InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records

PART SEVEN

STRUCTURING THE RELATIONSHIP
BETWEEN RECORDS CREATORS
AND PRESERVERS

Policy Cross-domain Task Force Report

[including Appendix 19]

by

Luciana Duranti, The University of British Columbia
Jim Suderman, City of Toronto Archives
Malcolm Todd, National Archives of the United Kingdom

Status: Final (public)

Version: Electronic

Submission Date: February 2007

Publication Date: 2008

Project Unit: Policy Cross-domain Task Force

URL: http://www.interpares.org/display_file.cfm?doc=ip2_book_part_7_policy_task_force.pdf

How to Cite: Luciana Duranti, Jim Suderman and Malcolm Todd, "Part Seven—Structuring the Relationship Between Records Creators and Preservers: Policy Cross-domain Task Force Report," [electronic version] in *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*, Luciana Duranti and Randy Preston, eds. (Padova, Italy: Associazione Nazionale Archivistica Italiana, 2008). <http://www.interpares.org/display_file.cfm?doc=ip2_book_part_7_policy_task_force.pdf>

Table of Contents

Introduction[1]

Policy considerations are all-pervasive in the aims of a research project such as InterPARES because record creation, maintenance and preservation are integral components of many human activities and need the same explicit directions as those for the activities themselves. Policy statements should be most explicit in the juridical dimensions of record authenticity that the archival dimensions must satisfy and in the moral and ethical requirements of records preservation arising from the function of carrying forward the traces of societal memory. Implementation of the findings and recommendations of the two phases of the InterPARES Project are dependent on engagement with the agendas of policymakers to advocate the benefits of implementing the Project's recommendations. This report reflects primarily the research undertaken by the members of the Policy Cross-domain Task Force, especially in relation to legislation relating to the authenticity and maintenance of records and the rights and obligations on creators and users of records.

In practice, all InterPARES areas of enquiry are touched by and have implications for policy at an international, national, sectoral and organizational level. This is the reason why the Project has established in its second phase, hereinafter called InterPARES 2, a dedicated Policy Cross-domain that inherited the function of the Strategy Task Force of its first phase, hereinafter called InterPARES 1, which defined policy as:

> a formal statement of direction or guidance as to how an organization will carry out its mandate, functions or activities, motivated by determined interests or programs.[2]

As such, policy may be expressed through laws, regulations, standards (professional, industry and technical), ethical codes, codes of conduct or practice and guidelines. The implementation of laws and regulations related to the creation and maintenance of records—in other words, public policy—is expressed, or should be, in an organization's records policy. The implementation of a records policy by an organization or an individual, in turn, results in, or should result in, the creation, maintenance and preservation of records, and their associated metadata, that can be used for further action and reference and as evidence of the activities from which they result.[3] InterPARES 2 researchers involved in the Description Cross-domain team explore elsewhere the relationship between records, records' metadata and preservation.

The Policy Cross-domain has examined the policies and strategies that affect the preservation of authentic digital records produced in the course of artistic, scientific and governmental activities. As organizational activities adopt increasingly rich *yet dynamic and thus somewhat unstable* technologies, the preservation challenge grows. The use of dynamic, interactive and experiential systems to carry out organizational activities reflects the common practice of adopting technologies without considering, let alone resolving, the preservation challenges that they present for the records generated and kept in them. Indeed, while this is clearest in e-government policy initiatives, it is also a fact of life in the sciences and the arts.

---

[1] The authors acknowledge the general contribution of all members of the Policy Cross-domain in the preparation of this report. In particular, we thank Mahnaz Ghaznavi, Ken Hawkins and Tracey Lauriault for their contributions to the text and editorial guidance. Any errors of representation or omission are the responsibility of the authors.

[2] Luciana Duranti et al., "Part Four – An Intellectual Framework for Policies, Strategies, and Standards: Strategy Task Force Report," note 1, in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 118. Online reprint available at http://www.interpares.org/book/interpares_book_g_part4.pdf.

[3] See, for example, a recent court case in the United States, *Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640 (D. Kan. 2005)*, which features the admissibility of metadata as an integral part of legal discovery. Available at http://www.ksd.uscourts.gov/opinions/032200JWLDJW-3333.pdf.

New models for collaboration and production, the outsourcing of activities and functions and the privatization of many parts of the public domain introduce new challenges for records retention. Legislation, case law and multi-national agreements form an intricate and often inconsistent and internally conflicting regulating infrastructure that, rather than facilitating the proper creation and use of digital objects, makes these activities increasingly complex.[4] Taken together, recent changes in technology, public policy and business models have put at risk the ability of organizations to undertake some of the activities necessary for the preservation of records. As a result, part of the Policy Cross-domain research has been to identify and counter specific *barriers* to preservation.

The final product of the Policy Cross-domain consists of an intellectual framework for policy development comprising two sets of principles that distil all the other research activities of the research team. The principles were conceived as instruments to fulfil the research goal of the Cross-domain—to develop model policies and strategies for the long-term preservation of authentic digital records—in the most concise and effective form. By following the appropriate set of principles, records creators or preservers will be able to develop organizational directives, in the form of guidelines, instructions or policy proper, capable of ensuring the continuing preservation of authentic digital records according to methods that reflect and allow for the correct implementation of the findings of the whole InterPARES research. Ideally, these principles should also be enshrined in supranational, national, sectoral or organizational policies, strategies and standards.

Other research units of InterPARES 2 have developed guidelines aimed at achieving the same sort of outcome in a much smaller organization, even down to individual practitioner level.[5]

Research team

Team objective

The InterPARES Project set out the following responsibilities for the Policy research team:

> The Policy Research Team will analyze the existing policies and strategies in each domain and focus of inquiry in light of the work being done by the working groups and then distil from the findings and products of the working groups policies, strategies and guidelines for the reliable and accurate creation and maintenance of the records under examination, and their authentic preservation within the context of each activity and culture generating them.[6]

Throughout the duration of the research, the three focus task forces (i.e., arts, science and government) carried out records creator-based case studies. The data from these case studies and the diplomatic analysis and modeling activities carried out on them are the core research data of InterPARES 2. However, to reach its own specific objectives, the Policy Cross-domain

[4] Susan Gutman, Luke Meagher and Adele Torrance (2006), "InterPARES 2 Project - Copyright Policy Annotated Bibliography, Draft version 4." Available at http://www.interpares.org/display_file.cfm?doc=ip2(biblio)_copyright-annotated.pdf.

[5] For the preserver's procedures, see the Domain 3 Task Force Report and the *Preserver Guidelines* in Appendix 21. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_4_domain3_task_force.pdf. The Guidelines also are available in booklet form at http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)preserver_guidelines_booklet.pdf. For the creator's procedures, see the Domain 1 Task Force Report and the *Creator Guidelines* in Appendix 20. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_2_domain1_task_force.pdf. The Guidelines also are available in booklet form at http://www.interpares.org/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf.

[6] See Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.

conducted research on international and national legislation, regulations, directives, etc., to determine the guidance presently provided to the development of policies and strategies and the issues they raise in relation to long-term preservation of authentic records.

Team composition

The Policy Cross-domain comprised researchers from a mixture of academic, archival and cultural heritage institutions, assisted by graduate research assistants from the universities of British Columbia and California, Los Angeles. During the first half of the Project, the Cross-domain was chaired by Sharon Farb of the University of California, Los Angeles, and Livia Iacovino of Monash University, Australia. In mid-2004, the chairmanship passed to two of the present authors, both of whom were previously team members. Tasks such as data collection and initial analyses were typically carried out by research assistants under the leadership of the researchers, who undertook more involved and complex analyses, wrote reports and liaised with other research units. Some doctoral students also participated in the latter tasks. The Cross-domain's international membership helped overcome language barriers where precision in recognizing the importance of juridical and other policy instruments is of the essence.

The following is a complete list of researchers and research assistants who participated in the Policy Cross-domain Task Force throughout the Project.

Co-chairs:

Sharon Farb	2001-2004
Livia Iacovino	2001-2004
Malcolm Todd	2004-2006
Jim Suderman	2004-2006

Researchers:

Howard Besser	New York University, USA
Hannelore Dekeyser	Katholieke Universiteit Leuven, Belgium
Luciana Duranti	The University of British Columbia, Canada
Philip Eppard	University of Albany, State University of New York, USA
Sharon Farb	University of California, Los Angeles, USA
Mahnaz Ghaznavi	J. Paul Getty Trust, USA
Kevin Glick	Yale University, USA
Elaine Goh	National Archives of Singapore
Maria Guercio	University of Urbino, Italy
Chenhui Hao	State Archives Administration of China
Livia Iacovino	Monash Univerisity, Australia
Terry Maxwell	University of Albany, State University of New York, USA
Evelyn McLellan	Insurance Corporation of British Columbia, Canada
Du Mei	State Archives Administration of China
Shelby Sanett	U.S. National Archives and Records Administration
Jim Suderman	Archives of Ontario, City of Toronto Archives, Canada
Kate Theimer	U.S. National Archives and Records Administration, USA
Malcolm Todd	National Archives of the United Kingdom

Research Assistants:

Barbara Bean	University at Albany, State University of New York, USA

Jessica Bushey	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Erin Coulter	The University of British Columbia, Canada
Seth Dalby	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Adam Farrell	The University of British Columbia, Canada
Fiorella Foscarini	The University of British Columbia, Canada
Susan Gutmann	The University of British Columbia, Canada
Peggy Heger	The University of British Columbia, Canada
Sarah Henshaw	The University of British Columbia, Canada
Sharif Khandaker	The University of British Columbia, Canada
Greg Kozak	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Luke Meagher	The University of British Columbia, Canada
Catherine Miller	The University of British Columbia, Canada
Rachel Mills	The University of British Columbia, Canada
Jane Morrison	The University of British Columbia, Canada
Emily O'Neill	The University of British Columbia, Canada
Cara Payne	The University of British Columbia, Canada
Hema Ramasamy	The University of British Columbia, Canada
Geneviève Shepherd	The University of British Columbia, Canada
Melissa Taitano	University of California, Los Angeles, USA
Adele Torrance	The University of British Columbia, Canada
Catherine Yasui	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada

Research Methodology

The following excerpt from the Policy Cross-domain's research statement describes the team's methodology:

> The Policy Research Team will research and analyze the existing policies, strategies, guidelines and standards in each of the focus areas in relation to each of the domains, examine how they may apply to the digital environments under investigation, compare them to recognize commonalities and differences and identify gaps, especially in relation to the new issues arising from the accessibility, use, manipulability and fragility of the types of records being studied. It will then examine the results of the case studies and of the work carried out in the three domains. On the basis of this analysis, it will articulate principles that should guide the development of policies, strategies and standards for the creation, maintenance, appraisal and preservation of the records in question and give them to the national and multinational teams for contextualization. Upon receiving the requested feedback, the Team will produce guidelines for those responsible for developing policies, strategies and standards at the international, national and organizational level.[7]

[7] InterPARES 2 Policy Cross-domain Methodologies. Available at http://www.interpares.org/ip2/ip2_policy.cfm.

Owing to the time required to carry out the separate rounds of case studies, typically one year to eighteen months, the early phases of the team's research had little case study data with which to work. As a result, the first tasks undertaken were concerned with gathering other sources of existing policy and with their analysis.

Policy of the highest level, such as national and supranational laws and directives, was given priority. Liaison with other research teams, particularly Focus 3 (government), which had several researchers in common with the Policy Team, was particularly helpful in this respect. Key themes were discussed, such as authentication methods, the issues—new for InterPARES in this second phase of the Project—of accuracy and reliability, and emerging technologies such as Digital Asset Management. Policy data were compiled in comparative tables. Individual researchers worked on issue papers or scholarly papers related to their own jurisdiction and interests, while an InterPARES 2 moral rights panel discussed key challenges at the 2004 Conference of the Association of Canadian Archivists in Montreal.

After the mid-term InterPARES plenary workshop of September 2004, with the emergence of the first case study data, five discrete studies, briefly outlined below, were undertaken by the Policy Team.

1. An annotated bibliography on policy related to intellectual property covering a selection of national and supranational jurisdictions[8] and a study of the case study data relevant to the following research questions: "To what extent does society restrict use and impede preservation to protect the interest of copyright holders? To what extent are limitations to copyright being eroded by amendments to existing laws that focus on digital content?"

   The annotated bibliography covers current changes to national legislation in a number of countries, changes that have been introduced as a result of efforts to implement provisions of transnational agreements to which the respective countries are signatories, most notably the World Intellectual Property Organization Copyright Treaty (WIPO WCT). The provisions of this treaty include copyright protection for software as well as digital works and introduce criminal penalties for infringement, which ranges from unauthorized copying of material placed on a Web site to the removal or alteration of rights management controls from digital works. The newly introduced restrictions on re-use are not balanced by adequate exemptions or protections that enable records preservation activities. This precarious balance is further complicated by the trend that sees terms of copyright coverage being extended in most countries by the addition of years, or scope of coverage, or both.

2. A study of policy on privacy and freedom of information policies,[9] which examined the challenge brought to record integrity and authenticity by privacy protection.

   The study was supported by two principal scholarly papers: a comparative regulatory study of Canada, the United States, Australia and the European Union and a second, more

[8] Gutman et al., "Copyright Policy Annotated Bibliography," op. cit.

[9] Malcolm Todd (2005), "InterPARES 2 Project - Policy Cross-domain: Information Policy - Privacy Report." Available at http://www.interpares.org/display_file.cfm?doc=ip2(policy)privacy_report.pdf. The main contributing papers are Livia Iacovino and Malcolm Todd (2007), "The Long-Term Preservation of Identifiable Personal Data: A Comparative Archival Perspective on Privacy Regulatory Models in the European Union, Australia, Canada and the United States," *Archival Science* 7(1): 107–127; and Malcolm Todd (2006), "Power, Identity, Integrity, Authenticity, and the Archives: A Comparative Study of the Application of Archival Methodologies to Contemporary Privacy," *Archivaria* 61 (Spring): 181–214.

theoretical discussion of the issue by triangulating multiple archival viewpoints. Both papers and the summary study propose detailed policy recommendations to promote the preservation of authentic digital records in a way compatible with privacy principles.

3. A study of general records-related legislation, including that enabling archival institutions, evidence acts, etc., from thirteen (13) selected jurisdictions, aimed at identifying commonalities affecting records preservation and potential barriers to preservation.[10]

   The study reviewed national and sub-national legislation as well as the regulatory environment of the European Union. The study examined how records were defined, assessed how comprehensively the records lifecycle was reflected in the rules and looked for consistency (or its lack) in multi-jurisdiction environments.

4. A study of record authenticity.[11]

   The study analyzed the juridical concepts embedded in evidence legislation in the North American, European and Chinese jurisdictions, compared them with the benchmark requirements issued by InterPARES 1[12] and evaluated the digital authentication requirements within the same systems.

5. A study of the potential contribution of open source software and open standards to the long-term preservation of digital records.[13]

   The study examined whether the acquisition policies and transfer procedures of a broad variety of archival institutions showed a coherent body of knowledge on the issues of file format selection generally and the use of open source and open standards specifically. Some highly developed open source policy material was observed in the science Focus data collected in association with case study 06 and general study 10. This is an example of highly specialized usage and high capital cost of unrepeatable data creation forcing the consideration of creation standards from the systems design stage and even in sectoral and funding policies.

The studies were presented at the InterPARES plenary workshop in Chicago one year later. The last study was spun off for completion into the Appraisal and Preservation Domain, Domain 3, as this was deemed a more appropriate research unit for the study. A second Policy panel presented the findings pertaining to the legislation studies at the Association of Canadian Archivists annual conference in June 2006.

---

[10] Jim Suderman, Fiorella Foscarini and Erin Coulter (2005), "InterPARES 2 Project - Archives Legislation Study Report."Available at http://www.interpares.org/display_file.cfm?doc= ip2(policy)_archives_legislation_report.pdf. Jurisdictions studied are Australia, Canada (including the provincial jurisdictions of Nova Scotia, Quebec, Manitoba, and British Columbia), China, the European Union, France, Hong Kong, Italy, Singapore and the United States. The underlying studies are available on the InterPARES 2 Web site at http://www.interpares.org/ip2/ip2_documents.cfm?cat=policy.

[11] Luciana Duranti (2005), "InterPARES 2 Project - Policy Cross-domain: Authenticity and Authentication in the Law." Available at http://www.interpares.org/display_file.cfm?doc=ip2(policy)authenticity-authentication_law.pdf. The underlying studies are available on the InterPARES 2 Web site at http://www.interpares.org/ip2/ip2_documents.cfm?cat=policy.

[12] Heather MacNeil et al., "Part One – Establishing and Maintaining Trust in Electronic Records: Authenticity Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 19–65. Online reprint available at http://www.interpares.org/book/interpares_book_d_part1.pdf.

[13] Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation." Available at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf (English); http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf (French).

Case Study Data

The case study data were incorporated into the research of the Policy Cross-domain in three stages. The first stage was a review of the responses contained in each of the case study reports to the four questions below:

20. To what extent do policies, procedures and standards currently control record creation, maintenance, preservation and use in the context of the creator's activity? Do these policies, procedures and standards need to be modified or augmented?

21. What legal, moral (e.g., control over artistic expression) or ethical obligations, concerns or issues exist regarding the creation, maintenance, preservation and use of the records in the context of the creator's activity?

22. What descriptive or other metadata schema or standards are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

23. What is the source of these descriptive or other metadata schema or standards (institutional convention, professional body, international standard, individual practice, etc.)?

The second stage involved another pass through the case studies with targeted explanatory prompts on intellectual property and privacy issues. It was executed by a small team of research assistants in late 2004, with the benefit of the policy studies then completed or nearing completion. The third stage is the composition of this report. This has involved a final review of data gathered as well as providing case study leaders with the opportunity to comment on the conclusions reached.

Except for this final stage, matching the fragmentary policy data coming from the case studies with those from the higher level studies was problematic. This can be partly attributed to the difficulty of interdisciplinary exchange between the perspectives of archival science and political science.

Policy Themes

The record creation environments that emerged from the case studies and the regulations for record creation, maintenance and preservation that emerged from the policy studies show patchy-to-nonexistent degrees of maturity. In terms of a comprehensive framework, few organizations and legislative jurisdictions show that they can deal adequately with the digital challenge, particularly where management requirements at any time in the records lifecycle involve complex multi-component records as appeared in many of the case studies. The main exception to this grim picture was, unsurprisingly, the sphere of evidence law: across a wide range of jurisdictions, legislation related to the use of records as evidence in a court of law shows a considerable congruity with the findings of the Authenticity Task Force of InterPARES 1.

Answers to the original research questions of the Policy Cross-domain follow in the next section. The four statements reflect the principle policy themes that emerged from the research done by the Policy Team, by other groups within the Project and from the case study reports.

Theme 1: An inclusive policy infrastructure for recordkeeping is required to support the activities of a society heavily reliant on information technology.[14]

A principal finding of InterPARES 2 is that preservation of records emerging from interactive, dynamic and experiential environments requires an inclusive policy infrastructure beyond the principles expressed by the InterPARES 1 Strategy Task Force. Concerns of intellectual property, privacy and security pre-exist the digital recordkeeping environment, but in relation to a minority of records. These concerns are now far more prevalent. The Policy Team has presented the necessary elements of the top level of such an infrastructure in the Framework of Principles for the Development of Policies, Strategies and Standards for the Long-Term Preservation of Digital Records. These are capable of implementation at a variety of levels of governance.

In the current networked/inter-connected environment, the following concerns become central because of the increasing transfer of information across organizational boundaries. While the first three concerns are external to records and traditionally provide the basis for archival preservation, the intellectual property rights concern applies to the record both externally (to its context) as well as internally (to its content); the last two concerns are internal to the record.

- Relationship to business process
- Relationship to specific transaction
- Relationship to creator
- Relationship to intellectual property rights (context and content)
- Relationship to privacy (content)
- Relationship to security (content)

Acceptance of a new conceptual understanding of the nature of the record that is extensible to these new environments and its use in tandem with the related policy principles will encourage the commonality of approach required to turn InterPARES' theoretical outputs into a robust foundation for formal standards development and policy creation at organizational, sectoral, national and international levels. Like any policy principles and any juridical instruments designed to support them, standards need to be facilitative and not specific to any particular technology to be useful.

A research project such as InterPARES has to define best or even ideal practice based on clearly articulated theoretical principles. Standardization in the national or international arena tends to focus on either setting acceptable baseline requirements or formalizing commonly accepted "best" practice as a norm. The proliferation of computing has tended to nudge standard-setting towards the second whereas arguably it ought to be confined to the first. This is particularly true in the area of promoting interoperability—across time and space—between digital systems, which is vital to support information transfers and consequently the archival process. Aside from the study of file formats already cited as spun off into Domain 3, the Policy

[14] In the InterPARES 2 Terminology Database, "recordkeeping" is defined as "The whole of the principles, policies, rules and strategies employed by the creator that establishes and maintains administrative, intellectual and physical control on its records," and "recordkeeping system" is defined as "A set of rules governing the storage, use, maintenance and disposition of records and/or information about records, and the tools and mechanisms used to implement these rules" (http://www.interpares.org/ip2/ip2_terminology_db2.cfm). These concepts are reflected in the InterPARES 2 Chain of Preservation (COP) Model (see the Modeling Cross-domain Task Force Report, available at http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf).

Cross-domain has not directly addressed the development of standards. However, other research units within the Project have observed and collaborated with standards-setting initiatives.[15]

Theme 2: An expanded and more detailed definition of record is necessary.

InterPARES 2 findings recommend preservation of all documents the creator treats as records; that is, all documents that the creator relies upon in the usual and ordinary course of affairs, associates with other records and refers to as the records of its affairs. This is more consistent with the inclusive definition of the term "record" used in statutes. It is the creator's judgement of what constitutes the record to be kept for action and reference, and the preserver has then to assess the feasibility of preserving it over the long term.

InterPARES 2 findings also point to a new category of records: potential records. Records have traditionally been identified as such retrospectively; that is, after having been completed and issued with a fixed form and stable content; but, with dynamic systems, there is the possibility of identifying "prospective" records. The digital objects that clearly manifest themselves as records since the moment they are created fulfil the traditional, memorial function of records to bear witness to or remember an action in which they participated or of which they were the residue. Rather than witnessing the past, prospective records guide the future through a set of instructions or actions to be carried out.[16] As such, prospective records may not be considered records when their process of development begins, but, since their content can be fixed and their documentary form and functionalities described to make it possible to re-create them in the future, they could become records. Establishing policies to manage recordkeeping for digital objects that are prospective records and *may* become records appears to fall into the context of guides, manuals and other directive or procedural documents.

[15] Throughout the duration of InterPARES 2, there have been significant developments in the digital longevity standards arena, particularly open standards. ISO 19005 (see International Organization for Standardization, ISO 19005-1:2005 - Document management—Electronic document file format for long-term preservation—Part 1: Use of PDF 1.4 (PDF/A-1)) is a file format specification derived from the PDF Reference, Third Edition, version 1.4 of Adobe Systems Incorporated's commercial software *Acrobat* (a matrix image format with some textual support capability). This is an encouraging example of a proprietary software format specification becoming an openly available specification once the owners of the intellectual property have replaced it with another format for their main revenue-generating markets. In this case, the intellectual property is to be managed by ISO for fifty years. In late 2006, Microsoft Corporation announced that future versions of its Microsoft Office System software, beginning with the 2007 version, will support saving documents in its XML-encoded "Office Open XML" (abbreviated as OOXML), which is a file format specification created by Microsoft for the storage of digital documents. The format was standardized by Ecma (*European Computer Manufacturers Association)* International as Ecma 376 in December 2006, which has since been submitted for adoption under the ISO/IEC JTC 1 process. An important distinction should be drawn from the archival perspective between widely adopted "industry" standards and those that are genuinely open: dependencies on the current computing environment may exist for both record content and metadata at encoding/syntactic, computer file, application and database levels as well as computer hardware. There has been formal collaboration between researchers in the Description Cross-domain developing the Metadata and Archival Description Registry and Analysis System (MADRAS) (see http://www.gseis.ucla.edu/us-interpares/madras/guidelines.php) and the working group within ISO Technical Committee 46/Sub-Committee 11 drafting the third part of ISO 23081 - Information and documentation—Records management processes—Metadata for records. MADRAS is a tool that is designed to increase the visibility of recordkeeping and archival metadata schemas and to facilitate the comparison of schemas against established requirements. Similarly, in the fall of 2006, the Project made a submission to the revision of ISO 14721, the *Open Archives Information System Reference Model*. Within the case studies, the most comprehensive and policy-driven observance of record creation standards was found to be in the science focus and especially case studies 06 and 19. Many of the government focus case study reports mention standards in response to direct case study questions, but they are either substantially irrelevant to recordkeeping and preservation requirements or not actually implemented (e.g., ISO 15836:2003 - *Information and documentation—The Dublin Core metadata element set* was frequently cited).

[16] Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 13–68 (Note: a reprint of this article is provided in Appendix 2. Available at http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_02.pdf).

Theme 3: Business processes are divided between many systems.

Deployment of dynamic, interactive and experiential systems to capture, handle and manage data is at present not always undertaken with due consideration of the various roles of records (as memorials; i.e., for reference, or as directions; i.e., instructions for future activities), nor of their sometimes distributed nature. Systems may be distributed across an organization, which may itself be distributed (e.g., multi-national organizations that cross national boundaries). Both of these scenarios are the norm in collaborative e-Science projects such as the Cybercartographic Atlas of Antarctica examined in case study 06.

Dynamic, interactive and experiential systems may also be deployed to achieve objectives that are not compatible with those of recordkeeping (e.g., providing a "window" on existing data at a particular point in time, as in the VanMap case study). Documents that may satisfy recordkeeping requirements can be instantiated at multiple points across modern systems. The Revenue On-Line Service (ROS), examined in case study 20, was essentially a conduit, enabling controlled input of data directly by citizens rather than by government staff working from paper forms mailed in by citizens. The digital objects of the ROS are records meant to establish and normalize a citizen's relationship with the revenue agency. That is, the business of citizens paying taxes was broken out into at least two systems: the ROS, which managed the relationship of the citizen with the revenue agency, and the mainframe computers, which actually assessed the taxes. In addition, documentary elements that convey the semantic of a record (metadata schemas, for instance), may exist in systems as digital objects separate from the record. In case study 19, an engineering experiment that used Web Ontology Language (OWL), an extension of XML that allows representation of semantics within metadata schemas to formulate a new logical preservation format for complex CAD records, metadata elements were stored in a segment of a pilot preservation system located on the opposite end of a national network shared by the experiment partners.[17]

The sub-division of a business process between systems, some of which may (a) be dynamic, interactive, or experiential and (b) exist across organizational boundaries, suggests a need for policy direction as comprehensive as the systems and business processes at hand. Records identified in one system must be considered along with records related to the same business process created by other system(s) to ensure the most effective management, disposition and preservation of records takes place. Policy should ensure that: (1) the identification of documentary entities, including but not limited to records/metadata/linkages, etc.,[18] is undertaken at the system design phase, (2) appropriate functions are incorporated to manage and preserve the entities identified at the outset of system development and (3) the process and outcomes of these activities are reviewed regularly as part of system operations.

[17] See Kenneth Hawkins (2006), "InterPARES 2 Project - Case Study 19 Diplomatic Analysis: Preservation and Authentication of Electronic Engineering and Manufacturing Records." Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_diplomatic_analysis.pdf; and Kenneth Hawkins (2006), "InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records," 14, 18. Available at http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf.

[18] Because semantic value can be derived from an understanding of how documentary entities relate to one another (for example, a registry to a series of records and the records themselves), additional entities of interest might include data and system models, domain-specific taxonomies and enterprise architecture models and specifications.

Theme 4: Preservation policies are inadequate or absent.

The digital objects considered to be records by their creators may not be preservable because of poorly considered use of digital signatures, for example, or may otherwise not be fit for functioning as retrospective or prospective records. The preservation activities examined by the case studies were mostly directed at keeping the data, not the record. For example, back-up and disaster recovery routines were found to be widespread, but the ability to restore the records except to an identical system (interoperability across time) was rarely addressed. In science Focus case study 06, the Cybercartographic Atlas of Antarctica, steps taken to ensure interoperability across systems performed many of the same purposes as preservation. Similarly, production of exhibits that are accessed via the Web may include elements of a preservation policy, albeit not a long-term one, because the produced objects are considered to be maintained accurate and authentic across space (i.e., from one system to another). Consideration of what constitutes a preservation policy needs to be inclusive, as shown by the Chain of Preservation Model.[19]

Addressing the Research Questions

The research statement of the Policy Cross-domain contains the following questions, to which the report provides a combined response to avoid repetition.

- To what extent do policies, procedures and standards currently control record creation, maintenance, preservation and use in each focus area? Do these policies, procedures and standards need to be modified or augmented?
- Can an intellectual framework or frameworks be developed to facilitate the translation of policies, procedures and standards into different national environments, sectors and domains?
- How can enhanced control over and standardization of record creation, maintenance, preservation, access and use be balanced against cultural and juridical differences and perspectives on issues such as freedom of expression, moral rights, privacy and national security?
- What legal or moral obligations exist regarding the creation, maintenance, preservation and use of the records of artistic and scientific activities?
- What principles should guide the formulation of policies, strategies and standards related to the creation of reliable, accurate and authentic records in the digital environments under investigation?
- What principles should guide the formulation of policies, strategies and standards related to the appraisal of those records?
- What principles should guide the formulation of policies, strategies and standards related to the long-term preservation of those records?
- What should be the criteria for developing national policies, strategies and standards?
- What should be the criteria for developing organizational policies, strategies and standards?

[19] See the Modeling Cross-domain Task Force Report, op. cit.

Recordkeeping and the current policy environment

The degree to which policies, procedures and standards control record creation, maintenance, preservation and use in the three focus areas examined varies from none at all to partial. It appears that the two key factors affecting the response to this research question are the nature of the organization and the phase or stage of the records lifecycle (i.e., record creation, maintenance, etc.), being considered.

Where organizational culture is conducive to the development of policies and procedures and the adoption of standards, there are controlled aspects of record creation, maintenance, preservation and use. Of the environments studied, government and some scientific organizations developed or adopted policies and procedures, while individual artists or small, temporary partnerships did not. In these latter organizations, the extent of control on recordkeeping through policies, procedures and the adoption of standards is effectively nil.

In half of the arts focus case studies, it was found that there was no consideration given to record maintenance and preservation. Among the remaining case studies, where there was some consideration given, there was no common motivating factor. In some instances, record maintenance and preservation were motivated by legal reasons, usually pertaining to the protection of intellectual property. Publicity and future performances were two other motives for record maintenance and preservation, although this did not necessarily extend to the development of policies or adoption of standards.[20]

Among the science focus case studies, it was noted that the development of rules and procedures around record creation and maintenance is driven by the immediate and foreseeable requirements of each scientific activity. It was also found that while sophisticated technologies were often adopted, those used to maintain and access the records tended to be rudimentary (e.g., Microsoft Windows or other proprietary software tools). A process to determine how long to keep project data was consistently in place. Procedures supporting retention included duplication and migration. While there was no consistent approach or procedure to achieve the determined retention requirement, enabling others to access the data on different systems (i.e., interoperability), was frequently a guiding consideration.

Of the three categories of case studies, those involving government organizations consistently had the most comprehensive recordkeeping policies and procedures in terms of all phases of the lifecycle. In most cases, organizations had an existing, formal relationship with an archives or other unit responsible for record preservation. As with the science focus case studies, however, it was found that maintenance and preservation processes were data- or system-oriented and not necessarily linked to the organization's specific recordkeeping requirements.

Another factor to consider in relation to where an expressed policy exists or standards are adopted is the phase of the records lifecycle. The case studies showed that organizations may adhere to policies, procedures and standards in one stage of the lifecycle (e.g., record creation), but not in others. The InterPARES concept of records preservation is comprehensive and includes all activities that affect the record since its creation. None of the organizations involved with the case studies displayed such a comprehensive approach.[21]

[20] See the Domain 1 Task Force Report, op. cit.

[21] The conclusion that preservation of digital records must begin at the creation stage has been reached by most, perhaps all, research in this field. This conclusion is thoroughly developed in InterPARES 1 and in that Project's strategic principles as follows: "…preservation of authentic electronic records is a continuous process that begins with the process of records creation..." (Duranti et al., "Strategy Task Force Report," op. cit., 4).

A study of records-related legislation concluded that laws in the jurisdictions studied are very inclusive in their definitions of records, in contrast with the much more specific archival definition adopted by the InterPARES 2 Project:

> 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.[22]

Inclusive and inconsistent definitions of record undermine not only an organization's ability to develop the policies and procedures it needs, but also decisions to adopt existing or proposed standards. They also compromise an organization's ability to correctly interpret precedents set in court decisions regarding records.[23]

The study of records-related legislation concluded that while all phases of the records lifecycle are addressed in the laws or directives of the jurisdictions studied, they are not addressed comprehensively within any single law, nor overall within the body of legislation examined within any one jurisdiction.[24] For example, land transactions are a highly regulated business activity. In Alsace-Moselle (France), case study 18, information technologies were adopted to carry out this activity. The system developed is very effective for the short to medium term, but presents unresolved long-term maintenance and preservation issues, especially with regard to maintaining the authentication function of the judge's digital signature over the long term. The same study also concluded that the records lifecycle phases most commonly addressed within legislation are those of creation and disposition.

Statutory recordkeeping requirements provide a strong impetus for organizations governed by them, but it cannot be expected that legislation will consistently and comprehensively address all phases of the records lifecycle. In the absence of a comprehensive guidance or direction from law, organizations may be willing to adopt general standards, such as the records management standard,[25] to help them effectively maintain and preserve their records.

Balancing cultural differences against a common approach

The widespread adoption of new technology in the three environments examined by InterPARES 2 does not appear to have necessarily or fundamentally changed the long-standing processes in those environments. The Domain 1 Task Force Report on record creation concludes that the processes occurring in the creation of records today are recognizable in those used in the pre-digital environment.

> What is being witnessed in the arts focus case studies is the *continuation* of the artistic tradition in the digital environment. The processes are largely the same, based on the long-established artistic principles of each field… [for most of] the science focus case studies, document creation takes place in a much more formalized and controlled environment, with pre-determined processes including the collection, analysis and preservation or communication of data…Most of the case studies in the government focus deal with a traditional activity being carried out in a new way. Therefore, the process of document creation is largely the same as for the traditional environment; it is simply transposed into the digital

[22] Definition for "record" from the InterPARES 2 Terminology Database, op. cit.
[23] Examples exist where organizations have been sued for large amounts while other, similar organizations continue to ignore the risk.
[24] Suderman et al., "Archives Legislation Study Report," op. cit., 24.
[25] See International Organization for Standardization, ISO 15489-1:2001 - Information and documentation—Records management—Part 1: General.

> environment with the possible addition of certain steps in the process to take the technology into account.[26]

If the processes are not changing at a fundamental level, then existing policies, strategies and standards may not need to be completely changed, but simply revised and extended. In the pre-digital environment, where records and physical media are inseparable, the point at which a record is created is well established (e.g., a film is made or a letter is written). InterPARES 1 concluded that in the digital environment the medium is no longer an essential part of a record. As a consequence, preservation must be directed to preserving the ability to reproduce digital records, moving them, as needed, from one medium to another. Therefore, record creation procedures and standards must set out when a record has been created as well as identify the intellectual and digital components comprising the record and their relationships to each other. For records created or existing in dynamic systems, procedures must outline how those components are determined and set out the acceptable range of variations on their relationships to each other.[27]

New principles that guide policies, procedures and standards on the identification and modification of created records are also required. Digital technologies have dramatically increased the opportunity to integrate record types formerly distinguished by their media (e.g., audio and text). Besides the well-known complications linked to maintaining and/or preserving the differently formatted digital components of which records are comprised, this capacity requires a significantly enhanced management of intellectual property rights existing within the records. These may include database rights, copyrights and patents. Similarly, the emergence of access and privacy legislation requires a more comprehensive management of record content than existed before. Laws governing personal information emphasize accuracy and enable a person identified within the record to request that information within a record be corrected. Rights inimical to the preservation of the records by a preserver may subsist at the record component level. Whereas previously the preserver could manage these issues by considering "sunset" periods for which entire records might be withheld until the rights had expired, they now need addressing in policies and corresponding rules applying from creation. Records creators must also have clear procedures in place for how those rules are implemented. These procedures must be explicitly understood by the individuals responsible or be built into the design of systems that maintain the records.

Digital technology has also dramatically enhanced the means to transmit information. In both the arts focus and science focus case studies, this was found to be a welcome characteristic. For the artist, the ease of transmission can dramatically increase the potential audience for a created work. For scientists, greater access to more data supports more effective research. The scientific community is motivated by "the desire and possibility of translating the collected data into a neutral or open source format."[28] By contrast, security concerns predominate in the governmental environment, where record transmission is emphasizing the need for security metadata and technologies to support legal non-repudiation by participants in the record creation process, as well as standards for secure storage technologies, such as encryption, secure digital signatures and biometrics. Governments also exchange information, of course. For records to be correctly accessed across space, explicit policies are required not only so that the receiver of the transmitted record can accurately reproduce it, but also so that records sent do not contravene

[26] Domain 1 Task Force Report, op. cit., 31, 33, 34 (emphasis in original).
[27] See Duranti and Thibodeau, "The Concept of Record," op. cit.
[28] Domain 1 Task Force Report, op. cit., 33.

requirements for security, privacy protection and intellectual property in either jurisdiction. There are three generic approaches to achieving this greater communication with appropriate safeguards: (1) harmonizing juridical frameworks, (2) implementing effective exemptions for the purposes of archival preservation and (3) ensuring comprehensive rights metadata accompany the record.

InterPARES 1 emphasized the importance to the preserver of assessing the feasibility of preservation during the records appraisal process. Feasibility assessment policies and procedures need to be guided by the technological requirements of the records as they relate to the capabilities of the preserver's preservation system. They must also take into account what residual rights or obligations—privacy, intellectual property, security, etc.—will have to be managed or administered by the preserver. In this way feasibility operates both specifically (i.e., in relation to an identified body of records), and generally, in that the preserver must develop or modify acquisition policies so that they are consistent with the capabilities of the preservation system.

The preserver who is maintaining authentic copies of created digital records must, in effect, be guided by the same concerns as the creator. That is, if the creator had to observe requirements of privacy, intellectual property and security while maintaining the records, the preserver must also observe those requirements within the preservation environment, unless explicitly exempted. The foremost principle that must guide the long-term preservation of digital records was established in InterPARES 1, which is to ensure that through preservation processes records remain authentic copies of the creator's records.

The literature reviewed in the annotated bibliography of intellectual property rights points to the issue of enhanced control over access to digital content in the service of commerce as a key one at play in the formulation of international treaties, national legislation, case law and policy debates. Ironically, the very same technical and legislative features that enable economic protection for rights owners and enhance the immediate access to records and information for consumers also make more restrictive the future uses of the content and ultimately impede the ability to preserve these for their "second noncommercial life."[29] The emergence of access and redistribution control technologies (also known as digital rights management or DRM) and attendant debates about these technologies and the challenges they introduce for preservation demonstrate well the precarious balance struck between enhanced control over access, communities' expectations and juridical perspectives on use.[30]

Reflecting legal and moral obligations in policy

Activities of records creators and records preservers are subject to legal and moral obligations as well as community expectations. Records preservers are "downstream" recipients of evidence of the activities of records creators of yesterday and today. At the same time, records preservers are not only recipients of records but also records creators in their own right. Where in the past records preservers, especially archives, managed the transfer of physical and intellectual property rights in analogue records, the picture is quite different today. As records creators use

[29] "Second noncommercial life" has been elaborated by legal scholar Lawrence Lessig as the period that begins when the copyright term expires and content becomes subject to re-use; see Lessig's *Free Culture* (The Penguin Press, 2004), and the Editorial, "The Coming of Copyright Perpetuity," *New York Times*, January 16, 2003, p. A28.

[30] For instance, see the Canadian Internet Policy Clinic (University of Ottawa) policy debates, available at http://www.cippic.ca/en/faqs-resources/digital-rights-management/, and "Digital Rights Issues" in the American Library Association, Washington Policy Office, available at http://www.ala.org/ala/washoff/WOissues/copyrightb/digitalrights/DRMissues.pdf.

software to create and/or apply rights management technologies to wrap or otherwise protect intellectual assets, they introduce a whole new layer for preservation management. Because this additional layer is itself subject to intellectual property rights and protections, the process of preservation takes on additional tasks and risks.

The emergence of access and redistribution control technologies comes at a time when moral rights are being trumped by commercial rights, and privacy rights are being overwritten by assertions of national security. Successive changes to national laws, international trade agreements and business models render the already considerable challenge of preserving digital records far more complex than simply overcoming issues of technological obsolescence.

Beyond ensuring that preserved digital records remain authentic copies of the creator's records, preservation should be seen and undertaken as a process that is compatible with the purpose for which the preserved records were created. If compatibility of purpose cannot be established, preservers may require specific or general exemptions from liability under intellectual property requirements, including moral rights, and privacy requirements. For example, anonymization of records containing personal information compromises the integrity of the created record. Preservation activities may result in changes to the records at the bit level, but not at a functional level. Such activities would contravene a rigid application of intellectual property rights.

The records created in some of the arts focus case studies were insufficiently well-defined for a preserver to demonstrate the authenticity of reproduced copies of the records. In *Obsessed Again...*, case study 13, a reproduction of the work was deemed not to be authentic by the creator. This suggests the importance of the principle that preservers must interact with creators from the outset to preserve authentic copies of records. Where that relationship does not exist, the preserver's procedures and standards must set out the extent of the authenticity of the reproductions of preserved copies.

Principles for appraisal and preservation

The combination of rapid technological change and the need to manage rights subsisting within the content or components of the records poses significant challenges to the long-term preservation of digital records. The development and adoption of common standards and stronger procedural controls for recordkeeping cannot by itself enable long-term preservation. A clear and ongoing relationship between creator and preserver is also necessary. Each aspect sustains the other. Standards and procedural controls inform the selection of record creation and maintenance technologies by the creator. The development of guidelines and procedures and the adoption of standards would comprise the principal aspects of the preserver's participation at the record creation phase. Established procedures and standards also help the preserver to develop and operate a preservation system and, much more importantly, demonstrate the authenticity of records maintained in that system. The European Union

> consider[s] standardization "an integral part of their policies to carry out 'better regulation', to increase competitiveness of enterprises and to remove barriers to trade at international level."[91] The directives on *Data Protection*, *Electronic Signature*, *e-Invoicing* and the regulatory framework for electronic communications networks and services (which consists of five additional directives) are a set of new legislation (categorized as Information Society legislation). They are issued under the aegis of the European Standards

> Organizations with the purpose of establishing a "legal framework to ensure the free movement of information society services between Member States."[92 31]

Standards and procedural controls are static in relation to the creativity of record users and deployment of new technologies and systems and so, by themselves, cannot accommodate national and cultural differences.[32] A sustained relationship between records creator and preserver is a means by which the creator can communicate innovative uses or procedural variations to the preserver. Such a relationship also informs the preserver of ethical behaviour of a community (e.g., that in the scientific research community, research data are to be shared as broadly as possible, but not until those who have prepared or gathered the data have had a reasonable opportunity to publish their findings).

The extent to which public, sectoral and organizational policy affects the participants in various record creating activities varies according to the legal, ethical and moral dimensions of their relationships with their correspondents. Thus, there are fewer legal and moral obligations affecting recordkeeping in the artistic and scientific environments in comparison with obligations existing in the governmental environment. Those in the artistic environment might be summarized as relating to intellectual property, while obligations in the scientific environment would centre more on accuracy and accessibility of research data.[33]

Among the arts focus case studies, some creators simply were unconcerned with the long-term risk of loss of their digital records. As noted above, even where the creating organizations were concerned with recordkeeping, obligations related to protecting or acknowledging intellectual property or meeting financial accountability requirements to a granting body.

Legal and moral obligations for recordkeeping may be increasing within the scientific community, driven primarily by policies of funding organizations. As a knowledge-based community, it is in its own interest to ensure that research data be maintained for future use. Some scientific communities have long-established recordkeeping standards (e.g., metadata standards for geospatial data), to which the community expects researchers to adhere. Likewise there is widespread use of creative commons licenses to support general access to and use of scientific research data while simultaneously establishing the ownership rights of the creator.

Recent legislation governing personal information has extended obligations in this regard beyond the governmental sector and highly regulated private sector activities, such as banking. This legislation imposes some additional obligations on any organization collecting and using personal information. Research communities already have rules in place for the ethical collection, use and maintenance of research data involving human subjects, so it is safe to say that legal obligations are increasing for records containing personal information in the scientific environment. It is unclear how this new legislation will affect records in the artistic environment.

Intellectual property laws pose a dilemma for records preservers where such rights subsisting in the preserver's custody may be aggressively protected. The moral obligation not to change the creation of an artist, for example, may make long-term preservation impossible if the created record relies on short-lived technological components. In some jurisdictions, laws have exempted specific preservation institutions, usually national archives, from liability arising from

[31] Suderman et al., "Archives Legislation Study Report," op. cit., 30. Note: footnote references in the quote are from the original text, and are not reproduced here.

[32] Case studies used in both phases of InterPARES were from many different jurisdictions. For specific jurisdictional studies on specific issues, such as authenticity, see the Policy Cross-domain studies summarized above in the section entitled "Research Methodology."

[33] The link to the recordkeeping practices of individuals is in their relationships with organizations, and InterPARES guidelines for their own recordkeeping and preservation are referred to in the *Policy Framework* (see Appendix 19).

copyright (e.g., Library and Archives Canada is explicitly allowed to "crawl" and capture Canadian Web content). In terms of personal information, preserving organizations or their clients may need to prove to an external authority that their use of records containing personal information is compatible with the purpose for which it was created.

Importance of a common basis for national policies

National policies, strategies and standards should be guided by common approaches and common purposes for all phases of recordkeeping. The directives developed in the European Union (EU) are an example of how common criteria can be set for specific implementation within each member nation. Criteria developed in this way need to be reviewed or measured against technological and economic limitations. While the general principles of the EU's e-Signature Directive are being implemented by member states "despite recognized limitations in the technology supporting e-signatures [it is observed that] 'there is currently no market demand for qualified certificates and related services.'"[34] The limitations of technology are reflected in the caution shown in the EU's e-Signature Directive, which "explicitly excludes certain categories of contracts."[35]

The relationships between records creators and preservers must be acknowledged and supported by national policies, strategies and standards. This will require rules for recognizing professionals and organizations that preserve digital records throughout the society, not just in terms of national institutions. Such rules will need to address the obligations preservers must meet; that is, make explicit the characteristics of a trusted custodian, in connection with the rights subsisting within the records being preserved and the transient nature of technologies used for recordkeeping, particularly those for record creation. National policies and standards must also be flexible enough to accommodate the norms of specialized communities, such as the creative arts and scientific research, which may themselves not be particularly bounded by national borders.

While it is recognized that national policies and standards need not be comprehensively addressed in legislation, it is important that legislation be developed within the most comprehensive information strategy possible. Establishing policies for which consistent implementation is impossible, as in the case of the European Union's e-Signature Directive, jeopardizes the rights of all. Without authentic, reliable and accurate records and rules about their use and transmission, rights such as those pertaining to privacy or intellectual property of citizens may be violated. Without record creation, maintenance and preservation policies in place, the state may itself participate in the violation of those rights.

Criteria for organizational policies

Organizational policies, strategies and standards for recordkeeping must obviously meet legislated requirements. To the greatest degree possible, organizations or communities of practice must codify how these requirements will be met. Where legal obligations subsist within records (e.g., privacy or intellectual property) and where these would be contravened by normal maintenance or preservation activities, organizational recordkeeping policies should incorporate a risk assessment component. Organizations may protect themselves at least to some degree by working collectively with similar organizations to establish common practices. Such an approach

[34] Suderman et al., "Archives Legislation Study Report," op. cit., 31.
[35] Ibid., 31–32.

will necessarily involve consideration of all phases of recordkeeping and all organizations within the community participating in any of those phases.

Another criterion for organizational recordkeeping policies, strategies and standards is the explicit consideration of long-term preservation requirements. This is essential not only to determine what those are, but also to determine whether preservation is even possible or desirable within the organization or whether an external preserver must be identified. In the latter case, the presence of long-term preservation requirements will form the basis of the relationship between the creating and preserving organizations.

Two sets of guidelines were produced by InterPARES 2 to assist individuals and organizations with establishing recordkeeping policies, strategies and standards. The first set of these is entitled *Creator Guidelines—Making and Maintaining Digital Materials: Guidelines for Individuals*,[36] and is intended to help individuals or small organizations who are making and maintaining digital materials, including records. As the case studies revealed, "the technology used by innovators and early adopters, regardless of the focus area in which they belonged, was proprietary and frequently customized," and that "[i]n many cases, the point of the work of these types of creators is to explore, test and push the limits of the available technology, be it hardware or software."[37] These guidelines are intended to inform creators who may not consider or be aware of digital record creation and maintenance concerns. In particular, evidence of authorship, with implications for preservation of intellectual property rights, is at risk even if technological obsolescence issues are addressed.

The second set of guidelines, *Preserver Guidelines—Preserving Digital Records: Guidelines for Organizations*,[38] provides more procedural guidance for any organization charged with providing preservation services (i.e., where preservation considerations are central to the organization). These guidelines are not specific to large, established archival organizations. They support the development of preservation procedures and systems that can maintain the accuracy and authenticity of the preserved records and are conceptually linked to the components described in the InterPARES 2 Chain of Preservation model.

Toward an Intellectual Framework for Policy Development

It is clear from the foregoing that not only can more comprehensive policies, procedures and standards be developed, but also that they are needed. Several of the products developed by InterPARES 2 contribute comprehensive guidance for all aspects of digital recordkeeping. In terms of policy, the main product of the Policy Cross-domain is the already mentioned Framework of Principles, comprising two complementary sets of principles for the creation and preservation of digital records. These principles are introduced and detailed in Appendix 19. The Framework provides the scope for developing a consistent and comprehensive policy environment in different jurisdictions, sectors and organizations. It may also help with the assessment of standards, existing and contemplated, and with the development of new standards, in terms of their applicability and utility for all aspects of recordkeeping.

The Framework extends the strategic principles established in InterPARES 1 in three key ways.[39] Firstly and most importantly, it sets out principles for record creation from the creator's

[36] See Appendix 20, op. cit.
[37] Domain 3 Task Force Report, op. cit., 18–19.
[38] See Appendix 21, op. cit.
[39] See Duranti et al., "Strategy Task Force Report," op. cit.

point of view. Secondly, because of the Framework's dual viewpoints (creator and preserver), it structures the relationship between a records creator and preserver. This relationship is seen as one of gradually transferring responsibility from the creator to the preserver.[40] While the idea of a shifting responsibility is not new, the principles clarify the dynamics of a seamless transfer. Thirdly, the Framework considers records emerging from three different environments (i.e., the arts, the sciences and government/administration) and which exist in dynamic, experiential or interactive systems. The scope of its application in terms of organization and system is thus much more inclusive than the strategic principles from InterPARES 1, which were based on governmental organizations where the most comprehensive recordkeeping policy environments existed.

[40] The theory of movable responsibility was developed decades ago in the United Kingdom by Felix Hull. It recognized that the records manager and archivist worked together throughout the records lifecycle but with the responsibilities of the former gradually diminishing as those of the latter grew.

Appendix 19

A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records[1]

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[2] and its designated preserver[3] must begin at the time the records are created.[4]

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;[5] 2) it is not possible to preserve digital records but only the ability to reproduce them;[6] and 3) the preserver 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

[1] 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 bitstreams 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.

[2] 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.

[3] 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.

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

[5] 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.

[6] 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.

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 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.

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"[7] 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, regardless of jurisdiction. This document, in combination with other products of the Project, especially the Chain of Preservation (COP) model,[8] 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.[9] 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

[7] 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.

[8] The COP model is available in Appendix 14 (http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_14.pdf) 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, op. cit.

[9] The initial draft of the principles relied heavily on the contributions of three research assistants: Fiorella Foscarini, Emily O'Neill and Sherry Xie.

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,"[10] 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 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,[11] fixed form does not imply that the bitstreams 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.[12] 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 consideration of the fact that what causes these variations also limits

[10] See InterPARES 2 Terminology Database, op. cit.

[11] 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).

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

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 bitstream and for presenting it for use over time. The bitstream can be altered, as a result of conversion for example, as long as it continues to be able to fulfil 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 fulfil, 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)[13] and the trusted records officer[14] or the recordkeeping office,[15] identifying access privileges[16] and access restrictions[17] 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.[18]

[13] 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.).
[14] 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.).
[15] Recordkeeping office is defined as "The office given the formal competence for designing, implementing and maintaining the creator's trusted recordkeeping system" (Ibid.).
[16] 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.).
[17] Access restrictions is defined as "The authority to read a record, granted to a person, position or office within an organization or agency" (Ibid.).
[18] See MacNeil et al., "Authenticity Task Force Report," op. cit.

(C5) A trusted record-making system should be used to generate records that can be presumed reliable.[19]

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 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

[19] There is no corresponding Preserver Principle.

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[20] (i.e., protecting and ensuring continuous access to) a creator's inactive records.[21] 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 fulfil 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,[22] 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.

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

[21] 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.).

[22] An active record is defined as "A record needed by the creator for the purpose of carrying out the actions for which it was created or for frequent reference" (Ibid). A semiactive record is defined as "A record that 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.).

(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 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.'"[23]
>
> 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 information that may be used to make copies on any kind of medium for all legal purposes."[24]
>
> 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.[25]

[23] 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.
[24] Italy, DPR 445/2000, art. 9, par. 1. Available at http://www.parlamento.it/parlam/leggi/deleghe/00443dla.htm.
[25] 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.

Principles for Records Preservers

(P1) A designated records preserver fulfils 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 fulfil 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 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*,[26] 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,

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

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 (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[27] for the record over time and/or the recordkeeping office and the designated preserver; the access privileges code[28] and the access restriction code;[29] and the list of annotations[30] and of format changes.[31]

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

[27] 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.).
[28] 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.).
[29] Access restriction code is defined as "The indication of the person, position or office authorized to read a record" (Ibid.).
[30] List of annotations is defined as "Recorded information about additions made to a record after it has been created" (Ibid.).
[31] 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.).

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 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.[32] 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 bitstream and for presenting it for use over time. The bitstream can be altered, as a result of conversion, for example, as long as it continues to be able to fulfil 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.

[32] 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.

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 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."[33] 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 bitstreams 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.

[33] 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 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,"[34] 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.[35]

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 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,

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

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.[36]

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*.[37]

(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

[36] 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.

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

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 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.