure the records' authenticity against the Benchmark Requirements supporting the presumption of authenticity (InterPARES 1). The module provides appraisal guidelines to provide for an analysis of legacy files to establish authenticity, (data leading to the presumption of authenticity, or if there is an insufficient basis for a presumption of authenticity, the verification of authenticity), and provides a template for conducting the appraisal and documenting appraisal decisions.

The second part of this module outlines a methodology and action plan for an organization to move from a record-creating environment where unstructured records and

documents are stored and maintained in network drives, to a controlled record-creating and keeping environment such as an ERMS or EDRMS. This module walks individuals through a process of evaluating their organization's record-creating and recordkeeping environment; identifying and appraising their organization's electronic records; and preparing their organization's electronic records for migration to an ERMS or EDRMS (or other structured and secure records management system).

6.6 Module 6: E-mail Management and Preservation

This module is designed to help organizations gain better control of their e-mail. The module walks individuals through a series of steps that will facilitate the implementation of new policies, procedures, and practices with regard to how staff manage and preserve their messages. The module reviews the e-mail management and preservation model (EMPM), a multi-phase process for implementing e-mail management and preservation policies and procedures. This model discusses the various factors that influence e-mail management and preservation, different e-mail management methods, ways to apply retention and disposition to e-mail, ways to preserve e-mail, and the design and implementation of e-mail policies and procedures.

6.7 Module 7: Management and Preservation of Records in Web Environments

This module introduces key issues involved in the management and preservation of records in web environments. Organizational websites may contain a mix of records, some of which require long-term preservation, and non-record materials, in increasingly complex forms. Websites that comprise static documents and incorporate little or no interactivity are relatively simple to deal with. However, sites that incorporate high levels of interactivity and comprise dynamically generated pages are complex and difficult to preserve effectively. This module helps readers identify records that exist on their organization's website, and analyze the management and preservation needs of these records. It identifies a workflow management process for managing the creation and movement of records to and from websites into preservation environments, and situates the process within the policy framework of the organization.

6.8 Module 8: Cloud Computing Primer

Still an emerging concept, cloud computing is on-demand computing services delivered over the Internet from a remote location or via an organization's servers. This module provides a primer on cloud computing and the associated records management issues and challenges that should be considered before an organization moves some or all of its records, services and/or processes to the cloud. It defines the characteristics of cloud computing and explains its three service and four deployment models, outlines a methodology and identifies tools for analyzing the risks when migrating information and processes to cloud. The module lays the groundwork to aid users in developing a cloud computing strategy: including educating users on identifying the issues relevant to the use of cloud computing when selecting processes, applications and records to be moved to the cloud; and business requirements, rules and compliance frameworks that must be examined in light of the issues cloud computing raises.

6.9 International Terminology Database

The terminology used in the modules is common to archival and records management communities of practice. To ensure common understanding, and minimize potential confusion that may arise from regional or jurisdictional practice, the modules are supported by a database of archival and records management terms that reflect common usage and practice in 16 languages. This database, developed jointly by the ICA and InterPARES, is available at www.web-denizen.com/. This dynamic resource will continue to grow and develop as members of the archival community worldwide have an opportunity to participate in adding and commenting on definitions as used in their region of practice. Certain specific terms not yet included in the database will be found in short glossaries in each module.

7 Key Concepts and Models

Digital records are documents created with digital technologies (most commonly with the use of a computer) that are treated and managed as records (documents made or received in the course of a practical activity as an instrument or by-product of such activity, and set aside for action or reference.) Although we tend to think of digital records in the same way that we think of analogue records, they are different in several fundamental respects. Digital records are comprised of physical and intellectual components that are no longer inextricably linked as they are, for example, in paper records. A digital record may be stored on different media but still exist as a record, capable of achieving the purpose for which it was initially created. Digital records, then, are comprised of digital components and the relationships that link them.

It is not possible to preserve the digital record itself, but only the ability to recreate the record. Digital preservation is the whole of the principles, policies, rules and strategies designed to ensure that digital objects remain accessible, intelligible, and usable over time and across technological change, and that their reliability and accuracy is protected and their authenticity is verifiable. A digital records preservation strategy sets out objectives and methods for protecting and maintaining digital components and related information of records over time, and for reproducing authentic records and/or archival aggregations. This is accomplished through a set of rules governing the intellectual and physical maintenance of records, and the tools and mechanisms used to implement these rules.[3]

While the stability of analogue records has meant that preservation decisions can be postponed to the end of a record's lifecycle or beyond, digital records by their nature are fragile and volatile. Preservation concerns must be addressed from the moment of record creation. Failure to do so may result in records whose reliability, accuracy and authenticity cannot be guaranteed over time.

---

[3] Luciana Duranti and Randy Preston, eds. (2008) "Glossary," International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records. Padova: Associatione Nazionale Archivistica Italiana.

The ability to preserve trustworthy digital records (i.e., records that can be demonstrated to be reliable, accurate and authentic), depends on records being created in such a way that it is possible to maintain and preserve them. This requires that a relationship between a records creator and its designated preserver must begin at the time the records are created. However, record creation in the digital environment is rarely guided by considerations of preservation over the long term. As a result, the reliability, accuracy and authenticity of digital records can either not be established in the first place or not be demonstrated over time. Such records cannot support the creator's accountability requirements, nor can they be effectively relied upon either by the creator for reference or later action or by external users as evidence of actions and transactions or sources for research and memory. Furthermore, they cannot be understood within an historical context, thereby undermining the traditional role of preserving organizations such as public archival institutions.

The majority of records and information-related legislation at any level of government in any jurisdiction has not recognized the unique requirements of digital records. In some cases, legislation has established significant barriers to the effective preservation of digital records over the long term, most notably that regarding copyright. Despite this, the InterPARES 2 researchers concluded that it is possible to develop a framework of principles to support record creation, maintenance and preservation, regardless of jurisdiction. These principles guide this and all subsequent modules in this set.

7.1 Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records[4]

The framework of principles for preservation of digital records can be applied in any juridical context, and is independent of the technology used to create or store the records. Furthermore, the framework establishes the relationship between the records creator and the designated preserver. The records preserver and records creator may be one and the same, or different agencies within the same organization, or at arms length of one another – the preserver operating under mandate or legislation as a trusted third party.

The principles are summarized in the following tables. Table 1 lists the principles for records creators, ordered in priority from highest priority (C1) downward, and linked to the corresponding principle for records preservers. Table 2 lists the principles for records preservers, ordered in priority from highest (P1), and linked to the corresponding principle for records creators.

See Appendix A for the complete Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records

[4] Ibid., Appendix 19 – reproduced in Appendix A below.

Creator Principle	Text	Preserver Principle
C1	Digital objects must have stable content and fixed documentary form	P5
C2	Records creation procedures should ensure that digital components can be separately maintained and reassembled over time	P4
C3	Records creation and maintenance requirements should reflect the purposes the records fulfill, rather than available or chosen technologies	P6
C4	Records creation and maintenance policies, strategies and standards should address the issues of record reliability, accuracy, and authenticity expressly and separately	P2
C5	A trusted record-making system should be used to create records that can be presumed reliable	
C6	A trusted recordkeeping system should be used to maintain records that can be presumed accurate and authentic	P11, P12
C7	Preservation considerations should be embedded in records creation and maintenance activities	P7
C8	A trusted custodian should be designated as preserver of a creator's records	P1
C9	All business functions and processes that contribute to the creation and/or use of record aggregations should be explicitly documented	P10
C10	Third party intellectual property rights should be explicitly identified and managed in the record-making and record-keeping systems	P8
C11	Privacy rights and obligations attached to the creator's records should be explicitly identified and protected in the record-making and record-keeping systems	P9
C12	Sharing records across jurisdictions should be based on the legal requirements under which the records are created	P13
C13	Reproductions made in the usual and ordinary course of business are considered records of the creator	P3

Table 1: Principles for Records Creators

Preserver	Text	Creator

Principle		Principle
P1	A designated preserver fulfills the roll of trusted custodian	C8
P2	Records preservation policies, strategies and standards should address the issues of record reliability, accuracy, and authenticity expressly and separately	C4
P3	Reproductions of a creator's records made for the purpose of preservation by the preserver are considered authentic copies	C13
P4	Records preservation procedures should ensure that digital components can be separately maintained and reassembled over time	C2
P5	Authentic copies should be made for preservation purposes only from the creator's records (that is, digital objects with stable content and fixed documentary form)	C1
P6	Preservation requirements should be articulated in terms of the purpose/desired outcome of preservation, not in terms of available technologies	C3
P7	Preservation considerations should be embedded in all phases of the records lifecycle to ensure their continuing authenticity	C7
P8	Third party intellectual property rights should be explicitly identified and managed in the preservation system	C10
P9	Privacy rights and obligations attached to the creator's records should be explicitly identified and protected in the preservation system	C11
P10	Archival appraisal should identify and analyze all business processes that contribute to the creation and/or use of the same records	C9
P11	Archival appraisal should assess the authenticity of records	C6
P12	Archival description should be used as a collective authentication of the records in an archival fonds	C6
P13	Procedures for providing access to records created in one jurisdiction to users in another jurisdiction should be based on the legal requirements under which the records are created	C12

Table 2: Priniciples for Records Preservers

7.2 The Chain of Preservation (COP) Model[5]

Preservation activities begin with an awareness of what is required to create records that are reliable and accurate and maintained authentic over time. This can be modeled over the lifecycle of records to represent the actions that will ensure records remain accessible and readable over time, and that their form, content and relationships are protected for as long as required.

InterPARES 2 developed the Chain of Preservation (COP) Model to depict and document all the phases or stages in the lifecycle of digital records, and all the activities that must be undertaken to ensure that digital records are created reliable and accurate, and maintained authentic. The model is relevant for records creators or records preservers, reflecting the understanding that the long-term preservation of authentic digital records constitutes actions undertaken throughout the records' lifecycle. All activities pertaining to preservation are interdependent – omission of actions and activities at any stage may imperil the reliability and authenticity of records over time and across technological change.

The COP model proceeds from an understanding of the concepts, methods and practices that together comprise archival science. The model balances constraints on record making, record keeping, and record preserving, mechanisms that are available to creators and preservers for record making and recordkeeping, inputs into and outputs from the system so controlled. (Figure 1)

Constraints on the model then include the laws and regulations of the juridical system in which the records are created, and the international, national or other standards to which the records must adhere. Technological constraints limit, and the preserver's mission and capacity also influence, records creation and preservation. Mechanisms reflect the many resources needed to create, maintain, control, and preserve digital records. These resources include facilities and infrastructure, technology, personnel, and financial resources. Inputs include digital objects created or received and information about the digital objects, the creator, preserver, available technology, and requests for access to the digital objects. Outputs are documents and records that result from the activities throughout the chain of preservation, starting from creation.

[5]Ibid.

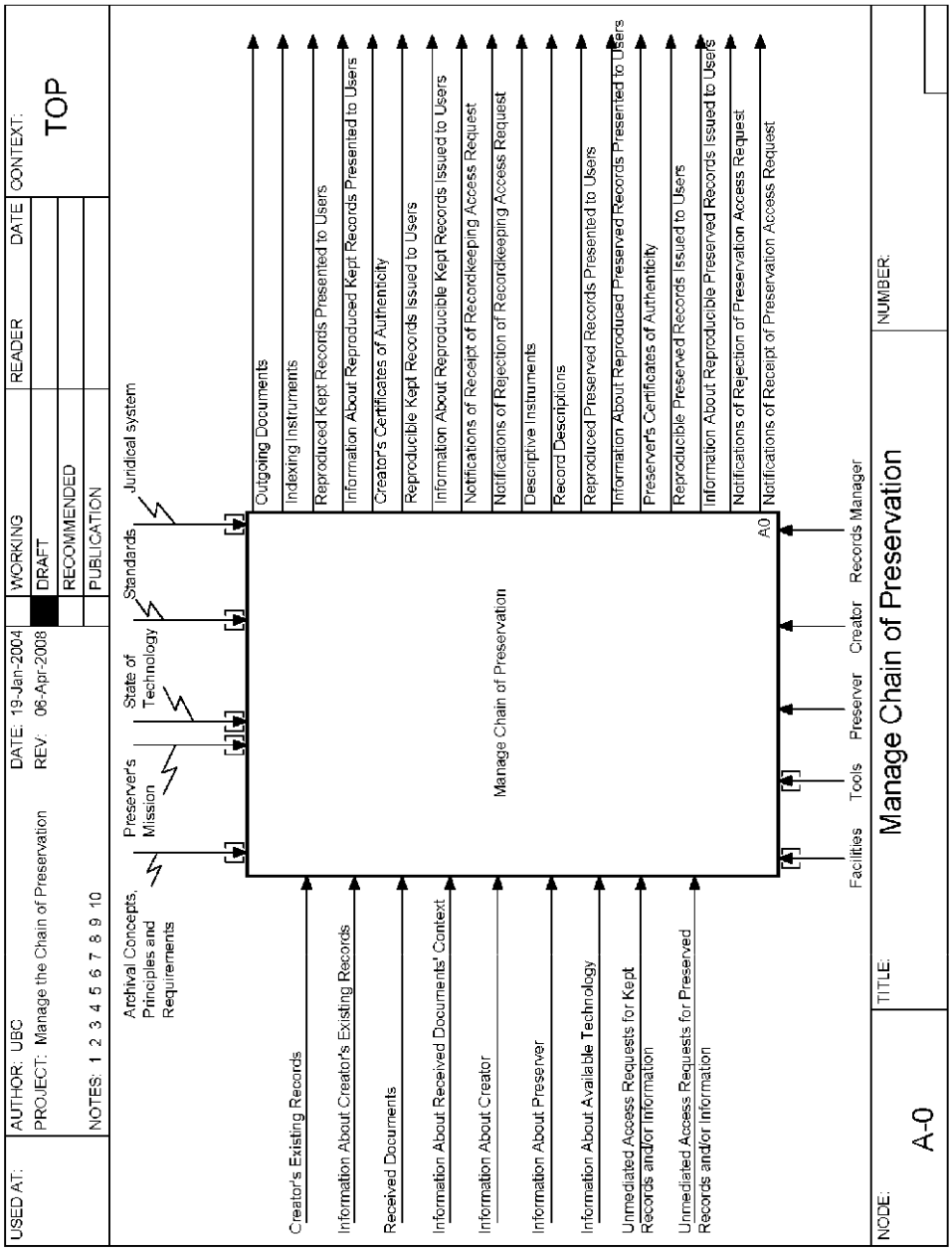

Figure 3: Manage Chain of Preservation A-0

Four high-level activities are delineated and further decomposed: (1) managing the framework for chain of preservation, (2) managing records creation, (3) managing records in a recordkeeping system, and (4) preserving selected records. (Figure 2)

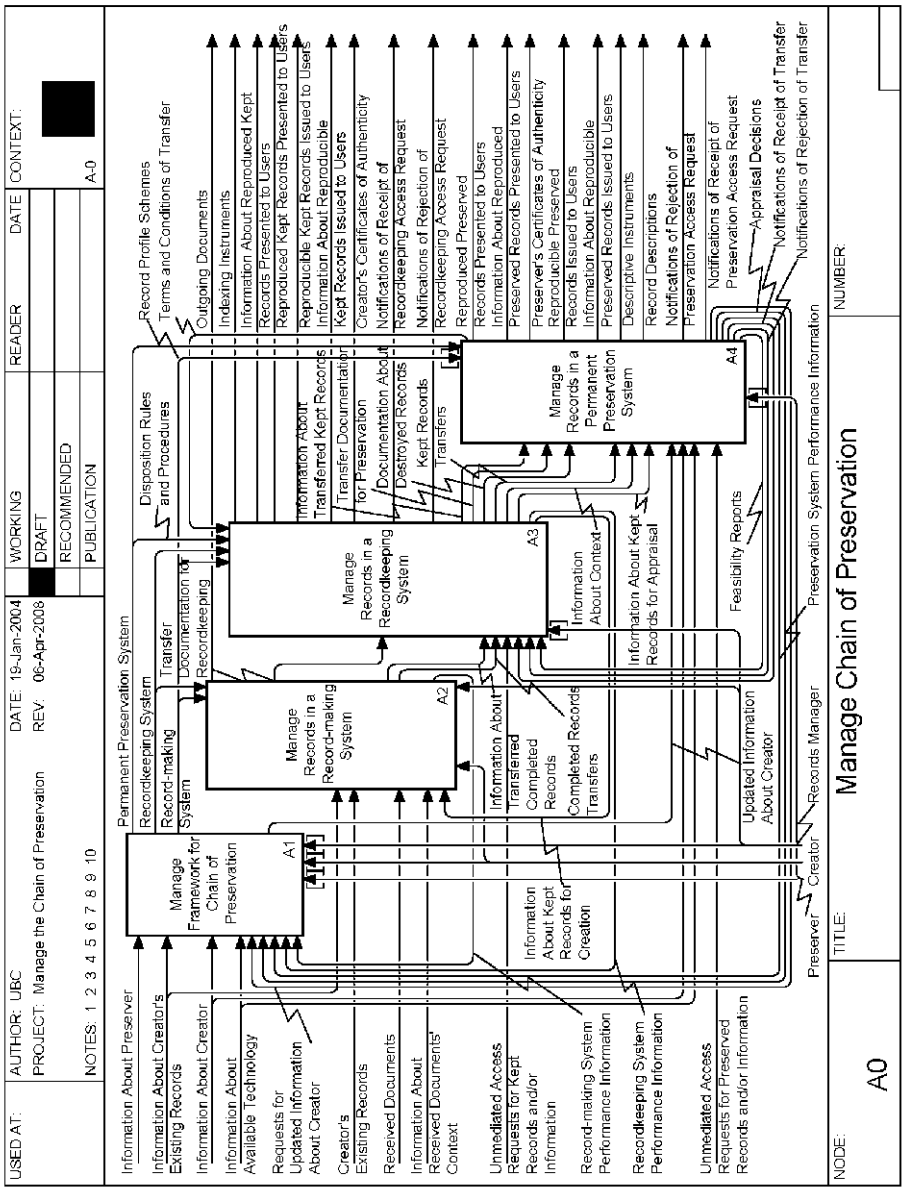

Figure 4: Manage Chain of Preservation A0

Go to http://www.interpares.org/ip2/book.cfm for full details on the Chain of Preservation Model

7.3 The Open Archival Information System (OAIS) Reference Model[6]

The Open Archival Information System (OAIS) Reference Model, an approved ISO standard and considered the benchmark for digital preservation systems, is a high-level model that defines the base functional components of a long-term preservation system and the key internal and external interfaces, and characterizes the information objects managed in the system. It addresses all aspects of long-term preservation of digital information: ingest, archival storage, data management, access, dissemination, and migration to new media and forms. The goal of an OAIS is to preserve information for a designated community over an indefinite period of time (CCSDS 2002). Digital preservation initiatives have adopted, adapted, or referenced the OAIS model since its inception as the foundation upon which to build, as has the preservation section of the COP model.

The OAIS Reference Model does not dictate means of implementation, but prescribes requirements to ensure that an OAIS-compliant repository is "an organization of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community." An OAIS archive, therefore, is situated in the context of a user community and answerable to that community. The Reference Model describes the external environment, the functional components or internal mechanisms that collectively fulfill the preservation responsibilities, and the information objects that are ingested, managed, and disseminated by the OAIS repository.

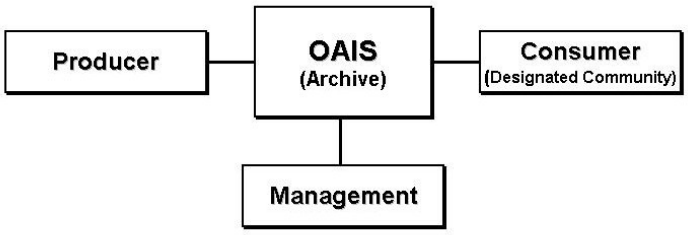

Figure 5: The OAIS Environment

The OAIS environment situates the repository in the context of the Producer – the individual, organization or system that sends content to the repository, and the Consumer – the end user of content from the repository.

---

[6] See Brian Lavoie (2004) *The Open Archival Information System Reference Model: Introductory Guide*. Dublin, OH: OCLC Online Computer Library Center, Inc., January. http://www.dpconline.org/docs/lavoie_OAIS.pdf

The Functional Model outlines six core repository functions: preservation planning, data management, administration, ingest, archival storage, and access.

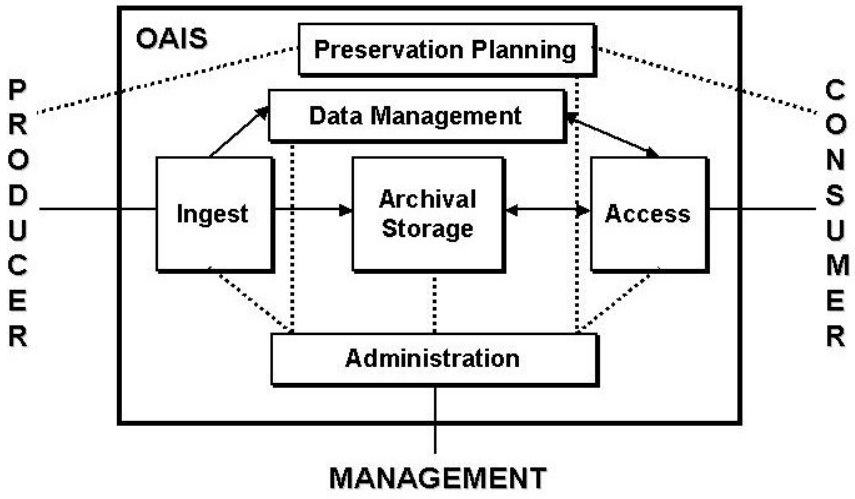

Figure 6: OAIS Functional Model

The Information Model is a high-level description of the information objects managed by the archive. An information object is made up of a data object – the bit stream of a digital object – and its representation information that allows the data object to be rendered as meaningful information. An information package is a conceptualization of the structure of information and consists of an object to be preserved and the metadata necessary for its long-term preservation and access, bound into a single logical information package. Three types of information objects exist – the submission information package (SIP), the archival information package (AIP) and the dissemination information package (DIP).

The Archival Information Package (AIP) is the focus of preservation activities. It must contain the complete set of metadata necessary to support long-term preservation and access. Four types of information make up the AIP: Content Information, Preservation Description Information, Packaging Information, and Descriptive Information. The first two contain metadata directly relevant to preservation. (Descriptive Information, while certainly required by the repository, is considered more relevant to resource discovery than to the preservation function itself.)

Metadata necessary for understanding the object is bound with the Content Data Object as Representation Information, giving meaning, in the case of digital objects, to the string of bits that comprise the object. Together they are referred to as the Content Information.

Metadata necessary to carry out preservation functions is found in the Preservation Description Information. It can be categorized in four types:

- Reference information – uniquely identifies the Content Information,
- Provenance information – details the history, documents any alterations to content and form over time, and chain of custody information,

- Context information – documents the relationship(s) of the Content Information to other Content Information objects,
- Fixity information – validates the authenticity or integrity of the Content Information (CCSDS 2002; OCLC/RLG Working Group on Preservation Metadata 2002; Lavoie 2004)

The OAIS Reference Model offers a useful starting point for the development of preservation metadata in a broad categorization of types. It is notable for its independence from data object type – not only can the model accommodate any digital object, but can also be applied to physical objects – and its independence from any technology or preservation strategy. Its very strengths in this regard are also its limitation – the high-level nature of its structure and concepts make implementation challenging.

Go to http://public.ccsds.org/publications/archive/650x0b1.pdf for full details on the OAIS Reference Model.

7.4 Metadata

Metadata affords intellectual and logical control over the data to which they relate. Preservation metadata is a subset of all the metadata associated with an object, supporting the functions of maintaining fixity, viability, renderability, understandability, and/or authenticity in the preservation context. It crosses the lines between structural and administrative metadata, as well as some rights management metadata.

Preservation metadata has been defined as information that supports and documents the process of digital preservation (Caplan 2006; Lavoie and Gartner 2005), or "information a repository uses to support the digital preservation process" (PREMIS Editorial Committee 2011). A preservation metadata framework is "an overview or description of the types of information (i.e. metadata) that should be associated with an archived object" (OCLC/RLG Working Group on Preservation Metadata 2001). A high-level framework must be comprehensive, covering all aspects of the preservation function, structured, and broadly applicable. It is supported by evidence of its development and creation, and information about the work processes that maintain it (see for example, the PREMIS Data Dictionary for Preservation Metadata.)

Go to the PREMIS Data Dictionary for Preservation Metadata, and the Preservation Metadata Maintenance Activity at http://www.loc.gov/standards/premis/

"Detailed trustworthy metadata are key to ensuring the creation of reliable, and preservation of authentic, records and other entities in electronic systems."[7] The InterPARES 3 project has proposed six functional requirements for metadata that ensure the authenticity of digital records. These functional requirements must support: 1) presumption of authenticity, 2) interoperability between systems and across time, 3) parsimony, 4) adequacy for archival description, 5) retrieval, and 6) meaningful display. This can be expressed as a sentence as follows: *these metadata should be necessary and sufficient to support the presumption of authenticity of records, interoperate between systems and across time, be adequate for archival description, and be useful for both retrieval and meaningful display of records.*

see Module 4: An Overview of Metadata for more information about metadata.

The creation of reliable, accurate records, and the protection of their authenticity and trustworthiness over time and across technological change can be guided by the instruments introduced here – the Framework of Principles, the Chain of Preservation Model, and, for trusted digital repositories, the OAIS Reference Model. The digital objects whose management is guided by these models are supported by "an end-to-end metadata management regime that addresses which metadata need to be created and/or carried forward in time."[8]

See Appendix B for an annotated bibliography of digital preservation programs

[7] InterPARES 2 Book, p. 340.
[8] Ibid.

8 Resources

Author: International Records Management Trust
Editors: General Editor, Laura Millar
Title: Understanding the Context of Electronic Records Management
Publication Date: 2009
Publisher: International Records Management Trust

Training in Electronic Records Management or *TERM*, was developed by the International Records Management Trust as part of a wider project to investigate issues associated with establishing integrity in public sector information systems. The focus of the study was pay and personnel records, since payroll control and procurement are the two major areas of government expenditure most vulnerable to misappropriation, and payroll control is, therefore, a highly significant issue for all governments. Providing route maps for moving from a paper-based system to an electronic environment, the project examined the degree to which the controls and authorizations that operated in paper-based systems in the past have been translated into the electronic working environment.

Author: International Records Management Trust
Editors: General Editor, Michael Roper; Managing Editor, Laura Millar
Title: Management of Public Sector Records Study Programme
Publication Date: 1999
Publisher: International Records Management Trust

Management of Public Sector Records Study Programme is a comprehensive set of training resources for archivists and records managers providing basic theoretical knowledge. The modules examine and stress the importance of good record keeping, particularly within the public sector, and discuss the need to manage information as a strategic resource. They present a rationale for developing an integrated records management program that restructures existing information and records systems and outlines the key activities undertaken in records and archives management. The information presented in this and the other modules can be used in government, corporate, organizational or personal settings; the principles apply equally whether the agency is public or private.

Author: Elizabeth Shepherd and Geoffrey Yeo
Title: *Managing Records: A Handbook of Principles and Practice*
Publication Date: 2003
Publisher: London: Facet Publishing

Managing Records: A Handbook of Principles and Practice is an essential text that "describes and discusses the principles of records management and its practical

implementation in contemporary organizations."[9] Intended for both experienced practitioners and newcomers to the field of records management, the book does not require any prior knowledge of the subject. The book includes a comprehensive bibliography of records management resources as well as lists of national and international records management standards and professional organizations for records managers.

Author: CCSDS
Title: Reference Model for an Open Archival Information System (OAIS)
Publication Date: 2002
Publisher: Consultative Committee for Space Data Systems
http://public.ccsds.org/publications/archive/650x0b1.pdf

InterPARES Creator and Preserver Guidelines; Benchmark Requirements Supporting the Presumption of Authenticity of Electronic records; and Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

[9] Elizabeth Shepherd and Geoffrey Yeo (2003), "Preface," *Managing Records: A Handbook of Principles and Practice*, London: Facet Publishing.

9 References

Caplan, Priscilla. 2006. Preservation Metadata. In *DCC Digital Curation Manual*, ed. Seamus Ross and Michael Day. July. http://www.dcc.ac.uk/resource/curation-manual/chapters/preservation-metadata.

———. 2009. *Understanding PREMIS*. Library of Congress, February. http://www.loc.gov/standards/premis/understanding-premis.pdf.

CCSDS. 2002. Reference Model for an Open Archival Information System (OAIS). Consultative Committee for Space Data Systems. http://public.ccsds.org/publications/archive/650x0b1.pdf.

Duranti, Luciana, and Randy Preston. 2008. Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential. Interactive and Dynamic Records. Padova: Associazione Nazionale Archivistica Italiana.

Lavoie, Brian. 2004. *The Open Archival Information System Reference Model: Introductory Guide*. Dublin, OH: OCLC Online Computer Library Center, Inc., January. http://www.dpconline.org/docs/lavoie_OAIS.pdf.

PREMIS Editorial Committee. 2011. *PREMIS Data Dictionary for Preservation metadata*. January. http://www.loc.gov/standards/premis/v2/premis-2-1.pdf.

APPENDIX A: A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records[10]

Introduction

The InterPARES research projects have examined the creation, maintenance and preservation of digital records. A major finding of the research is that, to preserve trustworthy digital records (i.e., records that can be demonstrated to be reliable, accurate and authentic), records creators must create them in such a way that it is possible to maintain and preserve them. This entails that a relationship between a records creator[11] and its designated preserver[12] must begin at the time the records are created.[13]

The InterPARES 1 research (1999-2001) was undertaken from the viewpoint of the preserver. Three central findings emerged from it: 1) there are several requirements that should be in place in any recordkeeping environment aiming to create reliable and accurate digital records and to maintain authentic records;[14] 2) it is not possible to preserve digital records but only the ability to reproduce them;[15] and 3) the preserver

---

[10] The term initially used in the InterPARES Project is "electronic records." In fact, the book resulting from InterPARES 1 is named *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (Luciana Duranti, ed.; San Miniato, Archilab, 2005), and the formal title of InterPARES 2 carries that terminology forward. However, in the course of the research, the term "electronic record" began to be gradually replaced by the term "digital record," which has a less generic meaning, and by the end of the research cycle, the research team had developed separate definitions for the two terms and decided to use the latter as the one that better describes the object of InterPARES research. The definition for "electronic record" reads: "An analogue or digital record that is carried by an electrical conductor and requires the use of electronic equipment to be intelligible by a person." The definition for "digital record" is, effectively, a digitally-encoded object and the metadata necessary to order, structure or manifest the object's content and form, where "digital object" is taken to mean "a discrete aggregation of one or more bit streams and the metadata about the properties of the object and, if applicable, methods of performing operations on the object." See the InterPARES 2 Terminology Database, available at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

[11] Records creator is the physical or juridical person (i.e., a collection or succession of physical persons, such as an organization, a committee, or a position) who makes or receives and sets aside the records for action or reference. As such, the term includes all officers who work for a juridical person, such as records managers, records keepers and preservers.

[12] Records preserver is a generic term that refers more to the function than to the professional designation of the physical or juridical person in question. Thus, the preserver might be a unit in an organization, a stand-alone institution, an archivist or anyone else who has as primary responsibility the long-term preservation of records.

[13] Records are created when they are made or received and set aside or saved for action or reference.

[14] See Authenticity Task Force (2002). "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204–219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.

[15] See Kenneth Thibodeau et al., "Part Three – Trusting to Time: Preserving Authentic Records in the Long Term: Preservation Task Force Report," ibid, 99–116. Online reprint available at http://www.interpares.org/book/interpares_book_f_part3.pdf. InterPARES 2 Project Book: Appendix 19 L. Duranti, J. Suderman and M. Todd InterPARES 2 Project, Policy Cross-domain Task Force Page 2 of 22

needs to be involved with the records from the beginning of their lifecycle to be able to assert that the copies that will be selected for permanent preservation are indeed authentic copies of the creator's records.

The InterPARES 2 research (2002-2006) took the records creator's perspective. The researchers carried out case studies of records creation and maintenance in the artistic, scientific and governmental sectors; they modeled the many functions that make up records creation and maintenance and records preservation according to both the lifecycle and the continuum models; they reviewed and compared legislation and government policies from a number of different countries and at different levels of government, from the national to the municipal; they analyzed many metadata initiatives and developed a tool to identify the strengths and weaknesses of existing metadata schemas in relation to questions of reliability, accuracy and authenticity; and, once again, they studied the concept of trustworthiness and its components, reliability, accuracy.

The case studies showed that record creation in the digital environment is almost never guided by considerations of preservation over the long term. As a result, the reliability, accuracy and authenticity of digital records can either not be established in the first place or not be demonstrated over periods of time relevant to the "business"[16] requirements for the records. These records cannot therefore support the creator's accountability requirements, nor can they be effectively relied upon either by the creator for reference or later action or by external users as sources. Furthermore, they cannot be understood within an historical context, thereby undermining the traditional role of preserving organizations such as public archival institutions.

The research undertaken in records and information-related legislation showed that no level of government in any country to date has taken a comprehensive view of the records lifecycle, and that, in some cases, legislation has established significant barriers to the effective preservation of digital records over the long term, most notably that regarding copyright.

It was the responsibility of the InterPARES 2 Policy Cross-domain research team (hereinafter "the Policy team") to determine whether it was possible to establish a framework of principles that could guide the creation of policies, strategies and standards, and that would be flexible enough to be useful in differing national environments, and consistent enough to be adopted in its entirety as a solid basis for any such document. In particular, such a framework had to balance different cultural, social and juridical perspectives on the issues of access to information, data privacy and intellectual property.

The findings of the InterPARES 1 research were confirmed by the research conducted by the InterPARES 2 Policy team, which further concluded that it is possible to develop such a framework of principles to support record creation, maintenance and preservation,

---

and authenticity and how it is understood, not just in the traditional legal and administrative environments, but in the arts, in the sciences and in the developing areas of e-government.

[16] The term "business" is used in its most general sense, since the object of the InterPARES research includes works of art and scientific data as well as standard types of business records.

regardless of jurisdiction. This document, in combination with other products of the Project, especially the Chain of Preservation (COP) model,[17] reflects this conclusion, while emphasizing the need to make explicit the nature of the relationship between records creators and preservers.

The Policy team developed two complementary sets of principles, one for records creators and one for records preservers, which are intended to support the establishment of the relationship between creators and preservers by demonstrating the nature of that relationship.[18] The principles for records creators are directed to the persons responsible for developing policies and strategies for the creation, maintenance and use of digital records within any kind of organization, and to national and international standards bodies. The principles for records preservers are directed to the persons responsible for developing policies and strategies for the long-term preservation of digital records within administrative units or institutions that have as their core mandate the preservation of the bodies of records created by persons, administrative units or organizations external to them, selected for permanent preservation under their jurisdiction for reasons of legal, administrative or historical accountability. They are therefore intended for administrative units (e.g., a bank, a city or a university archives) or institutions (e.g., a community archives or a state archives) with effective knowledge of records and records preservation.

Structure of the Principles

The principles are similarly presented, with the principle statement followed by an explanatory narrative, sometimes with illustrative examples. The principles are more often phrased as recommendations ("should") rather than imperatives ("must"), because some of them might not be relevant to some records creators or preservers. Each principle statement is followed by an indication of the corresponding principle in the other set (C stands for Creator, P stands for Preserver; the number is the principle number in the C or the P set). The reason why the principle numbers do not correspond in the two sets (C1=P1) is that the principles are listed in each set in order of relative importance.

Principles for Records Creators

(C1) Digital objects must have a stable content and a fixed documentary form to be considered records and to be capable of being preserved over time. (P5)

The InterPARES Project has defined a record as "a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference,"[19] adopting the traditional archival definition. This definition implies that, to be considered as a record, a digital object generated by the creator must first be a document; that is, must have stable content and fixed documentary

---

[17] The COP model is available in Appendix 14 and at http://www.interpares.org/ip2/ip2_models.cfm. A narrative discussion of the model is provided in the Modeling Cross-domain Task Force Report.
[18] The initial draft of the principles relied heavily on the contributions of three research assistants: Fiorella Foscarini, Emily O'Neill and Sherry Xie.
[19] See InterPARES 2 Terminology Database, op. cit.

form. Only digital objects possessing both are capable of serving the record's memorial function.

The concept of *stable content* is self-explanatory, as it simply refers to the fact that the data and the information in the record (i.e., the message the record is intended to convey) are unchanged and unchangeable. This implies that data or information cannot be overwritten, altered, deleted or added to. Thus, if one has a system that contains fluid, ever-changing data or information, one has no records in such a system until one decides to make one and to save it with its unalterable content.

The concept of *fixed form* is more complex. A digital object has a fixed form when its binary content is stored so that the message it conveys can be rendered with the same documentary presentation it had on the screen when first saved. Because the same documentary presentation of a record can be produced by a variety of digital formats or presentations,[20] fixed form does not imply that the bit streams must remain intact over time. It is possible to change the way a record is contained in a computer file without changing the record; for example, if a digital object generated in '.doc' format is later saved in '.pdf' format, the way it manifests itself on the screen—its documentary presentation, or "documentary form"—has not changed, so one can say that the object has a fixed form.

One can also produce digital information that can take several different documentary forms. This means that the same content can be presented on the screen in several different ways, the various types of graphs available in spreadsheet software being one example. In this case, each presentation of such a digital object in the limited series of possibilities allowed by the system is to be considered as a different view of the same record having stable content and fixed form.

In addition, one has to consider the concept of "bounded variability," which refers to changes to the form and/or content of a digital record that are limited and controlled by fixed rules, so that the same query, request or interaction always generates the same result.[21] In such cases, variations in the record's form and content are either caused by technology, such as different operating systems or applications used to access the document, or by the intention of the author or writer of the document. Where content is concerned, the same query will always return the same subset, while, as mentioned, its presentation might vary within an allowed range, such as image magnification. In

---

[20]Digital format is defined as "The byte-serialized encoding of a digital object that defines the syntactic and semantic rules for the mapping from an information model to a byte stream and the inverse mapping from that byte stream back to the original information model" (InterPARES 2 Terminology Database, op. cit.). In most contexts, digital format is used interchangeably with digital file-related concepts such as file format, file wrapper, file encoding, etc. However, there are some contexts, "such as the network transport of formatted content streams or consideration of content streams at a level of granularity finer than that of an entire file, where specific reference to "file" is inappropriate" (Stephen L. Abrams (2005), "Establishing a Global Digital Format Registry," *Library Trends* 54(1): 126. Available at http://muse.jhu.edu/demo/library_trends/v054/54.1abrams.pdf).

[21] See Duranti and Thibodeau, "The Concept of Record," op. cit.

consideration of the fact that what causes these variations also limits them, they are not considered to be violations of the requirements of stable content and fixed form.

Organizations should establish criteria for determining which digital objects need to be maintained as records and what methods should be employed to fix their form and content if they are fluid when generated. The criteria should be based on business needs but should respect as well the requirements of legal, administrative and historical accountability.

(C2) Record creation procedures should ensure that digital components of records can be separately maintained and reassembled over time. (P4)

Every digital record is composed of one or more digital components. A digital component is a digital object that is part of one or more digital records, including any metadata necessary to order, structure or manifest content, and that requires a given preservation action. For example, an e-mail that includes a picture and a digital signature will have at least four digital components (the header, the text, the picture and the digital signature). Reports with attachments in different formats will consist of more than one digital component, whereas a report with its attachments saved in one PDF file will consist of only one digital component. Although digital components are each stored separately, each digital component exists in a specific relationship to the other digital components that make up the record.

Preservation of digital records requires that all the digital components of a record be consistently identified, linked and stored in a way that they can be retrieved and reconstituted into a record having the same documentary presentation it manifested when last closed. Each digital component requires one or more specific methods for decoding the bit stream and for presenting it for use over time. The bit stream can be altered, as a result of conversion for example, as long as it continues to be able to fulfill its original role in the reproduction of the record. All digital components must be able to work together after they are altered; therefore, all changes need to be assessed by the creator for the effects they may have on the record.

Organizations should establish policies and procedures that stipulate the identification of digital components at the creation stage and that ensure they can be maintained, transmitted, reproduced, upgraded and reassembled over time.

(C3) Record creation and maintenance requirements should be formulated in terms of the purposes the records are to fulfill, rather than in terms of the available or chosen record-making or recordkeeping technologies. (P6)

Digital records rely, by definition, on computer technology and any instance of a record exists within a specific technological environment. For this reason, it may seem useful to establish record creation and maintenance requirements in terms of the technological characteristics of the records or the technological applications in which the records may reside. However, not only do technologies change, sometimes very frequently, but they are also governed by proprietary considerations established and modified at will by their developers. Both these factors can significantly affect the accessibility of records over

time. For these reasons, references to specific technologies should not be included in records policies, strategies and standards governing the creation and maintenance of an organization's records. Only the business requirements and obligations that the records are designed to support should be explicitly kept in consideration at such a high regulatory level. At the level of implementation, the characteristics of specific technologies should be taken into account to support the established business requirement and make possible its realization.

Technological solutions to record creation and maintenance are dynamic, meaning that they will evolve as the technology evolves. New technologies will enable new ways of creating records that meet an organization's business requirements. The rapid adoption of Web technologies to support business communication and transaction illustrates this. Specific activities for maintaining records will therefore require continuing adaptation to new situations drawing on expertise from a number of disciplines. To extend the example of the use of Web technologies, organizations creating and maintaining transactional records in a mainframe environment need to draw on knowledge of the new Web technologies from both connectivity (i.e., how to connect the mainframe to the Web) and security standpoints (i.e., how to protect the records from remote, Web-based attacks). As new technologies are used to create records, reference to new archival knowledge will continue to be required.

Technological solutions need to be specific to be effective. Although the general theory and methodology of digital preservation applies to all digital records, the maintenance solutions for different types of records require different methods. Therefore, they should be based on the specific juridical-administrative context in which the records are created and maintained, the mandate, mission or goals of their creator, the functions and activities in which the records participate and the technologies employed in their creation to ensure the best solutions are adopted for their maintenance.

Record policies that are expressed in terms of business requirements rather than technologies will need to be periodically updated as the organization's business requirements change, rather than as the technology changes. It is the role of a specific action plan to identify appropriate technological solutions for the maintenance of specific aggregations of records. The identified solutions must be monitored with regard to the possible need for modifying and updating. This requires the records creating body to be aware of new research developments in the archival and records management fields and to collaborate with interdisciplinary efforts to develop appropriate methods for the management of digital records.

(C4) Record creation and maintenance policies, strategies and standards should address the issues of record reliability, accuracy and authenticity expressly and separately. (P2)

In the management of digital records, reliability, accuracy and authenticity are three vital considerations for any organization that wishes to sustain its business competitiveness and to comply with legislative and regulatory requirements. These considerations should be directly and separately addressed in records policies and promulgated throughout the organization. The concept of reliability refers to the authority and trustworthiness of a

record as a representation of the fact(s) it is about; that is, to its ability to stand for what it speaks of. In other words, reliability is the trustworthiness of a record's content. It can be inferred from two things: the degree of completeness of a record's documentary form and the degree of control exercised over the procedure (or workflow) in the course of which the record is generated. Reliability is then exclusively linked to a record's authorship and is the sole responsibility of the individual or organization that makes the record. Because, by definition, the content of a reliable record is trustworthy, and trustworthy content is, in turn, predicated on accurate data, it follows that a reliable record is also an accurate record.

An accurate record is one that contains correct, precise and exact data. Accuracy of a record may also indicate the absoluteness of the data it reports or its perfect or exclusive pertinence to the matter in question. The accuracy of a record is assumed when the record is created and used in the course of business processes to carry out business functions, based on the assumption that inaccurate records harm business interests. However, when records are transmitted across systems, refreshed, converted or migrated for continuous use, or the technology in which the record resides is upgraded, the data contained in the record must be verified to ensure their accuracy was not harmed by technical or human errors occurring in the transmission or transformation processes. The accuracy of the data must also be verified when records are created by importing data from other records systems. This verification of accuracy is the responsibility of the physical or juridical person receiving the data; however, such person is not responsible for the correctness of the data value, for which the sending person is accountable. Thus, the receiving person should issue a disclaimer regarding accuracy of records using other persons' data.

The concept of authenticity refers to the fact that a record is what it purports to be and has not been tampered with or otherwise corrupted. In other words, authenticity is the trustworthiness of a record as a record. An authentic record is as reliable and accurate as it was when first generated. Authenticity depends upon the record's transmission and the manner of its maintenance and custody. Authenticity is maintained and verifiable by maintaining the identity and integrity of a record. The identity of a record is established and maintained by indicating at a minimum the names of the persons participating in the creation of the record (e.g., author, addressee); the action or matter to which the record pertains; the date(s) of compilation, filing or transmission; the record's documentary form; the record's digital presentation (or format); the relationship of the record to other records through a classification code or a naming convention; and the existence of attachments. The integrity of a record is established and maintained by identifying the responsibility for the record through time by naming the handling person or office(s)[22] and the trusted records officer[23] or the recordkeeping office,[24] identifying access

---

[22] Handling office (or person) is defined as "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" (InterPARES 2 Terminology Database, op. cit.).

[23] A trusted records officer (also called records keeper or records manager) is defined as "an individual or a unit within the creating organization who is responsible for keeping and managing the creator's records, who has no reason to alter the kept records or allow others to alter them and who is capable of implementing all of the benchmark requirements for authentic records" (Ibid.).

privileges[25] and access restrictions[26] and indicating any annotations or any modifications (technical or otherwise) made to the record by the persons having access to it.

Thus, record reliability is a quality that is established when a record is created and implies accuracy of the data contained in the record, while record accuracy and authenticity are qualities that are connected with the transmission and maintenance of the record. The latter are therefore the responsibility of both the records creator and any legitimate successor. Authenticity is protected and guaranteed through the adoption of methods that ensure the record is not manipulated, altered, or otherwise falsified after its creation, either during its transmission or in the course of its handling and preservation, within the recordkeeping system.[27]

(C5) A trusted record-making system should be used to generate records that can be presumed reliable.[28]

A trusted record-making system consists of a set of rules governing the making of records and a set of tools and mechanisms used to implement these rules. To generate reliable records, every record-making system should include in its design integrated business and documentary procedures, record metadata schemes, records forms, record-making access privileges and record-making technological requirements.

Integrated business and documentary procedures are business procedures linked to documentation procedures and to the classification system (i.e., the file management plan or taxonomy) established in the organization. This integration reinforces the control over record-making procedures: it supports the reliability of records by explicitly connecting records to the activities in which they participate and to the records organization system, thereby standardizing the procedures for creating and managing those records. The integration of business and documentary procedures also establishes the basis and central means to demonstrate ownership of and responsibility for the records. A record-making metadata scheme is a list of all metadata elements that need to be documented in the course of record-making processes for the purposes of uniquely identifying each record and enabling the maintenance of its integrity and the presumption of its authenticity. Such a scheme can also be used later to verify authenticity when questioned. Records forms are specifications of the documentary forms for the various types of records generated in the record-making system. Access privileges refer to the authority to compile, edit, annotate, read, retrieve, transfer and/or destroy records in the record-making system, granted to officers and employees by the records creator on the basis of position duties and business needs. Access privileges control access to the record-making system and are established

---

[24] Recordkeeping office is defined as "The office given the formal competence for designing, implementing and maintaining the creator's trusted recordkeeping system" (Ibid.).
[25] Access privileges is defined as "The authority to access a system to compile, classify, register, retrieve, annotate, read, transfer or destroy records, granted to a person, position or office within an organization or agency" (Ibid.).
[26] Access restrictions is defined as "The authority to read a record, granted to a person, position or office within an organization or agency" (Ibid.).
[27] See MacNeil et al., "Authenticity Task Force Report," op. cit.
[28] There is no corresponding Preserver Principle.

in the course of integrating business and documentary procedures through connecting specific classes of records to the office of primary responsibility for a business function or activity. The establishment and implementation of access privileges is the most important step towards ensuring that the reliability of records can be presumed. Record-making technological requirements include the hardware and software specifications for the record-making system that have a direct impact on the documentary form of records.

(C6) A trusted recordkeeping system should be used to maintain records that can be presumed accurate and authentic. (P11, P12)

A trusted recordkeeping system consists of a set of rules governing the keeping of records and a set of tools and mechanisms used to implement these rules. Every recordkeeping system should include in its design a recordkeeping metadata scheme, a classification scheme, a retention schedule, a registration system, a recordkeeping retrieval system, recordkeeping technological requirements, recordkeeping access privileges and procedures for maintaining accurate and authentic records.

A recordkeeping metadata scheme is the list of all necessary metadata to be attached to each record to ensure its continuing identity and integrity in the recordkeeping system. A classification scheme is a plan for the systematic identification and arrangement of business activities and related records into categories according to logically structured conventions, methods and procedural rules. A retention schedule is a document specifying and authorizing the disposition of aggregations of records as identified in the classification scheme. A registration system is a method for assigning a unique identifier to each created record, linked to its identity and integrity metadata. Recordkeeping access privileges refer to the authority to classify, annotate, read, retrieve, transfer and/or destroy records in the recordkeeping system, granted to officers and employees by the records creator based on position duties and business needs. Typically, access to records for purposes of classification, transfer and destruction is given only to the trusted records officer of the organization. A recordkeeping retrieval system is a set of rules governing the searching and finding of records and/or information about records in a recordkeeping system and the tools and mechanisms used to implement these rules. Recordkeeping technological requirements include the hardware and software specifications for the recordkeeping system. The procedures for maintaining accurate and authentic records are the procedures designed to ensure that the data in the records and the identity and integrity of the records in the recordkeeping system are protected from accidental or malicious corruption or loss.

To improve efficiency and reduce the potential for human-induced error, the record-making and recordkeeping systems should be designed to automate, as much as possible, the creation of the identity and integrity metadata both at the point of records creation or modification (e.g., when migrated to a new system or file format), and whenever the aggregations to which the records belong are created or modified—every record unit should automatically inherit the metadata of the higher level in the classification at the point of creation as well as whenever there are updates to the metadata of the higher level.

A records creator should indicate in its records management policy that it is the trusted records officer's responsibility to manage the recordkeeping system. The role of the trusted records officer is analogous to that of a trusted custodian; thus, the trusted records officer should have the qualifications for a trusted custodian as stated in principle C8.

A recordkeeping system that complies with the above requirements and procedures in its design and management is capable of ensuring the accuracy and authenticity of records after their creation, since these requirements and procedures establish the maximum degree of control with regard to the maintenance and use of the records.

(C7) Preservation considerations should be embedded in all activities involved in record creation and maintenance if a creator wishes to maintain and preserve accurate and authentic records beyond its operational business needs. (P7)

The concept of the records lifecycle in archival science refers to the theory that records go through distinct phases, including creation, use and maintenance and disposition (i.e., destruction or permanent preservation).

It is essential for records creators dealing with records in digital form to understand that, differently from what is the case with traditional records, preservation is a continuous process that begins with the creation of the records. Traditionally, records are appraised for preservation at the disposition stage, when they are no longer needed for business purposes. With digital records, decisions regarding preservation must be made as close as possible to the creation stage because of the ease with which they can be manipulated and deleted or lost to technological obsolescence.

The notion that records preservation starts at the creation stage requires that preservation considerations be incorporated and manifested in the design of record-making and recordkeeping systems. Each aggregation of records appraised for preservation should be identified in accordance with the classification scheme and records retention schedule established by the records creator, and this identification should be indicated among the records metadata. The aggregations of records so identified should be monitored throughout their lifecycle so that appraisal decisions and preservation considerations can be updated and/or modified to accommodate any possible change occurring after they are first made. To monitor and implement appraisal decisions and preservation considerations, the designated preserver should be given access to the organization's recordkeeping system. Policies and procedures should be established to facilitate constant interaction between the records creator and its designated preserver.

(C8) A trusted custodian should be designated as the preserver of the creator's records. (P1)

The designated records preserver is the entity responsible for taking physical and legal custody of and preserving[29] (i.e., protecting and ensuring continuous access to) a

[29] The term "preservation" is defined as "The whole of the principles, policies, rules and strategies aimed at prolonging the existence of an object by maintaining it in a condition suitable for use, either in its original

creator's inactive records.[30] Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a *trusted custodian* for a creator's records. To be considered as a trusted custodian, the preserver must:

- act as a neutral third party; that is, demonstrate that it has no stake in the content of the records and no reason to alter records under its custody and that it will not allow anybody to alter the records either accidentally or on purpose;
- be equipped with the knowledge and skills necessary to fulfill its responsibilities, which should be acquired through formal education in records and archives administration; and
- establish a trusted preservation system that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved.

For as long as the records are maintained by the creator in its recordkeeping system, they are active or semi-active records,[31] although under the responsibility of a trusted records officer. A records custodian trusted by the records creator as its designated preserver should maintain records that have been removed from the recordkeeping system for long-term or indefinite preservation. This trusted custodian will establish and maintain a preservation system to receive and preserve the creator's digital records. This involves ensuring that the accuracy and authenticity of the records received from the creator are assessed and maintained. Within the context of the preservation system, the designated preserver identifies appropriate preservation strategies and procedures, drawing on expertise from various disciplines, including archival science, computer science and law. The preservation procedures are implemented within the preservation system.

Only preservers that satisfy the requirements for trusted custodian are capable of fulfilling their duties of preserving authentic records over time and enabling a presumption of authenticity of the authentic copies they make for preservation purposes.

(C9) All business processes that contribute to the creation and/or use of the same records should be explicitly documented. (P10)

Records created in the course of carrying out one business function or one business process are often also used in the course of conducting other business functions or processes. In cases like this, records used in separate activities may be associated only with one activity in the records creator's record-making or recordkeeping system, or with none in some central "information" system or application. This practice creates difficulties for the records creator in identifying aggregations of records for

---

format or in a more persistent format, while leaving intact the object's intellectual form" (InterPARES 2 Terminology Database, op. cit.).

[30] An inactive record is defined as "A record that is no longer used in the day-to-day course of business, but which may be kept and occasionally used for legal, historical, or operational purposes" (Ibid.).

[31] An active record is defined as "A record needed by the creator for the purpose of carrying out the action for which it was created or for frequent reference" (Ibid). A semi active record is defined as "A record which is no longer needed for the purpose of carrying out the action for which it was created, but which is needed by the records creator for reference" (Ibid.).

accountability purposes and for its designated preserver in conducting appraisal and preservation activities.

It is recommended that policies and procedures be established that require detailed documentation of all business functions and processes contributing to the creation and use of the same records in any records creator's application or system and an explicit linkage between each record and the related workflow. Procedural manuals with such descriptions are effective in increasing the awareness of the impact of record-making and recordkeeping on the management of an organization. A subsequent different use of records after their creation can be captured by metadata, which are also capable of tracing the contexts in which records are generated.

(C10) Third-party intellectual property rights attached to the creator's records should be explicitly identified and managed in the record-making and recordkeeping systems. (P8)

Every records creator is usually aware that the records that it creates, or which are under its control or custody, contain information covered by intellectual property legislation. However, creators should also be aware that in some cases the intellectual property rights linked to a record may belong to a party other than the author and addressee.

All intellectual property rights attached to a record need to be documented in the metadata accompanying such record at the time that it is made or received and set aside. Intellectual property issues can significantly influence the reproduction of records, which is central to the processes of refreshing, converting and migrating records for either continuous use or preservation purposes. Subject to variations among different legislative environments, reproductions of records with intellectual property rights held by third parties may violate legislation that protects such rights. These issues must be identified and addressed at the stage of designing the record-making and recordkeeping systems. In the case of records identified for long-term preservation, long-term clearance of such rights should be addressed explicitly in the creator's record policy.

(C11) Privacy rights and obligations attached to the creator's records should be explicitly identified and protected in the record-making and recordkeeping systems. (P9)

Privacy legislation protects the rights of individuals with reference to personal data that may be part of any record used and maintained by a records creator with whom they have interacted. The limits of privacy depend on the legislative framework in which the records creator operates. The framework may be in conflict with the access policy linked to the mandate of the records creator and even with the access to information legislation in the same jurisdiction.

The presence of personal information within the records should be identified and documented within the metadata schema linked to the records in the record-making and recordkeeping systems of the creator. Metadata schemas that note and administer the use of personal information contained within the records must be embedded in record-making and recordkeeping systems. This will enable the protection of personal information

through the establishment of system-wide access privileges. In cases where records are to be preserved indefinitely, privacy issues relating to access to records must be expressly resolved (i.e., explicit permissions must be sought from the individuals concerned), ideally prior to record creation. This is the best way to ensure that the records are managed in accordance with privacy legislation and that the preserver will be able to effectively include the privacy issues relevant to the records in the preservation feasibility study during appraisal. The designated preserver for each records creator should, as a trusted custodian, be granted access to records containing personal information to perform preservation activities. Processing of personal information for maintenance or preservation purposes is different from the use of it for research or business purposes. Regardless of the legislative framework, the records creator should be able to demonstrate that processing of records containing personal information does not put such information at risk of unauthorized access.

Responsibility for processing records containing personal data for maintenance and preservation purposes must reside with the records creator and its legitimate successors. Although the practice of outsourcing these functions to specialized commercial operators is authorized and regulated under most existing privacy legislation, the practice should still be avoided whenever possible to minimize the number of individuals authorized to access and/or process the records, thus reducing the risk of unauthorized disclosure of personal information in the records and of jeopardizing the ability to obtain permission to process personal information for maintenance or preservation purposes.

In the case of records that are not yet designated for permanent preservation, appraisal decisions should be taken before the initial mandate for processing personal information has expired to ensure that the legal basis for retaining such records is still in force.

(C12) Procedures for sharing records across different jurisdictions should be established on the basis of the legal requirements under which the records are created. (P13)

Records creators with branches in geographically separate areas (i.e., areas that are covered by different legislation), must be aware that different access, privacy and intellectual property laws may have an impact on their records-sharing activities. Such sharing activities encompass records exchange within the records creator or with outside organizations, such as governments or business partners. This includes providing records to a trusted preserver, where the latter operates in a legal environment different from that of the records creator.

The fact that records are freely accessible in one jurisdiction does not imply that they can be accessed in the same way in other jurisdictions. Records creators must investigate such issues and address them in their policies.

(C13) Reproductions of a record made by the creator in its usual and ordinary course of business and for its purposes and use, as part of its recordkeeping activities, have the same effects as the first manifestation, and each is to be considered at any given time the record of the creator. (P3)

In the digital environment, the first manifestation of a record, be it a draft, an original or a copy, only exists when first composed in the creator's record-making system, if it is an internal record, or when first received in the creator's recordkeeping system, if it is transmitted from the outside. When the record is closed and saved into the record-making or recordkeeping system, its first manifestation technically disappears, as the saving action decomposes it into its digital components. Any later manifestation of the digital record is a reproduction resulting from an assembly of its digital components. Conceptually, however, records creators can use any reproduction of a record's first manifestation as if it were the record's first manifestation, as long as the reproduction is made in the usual and ordinary course of carrying out business activities and used for such activities. This means that each reproduction in sequence should have the same admissibility in court as the record's first manifestation and be given the same weight.

To establish that a record is reproduced in the usual and ordinary course of business, it is necessary to set out routine procedures in writing. In effect, if reliable records have been generated in a trusted record-making system and their accuracy and authenticity have been maintained together with that of the received records in the creator's recordkeeping system, then all records should have the same authority and effects as their first manifestation.

Although, according to the theory of the record (i.e., diplomatics), an "original" record in a digital system is the first manifestation of a received record and, if after closing such manifestation the original no longer exists, it might be useful to look at three examples of statutory laws pertaining to the meaning of "original." Common to all three variations is the principle that it is the relationship of a record to the business of the creator that determines whether the record in question has the authority and effects of an original.

> Example 1: The U.S. Federal *Rules of Evidence* distinguishes between originals and duplicates, with greater value as evidence given to originals. For digital records, it is noteworthy that if "data are stored in a computer or similar device, any printout or other output readable by sight, shown to reflect the data accurately, is an 'original.'"[32]
>
> Example 2: The quality of being original is acknowledged in Italian legislation in terms of adding weight or greater trustworthiness to records. Italian legislation emphasizes the difference between digital data (original) and any kind of output of those data (copy), by establishing that "any data or document electronically created by any public administration represents a primary and original source of

[32] United States House of Representatives, *Federal Rules of Evidence*, Article X. Contents of Writings, Recordings, and Photographs: Rule 1001. Definitions, Committee on the Judiciary, Committee Print No. 8 (December 31, 2004). Available at http://judiciary.house.gov/media/pdfs/printers/108th/evid2004.pdf. The same rule generalizes that "any counterpart" to the writing or recording "intended to have the same effect by a person executing or issuing it" is an original.

information that may be used to make copies on any kind of medium for all legal purposes."[33]

Example 3: The *Electronic Signatures Law of the People's Republic of China* regards a digital record as an original if it meets the two following qualifications: it must be 1) capable of presenting the content effectively and of being retrieved and consulted at any moment, and 2) capable of unfailingly showing the integrity of the content from the moment of its completion. However, annotations made to a data electronic document [digital record] and changes of presentation occurring in the process of data exchanging, storing and displaying are not considered to affect its integrity.[34]

Principles for Records Preservers

(P1) A designated records preserver fulfills the role of trusted custodian. (C8)

The designated records preserver is the entity responsible for taking physical and legal custody of and preserving (i.e., protecting and ensuring continuous access to) a creator's inactive records. Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a *trusted custodian* for a creator's records. To be considered as a trusted custodian, the preserver must:

- act as a neutral third party; that is, demonstrate that it has no stake in the content of the records and no reason to alter records under its custody and that it will not allow anybody to alter the records either accidentally or on purpose;
- be equipped with the knowledge and skills necessary to fulfill its responsibilities, which should be acquired through formal education in records and archives administration; and
- establish a trusted preservation system that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved.

The acquisition of a creator's records is undertaken by the preserver, who, after having assessed the accuracy and authenticity of the records, produces an authentic copy of them from the creator's recordkeeping system. Records that are acquired this way are authentic copies of the records of the creator identified for long-term preservation, because they are made by the designated preserver in its role of trusted custodian.

The authentic copies of the creator's records are then kept by the trusted custodian in a trusted preservation system, which should include in its design a description and a retrieval system. This trusted preservation system must also have in place rules and

---

[33] Italy, DPR 445/2000, art. 9, par. 1. Available at http://www.parlamento.it/parlam/leggi/deleghe/00443dla.htm.

[34] China, *Electronic Signatures Law of the People's Republic of China*, art. 5. Translated by Sherry Xie. See also Sherry Xie (2005). "InterPARES 2 Project - Policy Cross-domain: Supplements to the Study of Archival Legislation in China (Report I)," 3. Available at http://www.interpares.org/display_file.cfm?doc=ip2(policy)archival_legislation_CHINA_SUPPLEMENT.pdf

procedures for the ongoing production of authentic copies as the existing system becomes obsolete and the technology is upgraded. This requirement is consistent with the final recommendations of InterPARES 1, which developed the *Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records*,[35] a set of requirements to be implemented by the preserver. It should be noted that the simple fact of reproducing records in the preserver's preservation system does not make the results authentic copies; such designation must be provided by the preserver's authority.

A sustainable preservation strategy requires close collaboration between a records creator and its designated preserver as trusted custodian. It is the preserver's responsibility to take the initiative in collaborating with the creator to establish acquisition and preservation procedures and in advising the creator in any records management activities essential to the preserver's acquisition and preservation activities.

(P2) Records preservation policies, strategies and standards should address the issues of record accuracy and authenticity expressly and separately. (C4)

An accurate record is one that contains correct, precise and exact data. The accuracy of a record is assumed when the record is created and used in the course of business processes to carry out business functions, based on the assumption that inaccurate records harm business interests. However, when records are transmitted across systems, refreshed, converted or migrated for preservation purposes, or the technology in which the record resides is upgraded, the data contained in the record must be verified to ensure their accuracy was not harmed by technical or human errors occurring in the transmission or transformation processes. This verification of accuracy is the responsibility of the preserver who carries out the transmission or transformation process; however, such person is not responsible for the correctness of the data value, for which the creator remains accountable, just as is the case for the reliability of the records containing the data.

The concept of authenticity refers to the fact that a record is what it purports to be and has not been tampered with or otherwise corrupted. In other words, authenticity is the trustworthiness of a record as a record. A record is authentic if it can be demonstrated that it is as it was when created. An authentic record is as reliable and accurate as it was when first generated. Authenticity depends upon the record transmission and the manner of its preservation and custody. Thus, it is a responsibility of both the records creator and its legitimate successor (i.e., either the person or organization acquiring the function(s) from which the records in question result and the records themselves, or a designated records preserver).

Authenticity is protected and is verifiable by ensuring that the identity and the integrity of a record are maintained. The identity of a record is what distinguishes it from all other records. It is declared at the moment of creation by indicating at a minimum the following attributes: the names of the persons participating in the creation of the record

---

[35] See MacNeil et al., "Authenticity Task Force Report," op. cit., and, more specifically, Authenticity Task Force, "Appendix 2."

(e.g., author, addressee); the action or matter to which the record pertains; the date(s) of compilation, filing or transmission; the record's documentary form; the record's digital presentation (or format); the relationship of the record to other records through a classification code or a naming convention; and the existence of attachments. The record identity so declared must be maintained intact through time first by the creator and its trusted records officer while the record is in active or semi-active use, and subsequently by the designated records preserver when the record is designated as inactive. The integrity of a record is its wholeness and soundness and can only be inferred from circumstantial evidence related to the person who held responsibility for the record through time, from access privileges and access restrictions and from the indication of any annotation or modification (technical or otherwise) that such person(s) with access to record might have made to it. Thus, the establishment and maintenance of record integrity are supported by declaring the following record attributes: the names of the handling office(s), the office of primary responsibility[36] for the record over time and/or the recordkeeping office and the designated preserver; the access privileges code[37] and the access restriction code;[38] and the list of annotations[39] and of format changes.[40]

Authenticity is not a quality that can be bestowed on records after their creation and maintenance by any preservation process. A preserver can only protect and maintain what was transferred under its responsibility. Authenticity is protected and maintained through the adoption of methods that ensure that the record is not manipulated, altered, or otherwise falsified after its transfer. It is the preserver's responsibility to assess the authenticity of records considered for acquisition into a preservation system and to ensure that it remains intact after the transfer to such system by respecting within the preserving unit or organization the same *Benchmark Requirements* that bind the creator (e.g., access privileges, measure against corruption or loss) and the *Baseline Requirements* for preservers.

(P3) Reproductions of a creator's records made for purposes of preservation by their trusted custodian are to be considered authentic copies of the creator's records. (C13)

Reproductions of digital records in the creator's record-making and recordkeeping systems made in the usual and ordinary course of activity for either action or reference purposes can be considered to have the same authority and effects as the first manifestation of the same records. Reproductions of a creator's records for preservation

---

[36] Office of primary responsibility is defined as "The office given the formal competence for maintaining the authoritative version or copy of records belonging to a given class within a classification scheme" (InterPARES 2 Terminology Database, op. cit.).
[37] Access privileges code is defined as "The indication of the person, position or office authorized to annotate a record, delete it, or remove it from the system" (Ibid.).
[38] Access restriction code is defined as "The indication of the person, position or office authorized to read a record" (Ibid.).
[39] List of annotations is defined as "Recorded information about additions made to a record after it has been created" (Ibid.).
[40] List of format changes is defined as "Recorded Information about modifications to a record's documentary form or digital format after it has been created" (Ibid.).

purposes rather than in response to a creator's business need are considered authentic copies of the records of the creator, because they are never used in their present manifestation for action or reference by the creator itself. The creator's records and their authentic preservation copies are the same records but at different phases in their lifecycle and thus at a different status of transmission.[41] The former are used by their creator to achieve business goals, while the latter are made by the preservers for preservation purposes.

Copies of records in the preserver's preservation system may not be designated authentic if the preserver has made them for purposes other than preservation; for example, a copy from which personal identifiers are removed may be made for access purposes. Ultimately, only the preserver has the authority to designate a copy as authentic.

(P4) Records preservation procedures should ensure that the digital components of records can be separately preserved and reassembled over time. (C2)

Every digital record is composed of one or more digital components. A digital component is a digital object that is part of one or more digital records, including any metadata necessary to order, structure or manifest content and that requires a given preservation action. For example, an e-mail that includes a picture and a digital signature will have at least four digital components (the header, the text, the picture and the digital signature). Reports with attachments in different formats will consist of more than one digital component, whereas a report with its attachments saved in one PDF file will consist of only one digital component. Although digital components are each stored separately, each digital component exists in a specific relationship to the other digital components that make up the record.

Preservation of digital records requires that all the digital components of a record be consistently identified, linked and stored in a way that they can be retrieved and reconstituted into a record having the same presentation it manifested when last closed. Each digital component requires one or more specific methods for decoding the bit stream and for presenting it for use over time. The bit stream can be altered, as a result of conversion, for example, as long as it continues to be able to fulfill its original role in the reproduction of the record. All digital components must be able to work together after they are altered; therefore, all changes need to be assessed by the preserver for the effects they may have on the record.

The preserver must be prepared to advise the creator, directly or through development of recommended standards, on the types of digital components that the preserver's system is able to sustain. Where standards governing the types and formats of digital components are common to both the record-making and recordkeeping systems and the record preservation system, the preserver can directly influence the creator towards those

---

[41] In diplomatics, the status of transmission is the degree of perfection of record. There are three possible statuses of transmission: draft, original and copy. Copies are than further categorized according to their authority, and the most authoritative among the copies is the authentic copy; that is, a reproduction that is declared conforming to the reproduced entity by an officer having the authority to do so. Professional archivists are among such officers.

standards that will facilitate meeting the preservation requirements. Where no common standards exist or can reasonably be adopted, the preserver must understand the degree of interoperability of certain types and formats of digital components. This understanding will provide a basis for the preserver to assess the capability of the preservation system to preserve the digital components and their relationships as they emerge from the creator's record-making and recordkeeping systems.

Highly interoperable formats—that is, formats that are not tied to specific applications or versions of applications—are generally seen to provide a better basis for preservation work. It is important, however, not to focus exclusively on the interoperability of formats at the expense of the relationships between them that also must be preserved. For example, an HTML-based Web page may be comprised of digital components that are highly interoperable, but the version of HTML coding used to structure the components may be an old version with many deprecated terms (i.e., terms that are not recognized by current software browsers that may be used to reproduce the Web page).

(P5) Authentic copies should be made for preservation purposes only from the creator's records; that is, from digital objects that have a stable content and a fixed documentary form. (C1)

A record is defined by InterPARES, following the traditional archival definition, as "a document made or received in the course of a practical activity as an instrument or a by-product of such activity and set aside for action or reference."[42] This definition implies that, to be considered as a record, a digital object generated by the creator must first be a document; that is, must have stable content and fixed documentary form. Only digital objects possessing both are capable of serving the record's memorial function.

The concept of *stable content* is self-explanatory, as it simply refers to the fact that the data and the information in the record (i.e., the message the record is intended to convey) are unchanged and unchangeable. This implies that data or information cannot be overwritten, altered, deleted or added to. Thus, if one has a system that contains fluid, ever-changing data or information, one has no records in such a system until one decides to make one and to save it with its unalterable content.

The concept of *fixed form* is more complex. A digital object has a fixed form when its binary content is stored so that the message it conveys can be rendered with the same documentary presentation it had on the screen when first saved. Because the same documentary presentation of a record can be produced by a variety of digital presentations, fixed form does not imply that the bit streams must remain intact over time. It is possible to change the way a record is contained in a computer file without changing the record; for example, if a digital object generated in '.doc' format is later saved in '.pdf' format, the way it manifests itself on the screen—its documentary presentation, or "documentary form"—has not changed, so one can say that the object has a fixed form.

---

[42] See the InterPARES 2 Terminology Database, op. cit.

One can also produce digital information that can take several different documentary forms. This means that the same content can be presented on the screen in several different ways, the various types of graphs available in spread sheet software being one example. In this case, each presentation of such a digital object in the limited series of possibilities allowed by the system is to be considered as a different view of the same record having stable content and fixed form.

In addition, one has to consider the concept of "bounded variability,"[43] which refers to changes to the form and/or content of a digital record that are limited and controlled by fixed rules, so that the same query, request or interaction always generates the same result. In such cases, variations in the record's form and content are either caused by technology, such as different operating systems or applications used to access the document, or by the intention of the author or writer of the document. Where content is concerned, while, as mentioned, the same query will always return the same subset, its presentation might vary within an allowed range, such as image magnification. In consideration of the fact that what causes these variations also limits them, they are not considered to be violations of the requirements of stable content and fixed form.

Based on this understanding, any preservation policy should clearly state that reproductions of authentic copies for preservation purposes can only be made from the creator's records, as identified by the creator.[44]

The preserver should know (or help establish) the creator's criteria for identifying the digital objects that are maintained as records and the methods employed to stabilize their content and fix their form. This is consistent with the preserver's responsibility to advise the creator on its record creation processes and technologies. This advising activity will also provide the preserver with the critical information needed to understand the business activities and processes that caused the records to come into being and with the ability to assess their continuing identity and integrity.

(P6) Preservation requirements should be articulated in terms of the purpose or desired outcome of preservation, rather than in terms of the specific technologies available. (C3)

Digital records rely, by definition, on computer technology, and any instance of a record exists within a specific technological environment. For this reason, it may seem useful to establish record preservation requirements in terms of the technological characteristics of the records or the technological applications in which the records may reside. However, not only do technologies change, sometimes very frequently, but they also are governed by proprietary considerations established and modified at will by their developers. Both these factors can significantly affect the continued accessibility of digital records over time. For these reasons, references to specific technologies should not be included in preservation policies and standards. Only the requirements and obligations that the records are designed to support should be explicit within record preservation policies and

[43] See Duranti and Thibodeau, "The Concept of Record," op. cit.
[44] See principle C1 in the Principles for Creators regarding the identification of records.

standards. It is only at the level of implementation that specific technologies should, indeed must, be named.

Technological solutions to record preservation issues are dynamic, meaning that they will evolve as the technology evolves. This affects record preservation in two ways. First, it makes it possible to adopt new strategies to meet preservation needs, as happened with the use of XML to support the long-term preservation of structured records. Second, it creates opportunities for drawing on expertise from a number of disciplines. These two issues are interconnected. Thus, for example, while utilization of XML is, by itself, only one activity for preservation, it might be matched with using data grid technology as a stable and enduring platform to support XML-based records. By experimenting with these combinations, new archival knowledge will continue to be both acquired and required.

Technological solutions also need to be specific to be effective. Although the general theory and methodology of digital preservation applies to all digital records, the preservation solutions for different types of records require different methods. These should be based on the specific context in which the records are created and maintained, the functions and activities to which the records are linked and the technologies employed for record-making and recordkeeping to ensure the best solutions are designed for preserving each type of record.

Preservation policies that are expressed in terms of record requirements rather than technologies will be more stable, needing updates only if the record requirements change, rather than as the technology changes. Preservation action plans will likely need to be updated more frequently to identify appropriate technological solutions for the digital preservation of specific aggregations of records. The identified solutions must be monitored with regard to the possible need for modifying and updating.

(P7) Preservation considerations should be embedded in all activities involved in each phase of the records lifecycle if their continuing authentic existence over the long term is to be ensured. (C7)

The concept of the records lifecycle in archival science refers to the theory that records go through distinct phases, including creation, use and maintenance and disposition (destruction or permanent preservation).

It is essential for preservers who acquire digital records to understand that, differently from what is the case with traditional records, preservation is a continuous process that begins with the creation of the records. Analogue records are appraised for preservation at the disposition stage, when they are no longer needed by the creator for business purposes. With digital records, decisions relevant to preservation must be made as close as possible to the creation stage because of the ease and the speed with which digital objects can be manipulated, deleted by accident or on purpose, or lost to technological obsolescence.

The notion that records preservation starts at the creation stage requires that preservation considerations be incorporated and manifested in the design of record-making and

recordkeeping systems. Each aggregation of records appraised for preservation should be identified in accordance with the classification scheme and the records retention schedule established by the records creator in collaboration with the preserver, and this identification should be indicated in the records metadata. The records so identified should be monitored throughout their lifecycle by the preserver, so that appraisal decisions and preservation considerations can be updated to accommodate any possible changes occurring after they are first made. Appraisal decisions need to be reviewed to ensure that the information about the appraised 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. To monitor and implement appraisal decisions and preservation considerations, the designated preserver should obtain continuing access to the records creator's recordkeeping system within limits agreed upon with the creator and reflected in the preserver's access privileges. The preserver should establish procedures to facilitate constant interaction with the records creator.

(P8) Third-party intellectual property rights attached to the creator's records should be explicitly identified and managed in the preservation system. (C10)

Preservers know that records under records creators' control usually contain information covered by intellectual property legislation. They should also be aware that, in some cases, the intellectual property rights attached to records belong to a party other than the author; that is, the intellectual property rights reside with a third party. Third-party intellectual property rights should be documented in the metadata accompanying such records because they influence the processes of refreshing, converting and migrating them for either continuous use or preservation purposes. Subject to variations in different legislative environments, reproductions of records with third-party intellectual property rights attached to them may violate legislation that protects such rights. In the case of records identified for long-term preservation, long-term clearance of such rights should be addressed explicitly with the records creator.

Because preservation in a digital environment involves making copies, intellectual property rights have become an issue, not just for access as in the past, but for preservation. It is the preserver's responsibility; first, to advise the creator on how to address intellectual property issues in its record-making and recordkeeping systems, and, second, to ensure that intellectual property issues are addressed in the design of the preservation system. In particular, any issues relevant to third-party intellectual property rights should be cleared before the transfer of records to be preserved from the creator to the preserver. The latter must consider these issues as a part of the assessment of feasibility of preservation.

(P9) Privacy rights and obligations attached to the creator's records should be explicitly identified and protected in the preservation system. (C11)

Privacy legislation protects the rights of individuals with reference to personal data that may be part of any record used and maintained by a records creator with whom they have interacted. The limits of privacy depend on the legislative framework in which the records creator operates. It may be in conflict with the access policy linked to the

mandate of the records creator and even with the access to information legislation in the same jurisdiction. Besides lobbying for exceptions, the designated preserver should ensure that the consequences of the existing situation for preservation and access are clearly understood.

The presence of personal information within the records should be identified and documented among the metadata linked to the records in the record-making and recordkeeping systems of the creators. This is the best way to ensure that the records are managed in accordance with privacy legislation and that the preserver will be able to effectively include the privacy issues relevant to the records in the preservation feasibility study during appraisal. The designated preserver for each creator should, as a trusted custodian, obtain access to records containing personal information to perform preservation activities. Archival processing of personal information for preservation purposes is different from the use of it for research or business purposes. Regardless of the legislative framework, the creator and the preserver should be able to demonstrate that archival processing of records containing personal information does not put such information at risk of unauthorized access.

Preservers should also insist that responsibility for processing records containing personal data for preservation purposes must reside with the records creator and its legitimate successors. Although the practice of outsourcing these preservation functions to specialized commercial operators may be authorized and regulated under most existing privacy legislation, the practice should still be avoided whenever possible to minimize the number of individuals authorized to access and/or process the records, thus reducing the risk of unauthorized disclosure of personal information in the records and of jeopardizing the ability to obtain permission to process personal information for preservation purposes.

In the case of records that are not yet designated for permanent preservation, appraisal decisions should be taken before the initial mandate for processing personal information has expired to ensure that the legal basis for retaining such records is still in force.

(P10) Archival appraisal should identify and analyze all the business processes that contribute to the creation and/or use of the same records. (C9)

A record may be created for one purpose and then subsequently used for different purposes by different persons. Any appraisal decision should consider all uses of the record and be aware of the business processes behind them. This is necessary to make an informed decision about what to preserve as well as to be able to dispose effectively of all possible copies of the records that have not been selected for preservation.

The use of records or information within records by different business processes may be desirable from the creator's standpoint in terms of providing a degree of interoperability among the creator's information and record systems. In such situations, the preserver should advise the creator that metadata attached to records used by many business processes must identify each relevant business process. This is critical for the creator because it ensures the authenticity of the records by establishing their identity and integrity in each context. It is also critical for the preserver who must understand all

contexts in which the records were used to effectively undertake appraisal and also to meet the baseline requirements for maintaining authenticity for any records acquired into the preservation system.

(P11) Archival appraisal should assess the authenticity of the records. (C6)

Appraisal decisions should be made by compiling information about kept records and their context(s), assessing their value and determining the feasibility of their preservation.[45]

As part of the assessment of value, preservers must establish the grounds for presuming that the records being appraised are authentic. This means that preservers must ensure that each record identity has been documented and maintained as documented and must ascertain the degree to which the records' creator has guaranteed their integrity by making sure that its records are intact and uncorrupted. The evidence supporting the presumption of authenticity must be measured against the InterPARES *Benchmark Requirements*.[46]

(P12) Archival description should be used as a collective authentication of the records in an archival fonds. (C6)

Archival description of a fonds emerges from the comprehensive analysis of the various relationships interwoven in the course of the formation and accumulation of records and therefore is the most reliable means of establishing the continued authenticity of a body of interrelated records. While the authenticity of individual records can be in part established through their metadata, the authenticity of aggregations of records (i.e., file, series or fonds), can only be proved through archival description.

It has always been the function, either explicit or implicit, of archival description to authenticate the records by perpetuating their administrative and documentary relationships; but, with digital records, this function has moved to the forefront. In fact, as original digital records disappear and an interminable chain of non-identical reproductions follows them, the researchers looking at the last of those reproductions will not find in it any information regarding provenance, authority, context or authenticity.

The authentication function of archival description is different from that of a certificate of authenticity, because it is not simply an attestation of the authenticity of individual records, but a collective attestation of the authenticity of the records of a fonds and of all their interrelationships as made explicit by their administrative, custodial and technological history (including a description of the recordkeeping system(s) within which they have been maintained and used), the scope and content and the hierarchical representation of the records aggregates. It is also different both from the identity and integrity metadata attached to individual records, which are part of the record itself and

---

[45] See Terry Eastwood et al., "Part Two – Choosing to Preserve: The Selection of Electronic Records: Appraisal Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 67–98. Online reprint available at http://www.interpares.org/book/interpares_book_e_part2.pdf.

[46] See the already cited benchmark requirements in MacNeil et al., "Appraisal Task Force Report," op. cit.

are reproduced time after time with it and from the additional metadata attached to records aggregations (e.g., file, series) within the recordkeeping system to identify them and document their technological transformations.

The unique function of archival description is to provide an historical view of the records and of their becoming, while presenting them as a universality in which each member's individuality is subject to the bond of a common provenance and destination.

(P13) Procedures for providing access to records created in one jurisdiction to users in other jurisdictions should be established on the basis of the legal environment in which the records were created. (C13)

Different jurisdictions may have different laws and regulations with regard to access rights in relation to the protection of privacy, intellectual property and any other kind of public or private interests (e.g., market sensitive records). Preservers who are a unit of a records creator (e.g., in-house archival programs or archives) that has geographically separated branches falling under different legislation must be aware of the impact of such diverse legal contexts on their records-sharing activities. This will affect access policies relevant to both internal and external sharing activities.

Appendix B: Annotated Bibliography: Survey of Existing Educational Resources

A. Professional Associations

1. Association for Information and Image Management (AIIM) Information Management Training Courses – Online, In-Person
http://www.aiim.org/Education/Information-Management-Training-Online-Courses-IT-Systems.aspx

The Association for Information and Image Management (AIIM) was founded in 1943 and is dedicated to platform and vendor neutral dissemination of enterprise content management (ECM) and to providing a vehicle to promote discussion among professionals and users. AIIM offers two streams of education; Market Education and Professional Development. The Market Education stream utilizes tools such as a bi-monthly magazine, user guides and Webinars. The Professional Development stream provides Certificate Training Programs, AIIM Essential Courses, AIIM Expo and Conference, and ECM Solutions Seminars.

The certificate program courses are in Enterprise Content Management, Electronic Records Management, Business Process Management, Information Organization and Access, Enterprise 2.0, and Email Management. These courses can be taken online or in person.

The AIIM Essential Courses are online courses on specific topics, including E-Discovery, PDF/A, evaluating SharePoint, and digital asset management technologies.

AIIM charges a fee for membership (both individual and corporate) as well as for courses. Membership is not required to take the courses but there is often a discount for members. The online courses are delivered through audio and PowerPoint over the Web and downloadable hand-outs. Registrants are also pointed to related resources. Each module has an online exam and, with a grade of 70% or higher, registrants receive a certificate of completion.

Depending on the track selected (Strategy, Practitioner, Specialist or Master), the ERM Certificate Program is available as either an online course or a classroom course. Online courses consist of modules that last approximately 60 minutes. The courses are directed towards a number of stakeholders in RM, including Business Analysts, IT Management, Technical Staff, Records Management Personnel, Business Units, Vendors, Executives, Change Agents and Users.

Costs for both online and classroom classes vary, depending on whether the participant is a member of AIIM. Participants receive a certificate upon completion. The courses are intended for an international audience, with materials relevant for North America, Europe and Australia. Materials are delivered in English.

2. Association of Records Managers and Administrators (ARMA) a) Electronic Records and E-Discovery

Online Courses

http://www.aiim.org/Education/Information-Management-Training-Online-Courses-IT-Systems.aspx

ARMA International has four online courses listed on their Website. These courses are aimed at records management professionals. The courses are grouped into four sections: E-Mail, Voice Mail & Instant Messaging: A Legal Perspective, Electronic Discovery in 2010, Issues and Approaches in Archiving Electronic Records, and RIM 101: Fundamentals of Professional Practice. Some of these courses are free to members, while some are not; all non-members are charged a fee.

Web Seminar On Demand

https://www.arma.org/eWeb/DynamicPage.aspx?Webcode=ARMAISeminarArchive

ARMA International hosts several Web seminars on their Website which can be viewed at any time. These courses are aimed at records management professionals. Some of these Web seminars cover topics such as e-discovery, collaboration in electronic records management, software selection, developing retention schedules for electronic records, and creating the case for electronic records management. These Web seminars are free for members, and $35 for non-members.

3. International Council of Archives (ICA) Electronic Records: A Workbook for Archivists

http://www.ica.org/sites/default/files/Study16ENG_5_2.pdf

In 2005 the International Council of Archives Committee on Current Records in an Electronic Environment published this document aimed at "everybody who has an interest in the management and preservation of electronic records with a view to their accessibility over the long-term."

Electronic Records: A Workbook for Archivists is available online as a PDF in English and French. It presents a practical approach to electronic records management, with the viewpoint that archivists should be involved throughout the entire life cycle. Its governing principles and aims come from ICA's Guide for Managing Electronic Records from an Archival Perspective; terminology and definitions from ISO 15489-1 (Records Management). The workbook covers terminology, influencing strategies in records management, implementing recordkeeping requirements, preservation and access.

4. International Records Management Trust (IRMT) Training in Electronic Records Management (TERM)

http://www.irmt.org/educationTrainMaterials.php

The International Records Management Trust is a non-profit, UK-based organization dedicated to making information about managing records available at no cost to developing countries. Training consists of five modules (Understanding the Context of Electronic Records Management, Planning and Managing an Electronic Records

Management Programme, Managing the Creation, Use and Disposal of Electronic Records, Preserving Electronic Records, and Managing Personnel Records in an Electronic Environment), a resource list, a glossary of terms, good practice indicators for integrating records management with ICT systems, and route maps. The documents are posted on the IRMT Website in the form of PDF and Word documents.

5. UNESCO: E-Heritage

http://portal.unesco.org/ci/en/ev.php-URL_ID=1539&URL_DO=DO_TOPIC&URL_SECTION=201.html

The E-Heritage initiative is a United Nations Educational, Scientific and Cultural Organization's (UNESCO) effort aimed at preserving valuable archival holdings and library collections worldwide. UNESCO recognizes that digital information is subject to technical obsolescence and physical decay, coupled with the instability of the Internet. Consequently, UNESCO sought "international consensus on its collection, preservation and dissemination which resulted in the adoption of the UNESCO Charter on the Preservation of the Digital Heritage." The Charter was published in 2003 in seven languages namely: English, French, Spanish, Russian, Arabic, Chinese and German.

Also in March 2003, UNESCO developed a Guideline for the Preservation of Digital Heritage. The guideline covers areas in digital heritage, preservation, and how to manage preservation programs. It further touches on topics including how to build effective teams, team work, responsibilities of digital preservation managers, digital materials, and how to manage rights in the digital world. The guide is freely available online in English, French and Spanish.

6. Archives Association of British Columbia (AABC) The Archivist's Toolkit: Electronic and Born-digital Records

http://aabc.bc.ca/aabc/TK_08_electronic_records.html

AABC's "Archivist's Toolkit" is designed for archivists working in small to medium sized institutions. In a section entitled "Electronic and Born-digital Records," the Toolkit offers links to articles and talks given about electronic records, including topics divided as general resources, preserving authenticity, and physical preservation.

7. Society of American Archivists (SAA) Conference/Workshop Calendar

http://saa.archivists.org/Scripts/4Disapi.dll/4DCGI/events/ConferenceList.html?Action=GetEvents

The SAA puts on a variety of workshops devoted to archival training. A few of these are Web seminars – users pay a fee and are given two months of access to a 90-minute presentation on rotating topics. At the moment there are Web seminars entitled

"Electronic Records: Preservation of PDF" and "Thinking Digital… Practical Session to Get You Started."

SAA Archival Education Directory

http://www.archivists.org/prof-education/edd-index.asp#about

8. Australian Society of Archivists Inc. (Australia)

http://www.archivists.org.au/

The Australian Society of Archivists (ASA) Inc. offers occasional seminars and workshops on a range of archival and records management issues of which preservation is a component. One educational resource by the Society that addresses preservation is its flagship publication, *Keeping Archives*, 3rd Edition (*KA3*) at the cost of $130 for non-members and $100 for members.

9. The Japan Society of Archives Institutions (Asia)

http://www.jsai.jp/

B. Major Research Projects on Digital Preservation

10. Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR)

CASPAR Digital Preservation User Community

http://www.casparpreserves.eu/

CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval) is a digital preservation project begun in April 2006. It is developing and testing preservation strategies based on the OAIS (Open Archival Information System) reference model (IS0: 14721:2003) and is funded in part by the European Union. It offers training modules to its members (membership is freely available to individuals and organizations through a registration form on the Website). It also has papers in PDF form and videos available on their general Website.

11. Digital Preservation Coalition (DPC)
Digital Preservation Roadshows 2009 – 2010
http://www.dpconline.org/training/roadshows-2009-2010.html

The UK based Digital Preservation Coalition was established in 2001 as a not for profit organization dedicated to preserving digital resources in the UK in particular, and globally in general, through international collaboration on preservation issues and open dissemination of the results of this collaboration. Membership to the Coalition is free but is contingent upon collective or not for profit status. The Coalition provides training in the form of a handbook which is available either in PDF form or online. Part of this

handbook is the "Decision Tree" which is a tool to build policies for selecting digital materials for long-term preservation. It also has a section on media formats and how they rate in terms of long-term preservation.

What's New in Digital Preservation
http://www.dpconline.org/docs/whatsnew/whatsnew19.pdf

12. Creative Archiving at Michigan & Leeds: Emulating the Old on the New (CAMiLEON)
13.
http://www.si.umich.edu/CAMILEON/

The CAMiLEON project was started in 1999. The UK component of the project ended in 2002, while the US component ended in 2003. The project had three main objectives to: 1) explore the options for long-term retention of the original functionality and 'look and feel' of digital objects, 2) investigate technology emulation as a long-term strategy for long-term preservation and access to digital objects, and 3) consider where and how emulation fits into a set of digital preservation strategies.

Some deliverables of the project are guidelines for using different technical strategies (migration and emulation) to preserve digital entities and also, strategies for preserving digital objects that have no known method of long-term preservation and access. The project provided a collection of links to a range of information related to emulation and preservation. The links were categorized under the headings: publications, emulation and open source projects; old and new technological developments; computer games and emulation; miscellaneous Websites; and news articles. Although most of these preservation resources are free online, some of the sites require subscription or are in a language other than English.

13. Digital Curation Centre (UK)

http://www.dcc.ac.uk/

14. Electronic Resource Preservation and Access Network (ERPANET)

http://www.erpanet.org/

15. European Commission on Preservation and Access

http://www.knaw.nl/ecpa/

The European Commission on Preservation and Access (ECPA) was a set up in 1994 with the mandate to "promote the preservation of the documentary heritage in Europe." A key component of the project was the Training for Audio-visual Preservation in Europe, (TAPE) program which was to "explore the requirements for continued access to audio-visual materials and the application of new technologies for opening up collections that

provide living documentation of the world of the 20th century." The ECPA program was brought to a close in 2008. Some publications resulting from the program are online as PDFs and include short guidelines for video digitization and reports on audio visual issues.

16. Preservation and Long-Term Access Through Networked Services (PLANETS)

http://www.planets-project.eu/

Digital Preservation—The PLANETS Way is a three-day workshop, incorporating lectures, demonstrations, hands-on exercises, case studies, speaker panels, social networking and facilitated discussions on digital preservation. PLANETS' educational programs are aimed at archival organizations in Europe, with workshops held in various locations throughout Europe. Each workshop is targeted and tailored to the European region in which it is being held. More specifically, PLANETS targets consortium partners, European national libraries and archives and large data-holding institutions in Europe. Within these organizations, PLANETS targets CEOS, Head of IT, Heads of Preservation or Conservation, and Digital Preservation Managers and Repository Managers. The first day of the three-day workshop is targeted at both decision makers and preservation staff, thus encompassing all three target audience groups. The second and third days consist of hands-on workshops geared towards preservation staff.

Bibliography

Snow, Kellie, *Revised Training Plan* (London: Planets, 2009). http://www.planets-project.eu/docs/reports/Planets_DT6-D4_Training_Plan.PDF (accessed February 21, 2010).

17. Digital Preservation Europe Training and Education

http://www.digitalpreservationeurope.eu/publications/presentations/Jelena_Saikovic.PDF

18. Cornell University: Digital Preservation Management

http://www.icpsr.umich.edu/dpm/dpm-eng/eng_index.html

This initiative began at the University of Cornell and is now part of the Inter-university Consortium for Political and Social Research (ICPSR). It consists of an online tutorial in English, French and Italian which is designed to both stand alone and to serve as an introduction to a five day workshop held in Cambridge, MA. It is designed to help cultural institutions develop a plan to create a sustainable preservation program of their digitized and born-digital material.

The tutorial has an introduction, a conclusion, and six sections; Setting the Stage, Terms & Concepts, Obsolescence and Physical Threats, Foundations, Challenges, and Program Elements. There is an option to provide feedback and ask questions. These options are always available on the left of the site so the user can negotiate through them at any time.

The project is based around three interconnected themes to achieve successful digital preservation; organizational and technological resources and dedicated resources. The entire tutorial is available in PDF form but it is best delivered online as it is highly interactive with quizzes and many links to real world examples and further resources.

19. Inter-University Consortium for Political and Social Research (ICPSR)

Digital Preservation
http://www.icpsr.umich.edu/icpsrWeb/ICPSR/curation/preservation.jsp

The Digital Preservation Training Program was developed by Cornell University Library at Ithaca, New York. The goal of the program is to enable "effective decision making for administrators" so they can develop the capacity to make responsible preservation decisions that will ensure the longevity of digital objects.

Currently, the program is being administered by the Inter-University Consortium for Political and Social Research (ICPSR) in support of their mission to provide "*leadership and training in data access, curation, and methods of analysis for a diverse and expanding social science research community." The program has two training components: 1)* Digital Preservation Management Workshop, and 2) Digital Preservation Management Tutorial. The workshop is purposed "to foster critical thinking in a technological realm and provide the means for exercising practical and responsible stewardship of digital assets." It is intended for managers in libraries, archives and other cultural institutions. The content of the workshop is made up of presentations, group discussions, exercises and individual assignments. Attendees are required to take the Digital Preservation Online Tutorial – a prerequisite on which the content of the workshop is built. Before enrolment, interested attendees follow a simple procedure: application, notification and payments. Cost: $750.

20. Digital Preservation for the UK HE/FE Web Management Community

The Preservation of Web Resources Handbook (PoWR)

http://jiscpowr.jiscinvolve.org/files/2008/11/powrhandbookv1.PDF

JISC Preservation of Websites (PoWR) Handbook was developed in 2008 and is available for free download. It is aimed at information and asset managers; Webmasters, IT specialists and system administrators; records managers and archivists. The *Handbook* aims to raise awareness of Web site preservation and to lay out strategic principles and practical steps in achieving it. It deliberately stays away from the specifics of the technology in order to remain relevant through technological change and obsolescence. The *Handbook* consists of two parts, Web sites in general and Web sites in the context of organizations. This second part consists of strategies to work within your organization's workplace culture and how to raise awareness and gain management buy-in for the importance of preserving Web sites.

21. Arts and Humanities Data Service (AHDS)

AHDS Preservation Consultancy
http://ahds.ac.uk/preservation/preservation-consultancy.htm

The Arts and Humanities Data Service (AHDS) through a consultancy has produced a number of reports on digital preservation. Some of the reports are educational and address preservation issues such as taxonomy of data types, preservation metadata and ingest procedures framework. Also, AHDS has a comprehensive bibliography on digital preservation addressing preservation concerns as it regard text, sound, video, databases, etc. and a preservation glossary. These resources are free online in PDF and html formats.

22. Digitale duurzaamheid

Digital Preservation Testbed
http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=298&catagorie=6

In an effort to identify best strategies for preserving digital records, the Digital Preservation Testbed, in 2001 and 2003, "carried out a series of experiments with different types of digital files..." Materials such as emails, text documents, databases and spread sheets were collected from various organizations including government ministries and agencies as test beds for the experiments.

The results of the research were published in a series of publications entitled: From digital volatility to digital permanence and were mainly recommendations for preserving digital records in the areas of database, spread sheets, text documents and emails. The program also investigated three main strategies by which digital information could be preserved: migration, emulation and XML. This was an effort "to ensure that Dutch government digital information is sustainable, properly managed, and can be preserved in an authentic and re-usable manner for the long-term."

23. Regional Universities Building Research Infrastructure Collaboratively (RUBRIC)

Digital Preservation Management
http://www.rubric.edu.au/packages/RUBRIC_Toolkit/default.htm

Based on a recognition of the needs of regional and smaller universities to access, manage and disseminate research information; and the fact that significant research output is now often born digital, RUBRIC in 2007, launched a preservation toolkit. The toolkit which was a major deliverable of the RUBRIC project which was aimed at establishing institutional repositories. The content of the toolkit includes "information on building and managing a repository, publicity and marketing, Metadata, the Research

Quality Framework (RQF), content population and data issues." The toolkit is freely available online.

C. National Archives'/Libraries' Initiatives

24. National Archives and Records Administration (NARA)

Electronic Records Management Initiative
http://www.archives.gov/records-mgmt/initiatives/erm-overview.html

NARA's Electronic Records Management Initiative is intended to help government organizations create, use and maintain electronic records. In terms of preservation, it is intended to help organizations keep their electronic records accessible for as long as they need them and also to aid in transferring to NARA records deemed worthy of permanent retention. It provides an online toolkit, a handbook, policies, and links. Rather than true tutorials however, it mainly serves as a resource of links.

Records Management brochures and pamphlets
http://www.archives.gov/publications/records-mgmt.html

NARA Toolkit
http://www.archives.gov/records-mgmt/toolkit/

Electronic Record Archives (ERA)
http://www.archives.gov/era/

25. National Library of Australia

Preserving Access to Digital Information (PADI)
http://www.nla.gov.au/padi/

Preserving Access to Digital Information (PADI) is a preservation initiative of the National Library of Australia. The aim of the program is "to provide mechanisms that will help to ensure that information in digital form is managed with appropriate consideration for preservation and future access."

Two main channels used to access the program's digital information resources are 1) a subject gateway Website, and 2) a discussion list, which enables the sharing and exchange of ideas on digital preservation concerns. The subject gateway Website provides a pool of resources on digital records including digitization, rights and management, Web archiving and digital records. Other forms of preservation resources that are enlisted are books, articles, surveys, policies, strategies and guidelines, projects and case studies.

26. National Archives of Australia

Preserving electronic records: Strategies to ensure records remain accessible in the long term

http://www.naa.gov.au/records-management/secure-and-store/e-preservation/index.aspx

The National Archives of Australia offer advice to agencies on strategies that helps to ensure continued access to electronic records over the long-term. The Archives also provide information on long-term access to archival records of Commonwealth countries.

Digital Recordkeeping: Guidelines for Creating, Managing and Preserving Digital Records
http://www.naa.gov.au/records-management/publications/Digital-recordkeeping-guidelines.aspx

This digital recordkeeping guideline was developed in 2004 to provide the Australian Government and agencies a parameter for "creating, managing and preserving their digital records." It also aims to help managers of digital records and information to understand various issues of digital records management. The guidelines contain information on problems confronting digital recordkeeping in the Australian context and offer solutions accordingly. The guide is available online as a free download in PDF format.

27. Public Record Office of Victoria VERS project

http://www.prov.vic.gov.au/vers/vers/default.asp
http://www.prov.vic.gov.au/vers/standard/

28. National Library of Australia

http://www.nla.gov.au/preserve/index.html

The National Library of Australia has a responsibility "to preserve Australia's documentary heritage, and to make sure it is available for people to use for as long as possible." The Library places emphasis not only on material collection but also on preserving it for future use. The preservation program is therefore aimed at maintaining and preserving items according to their use and their significance. Some of the activities that the Library undertakes to achieve its preservation goals are digitization, digital preservation and digital archiving.

29. The National Archives (UK)

http://www.nationalarchives.gov.uk/information-management/

The National Archives of the United Kingdom provides Information Management for government departments and employees, however, the content is available to everyone. There are several ways to access the information, for example through quick links and through the section *Guidance and Standards*, which is a subject based dictionary. The Electronic Records Management entry has PDF documents on topics such as designing a business classification scheme and a series of eight workbooks on the creation, maintenance and disposition (including transfer to the archives) of electronic records. The workbooks present the information and have a questionnaire for the user to determine

how their department or organization measures up to the guide and to identify areas that need improvement. There are also several documents on managing digital continuities including a guide, *Managing Digital Continuity*, which discusses how to develop strategies to ensure access to digital assets now and in the future.

The National Archives also provides a list of links to other resources and several documents called *Guiding Notes* that are intended to give general information on digital preservation to a broad audience.

Links to other preservation Web sites:

http://www.nationalarchives.gov.uk/recordsmanagement/related-Websites.htm

- **Selecting File Formats for Long-Term Preservation**
  http://www.nationalarchives.gov.uk/documents/selecting-file-formats.PDF
- **Selecting Storage Media for Long-Term Preservation**
  http://www.nationalarchives.gov.uk/documents/selecting-storage-media.PDF
- **Care, Handling and Storage of Removable Media**
  http://www.nationalarchives.gov.uk/documents/removable-media-care.PDF
- **Graphic File Formats**
  http://www.nationalarchives.gov.uk/documents/graphic-file-formats.PDF
- **Image Compression**
  http://www.nationalarchives.gov.uk/documents/image-compression.PDF
- **Digital Preservation**
  http://www.nationalarchives.gov.uk/preservation/digital.htm
- **The National Digital Archive of Datasets (NDAD)**
  http://www.ndad.nationalarchives.gov.uk/
- **NSF-DELOS Working Group on Digital Archiving and Preservation Invest to Save**
  http://eprints.erpanet.org/48/01/Digitalarchiving.PDF

30. New York State Archives Digital Preservation

http://www.archives.nysed.gov/a/records/mr_erecords.shtml

The New York State Archives organizes workshops and Webinars on digital records as a training opportunity for various audiences. The workshops and Webinars are administered by the Local Government's office and they cover a range of electronic records preservation topics. The goal of the workshop is to educate attendees on electronic media and techniques for preserving electronic records. The workshops are usually four hours in length and require free online registration. They are primarily free

and open to the general public with some exception. The Archives also provides free downloadable instructional videos on government records management.

31. Utah State Archives Electronic Records

http://www.archives.state.ut.us/recordsmanagement/ERM/electronic-records-links.html

The Utah Archives and Records Service Division conducts records management training workshops for state employees and the PowerPoint slides for these workshops are available to everyone in PowerPoint format or as PDFs. Their Website has a section on electronic records that provides a business case for electronic records management, general electronic records guidelines, e-mail guidelines, and a large list of further resources. The material is designed to help government employees manage records in general and provide advice and guidance on how to treat electronic records scheduled for transfer to the archives for permanent retention in particular. The archives preservation strategy is magnetic tapes.

32. The National Archives (UK)

http://www.nationalarchives.gov.uk/information-management/

33. Library of Congress

Digital Preservation
http://digitalpreservation.gov/

This resource distributes the work of the National Digital Information Infrastructure and Preservation Program (NDIIPP) of the Library of Congress and its partners who are researching methods and strategies for preserving digital content whether digitized or born-digital. It aims to educate both cultural institutions and individuals about preserving digital heritage. It delivers content through a variety of formats including articles in PDF form, videos, podcasts and lists of tips. Content ranges from reports from preservation partners (for example the Preserving Digital Public Television Project) to presentations by, and interviews with, digital preservation leaders (for example David Rosenthal of Lots of Copies Keeps Stuff Safe: LOCKSS).

34. Digital Library Federation

Digital Preservation
http://www.diglib.org/preserve.htm

35. Library Preservation at Harvard

Guidance for Digitizing Images
http://preserve.harvard.edu/guidelines/imagedig.html

As part of its institutional commitment to maintain the usability of content in electronic form, library and archival material, the Library Preservation at Harvard (through the Weissman Preservation Center) engages in digital preservation education and training programs. The programs are offered through workshops, master classes and internships. They are offered two or three times per year for general and specialist audiences. The audience for these programs are largely made up of the Harvard community.

Digital Preservation: A Brief Resource List
http://preserve.harvard.edu/bibliographies/digpresintro.PDF

This list contains links to other digital preservation communities and portals. It also includes a list of selected articles, books standards and preservation repositories.

D. Other Miscellaneous Projects

36. U. S. Environmental Protection Agency

http://www.epa.gov/records/index.htm

The U.S. Environmental Protection Agency (EPA) provides training material for its employees online and as PDFs. The organization takes a lifecycle approach. Topics covered include policy, legal requirements, metadata, how to approach different record formats including databases, e-mail and Websites, and how to schedule records including how to prepare permanent records for eventual transfer to NARA. The material is freely available although there are some links that are only available to employees.

37. United States Department of Agriculture Records Management

http://www.ocio.usda.gov/records/electronic.html

38. University of the Pacific

Establishing & Managing Successful Records Management Programs
http://Web.pacific.edu/Documents/school-cpce/2010 Records Mgmt Seminar Broch-Web-Email.PDF

This is a two-day workshop on records management offered for credit by the university. The course targets managers, business consultants, vendors and records management professionals in government, schools, legal and corporate organizations. Its scope includes both paper and electronic records, providing a beginning understanding of records management. The emphasis seems to be on the development of a records management program.