

Educating Records Professionals on Topics in Digital Preservation

Corinne Rogers
University of British Columbia

Luciana Duranti
University of British Columbia

1. Overview

Digital Records Pathways: Topics in Digital Preservation (hereafter *Digital Records Pathways* or ‘the modules’) is an educational initiative developed jointly by the International Council on Archives (ICA) Section for Archival Education (SAE) and the InterPARES project.¹ It contributes to the education and training of archivists and records professionals responsible for carrying out the preservation of authentic, reliable and usable digital records, based on the research findings of InterPARES. As InterPARES 1 and 2 demonstrated, the most effective records management environments are those guided by overarching statutory or normative requirements to maintain and preserve records, and an underlying framework of policies, procedures and guidelines for supporting and enforcing those requirements. *Digital Records Pathways* acknowledges this by treating topics in preservation as parts of a holistic system encompassing the entire records life-cycle that starts with sound policy, recognizes the role of organisational culture, and provides practical knowledge for specific aspects of digital preservation and digital preservation program development.

Digital Records Pathways consists of eight independent modules, supported by the ICA Multilingual Archival Terminology (MAT). The modules can be used as a set, or individually, for the purpose of covering the range of competencies required. Each module consists of theoretical and methodological knowledge and its practical application, illustrated through case studies and model scenarios. Several of the modules also contain templates for organisational use or for curriculum development by universities and professional associations to adapt and to develop specific training

¹Many people have contributed to the creation of these modules and should be expressly recognized and thanked. In particular, students in the PhD program at the University of British Columbia – Elizabeth Shaffer, Corinne Rogers, Donald Force, and Elaine Goh, who have drafted the contents, based on the work of InterPARES 1 and 2, and case studies conducted in InterPARES 3. ICA PCOM funded a Section on Archival Education (SAE) Workshop held in Rome in June 2011, led by Corinne Rogers and Elizabeth Shaffer, where SAE Steering committee members and other educators evaluated the proposed structure and content of the modules and contributed suggestions. Acknowledgment should also be made of the many Graduate Research Assistants who conducted the case studies, and therefore supported the development of these modules, and the many international researchers involved with InterPARES over the past fifteen years, under the direction of Luciana Duranti. Finally, thanks go to all who reviewed and commented on these modules, with special mention of InterPARES researchers: John McDonald, Information Management Consultant (modules 1, 2, 7, and 8), Jim Suderman, Director, Information Access from the Toronto’s City Clerk Office (module 3), Evelyn McLellan, Systems Archivist, Artefactual Systems Inc., and Paul Hebbard, Records Management Archivist, Simon Fraser University (module 6). This resource is freely available under a Creative Commons License also on the InterPARES web site and the Centre for the International Study of Contemporary Records and Archives (www.ciscra.org).

materials for students and professionals on digital recordkeeping and preservation issues. The modules have been developed by InterPARES Team Canada with the advice of the SAE educators who met in Rome in 2011 to discuss the first draft, and are illustrated with examples primarily from the Canadian context. However, the information in each module is transferrable to many different organisations and jurisdictions, and is customizable to specific domains or juridical contexts.

The modules in their current form can be self-administered by individuals, or offered through professional associations or workplace training. While not everyone who uses them will have the same knowledge base of archival concepts in general and digital records in particular, *Digital Records Pathways* assumes that the user has a solid grounding in the key concepts and methods of records management and archival science, and builds on that knowledge. To assist the users in determining the applicability of the modules to their situation, links to several self-assessment and organisational readiness tools are given in the first module.

In 2010, the SAE also embraced the challenge of creating an interactive, online, multilingual archival terminology database. This database, which can be found here <http://icarchives.webbler.co.uk/14282/ica-multilingual-archival-terminology/>, supports the education modules.

3. Objectives

The educational modules have the following objectives:

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

It is not the intention of these modules to be fully comprehensive for each topic. The depth of knowledge required, and the speed with which technology changes preclude such a goal. However, these modules offer a solid foundation based on the work of InterPARES, and a high level overview of other significant research initiatives.

4. Architecture of the set

Architecture of the set

1. A Framework for Digital Preservation			Foundational
2. Developing Policy and Procedures for Digital Preservation			
3. Organisational Culture	4. An Overview of Metadata	5. Appraisal Strategies	General purpose
6. E-mail	7. Websites	8. Cloud Computing	Specific purpose
9. International Terminology Database			Foundational

The first two modules offer information fundamental to any program of digital records preservation. They provide a foundation upon which the subsequent modules are built. The next three modules offer general information on topics common to digital preservation – the role of organisational culture, an overview of metadata, and an overview of appraisal in the context of managing records outside an electronic recordkeeping system (ERMS). The final three modules address specific (but by no means the only) topics of concern – the management of e-mail, preservation of records in web environments, and the issues arising from the increasing move to and reliance on cloud computing. These three modules provide the context for archivists and records professionals to understand the use of the technologies and their implications in the recordkeeping and record preservation environments. More importantly, the modules are toolkits for records professionals to ensure the trustworthiness of records over time.

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

- Overview of the topic and scope of the module;
- Learning objectives and expected level of knowledge upon completion;
- Methodology or the procedures to follow to apply the module;
- Templates (where appropriate) to facilitate the implementation of the module;
- Examples and Case Study(ies)/Scenarios (where appropriate) that provide real-world examples of module topic²
- Exercises covering key learning points
- Review questions to enhance comprehension and understanding of the topic
- Additional Resources for each module
 - Links to current research and online resources
 - Readings, standards and other templates for reference
 - Annotated bibliographies

Where appropriate, distinctions are drawn between the management and preservation activities involving active records and responsibilities for records that are no longer

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

required for business purposes, whether they are preserved by their creator or by a trusted third party.

5. Topics

5.1 Module 1: Introduction – A Framework for Digital Preservation

This module introduces the set as a whole. It explains how the modules are intended to be used, outlines objectives, and presents a summary of the contents of each module. It includes resources for institutional readiness and self-assessment to assist individuals and organisations in establishing their readiness or capacity for digital preservation. It introduces two complementary models for digital preservation: the InterPARES Chain of Preservation model, and the Open Archival Information System Reference Model (OAIS), and an annotated bibliography of digital preservation research and learning resources.

The Chain of Preservation (COP) Model (Duranti and Preston 2008) depicts and documents all the phases or stages in the lifecycle of digital records, and all the activities that must be undertaken to ensure that digital records are created reliable and accurate and maintained authentic over time. The model is relevant for records creators and records preservers, reflecting the understanding that the long-term preservation of authentic digital records comprises actions undertaken throughout the records' lifecycle. All activities pertaining to preservation are interdependent – omission of activities at any stage may imperil the reliability, accuracy, and authenticity of records (i.e. their identity and integrity) over time and across technological change. The COP model proceeds from an understanding of the concepts, methods and practices of archival science. The model balances constraints on record making, recordkeeping and record preservation, mechanisms that are available to creators and preservers for carrying out the related activities, and inputs into and outputs from the systems so controlled.

An approved ISO standard and considered the benchmark for digital preservation systems, the OAIS is a high-level model that defines the base functional components of a long-term preservation system and the key internal and external interfaces, and characterizes the information objects managed in the system. It addresses all aspects of long-term preservation of digital information: ingest, archival storage, data management, access, dissemination, and migration to new media and forms. The goal of an OAIS is to preserve information for a designated community over an indefinite period of time (CCSDS 2002). Digital preservation initiatives have adopted, adapted, or referenced the OAIS model since its inception as the foundation upon which to build, as has the A4_Preservation component of the COP model.

5.2 Module 2: Developing Policy and Procedures for Digital Preservation

A digital preservation policy, incorporated into an existing records management policy, or reflected in a stand-alone instrument, provides the framework for action and planning to ensure the long-term maintenance and preservation of an organisation's records. The capacity to preserve reliable, accurate and authentic digital records begins at the point of creation, so following a digital preservation policy through the records' active life will

facilitate preservation over the long-term for inactive records whether it is the creator or a trusted third party who is responsible for final disposition or preservation.

This module explains the purpose and benefits of a digital preservation policy, and offers a framework to create it. It provides a template for policy development based on the *InterPARES Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records* (Duranti and Preston 2008).

5.3 Module 3: Organisational Culture and its Effects on Records Management

Organisational culture is defined as a system of shared values and assumptions that may be explicit or implicit in nature. This system also involves an interaction of the values and assumptions of multiple stakeholders, who bring with them their own knowledge and experience. Organisational culture shapes and influences the manner in which individuals behave and act. It is also reflected in an organisation's goals, objectives, leadership style, system of incentives and disincentives, work processes, and the technology it adopts and uses.

Records professionals often overlook the impact of organisational culture on development and implementation (and success) of record systems. This module introduces five types of organisational culture that enable or constrain the adoption of a record making, recordkeeping or record preservation system, and explains the influencing factors in each. An organisational culture assessment checklist provides records professionals and other interested stakeholders with a template for analyzing the organisational culture in their workplace, and includes strategies on how to promote the management of records throughout their life cycle within each type of culture.

Additional resources listed guide the reader to some of the key literature on organisational culture generally, and within the archives and information management fields specifically.

5.4 Module 4: An Overview of Metadata

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

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

Appraisal of digital records consists of four distinct activities about records: compiling information; assessing value; determining feasibility of preservation; and making the appraisal decision. Assessment of authenticity in the context of assessing value is an

integral part of records' appraisal. Appraisal rests on a foundation of solid research into the five contexts of records: juridical/administrative, provenancial, procedural, documentary, and technological. Contextual data assists in assessing record value and authenticity, and identifying digital components that must be preserved. This module gives a background into the development of appraisal strategies for digital records, and introduces the reader to the recommended process of appraisal guided by the COP Model. The module provides guidelines to analyse legacy files to establish authenticity, that is, to examine data leading to the presumption of authenticity, and, if there is an insufficient basis for a presumption of authenticity, to carry out the verification of authenticity. It then provides a template for conducting appraisal and documenting appraisal decisions.

The second part of this module outlines a methodology and action plan for an organisation to move from a record-creating environment where unstructured records and documents are stored and maintained in network drives to a controlled record-creating and keeping environment such as an ERMS or EDRMS. This module walks individuals through the processes of evaluating their organisation's record-creating and recordkeeping environment; identifying and appraising their organisation's digital records; and preparing them for migration to an ERMS or EDRMS (or other structured and secure records management system), based on a case study from InterPARES 3.

5.6 Module 6: E-mail Management and Preservation

Managing and preserving electronic mail (e-mail) has become a formidable challenge for most organisations. This module is designed to help organisations gain better control of their e-mail based on the E-mail Management and Preservation Model (EMPM)³, a multi-phase process for implementing e-mail management and preservation policies and procedures. No one technological solution is promoted – successful management of email depends on analysis of the organisation's context, goals, and requirements. This module discusses the various factors that influence e-mail management and preservation, different e-mail management methods, ways to apply retention and disposition to e-mail, ways to preserve e-mail, and the design and implementation of e-mail policies and procedures.

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

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

³ The E-Mail Management and Preservation Model was developed through the course of InterPARES and for this module by Donald Force, PhD candidate, University of British Columbia.

to and from websites into preservation environments, and situates the process within the policy framework of the organisation.

5.8 Module 8: Cloud Computing Primer

Cloud computing consists of on-demand computing services delivered over the Internet from a remote location or via an organisation's servers. The National Institute of Standards and Technology (NIST) defines cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or services provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models" (NIST 2010).

This module summarizes the current state of cloud computing and identifies some of the records management issues and challenges that should be considered before an organisation moves its records, services and/or processes to the cloud. The first section provides a summary of the characteristics, service and deployment models of cloud computing. The module then addresses some of the key concerns for records professionals, including interoperability, business continuity, privacy and confidentiality, intellectual property and copyright, retention and disposition, and ownership of data.

6. Delivery

Digital Records Pathways is predicated on the theoretical and methodological findings of InterPARES, with acknowledgment of and references to other international research and best practices in digital preservation. The modules are intended to be expanded and adapted to specific organisational and national contexts and updated on a continuing basis (depending on personnel, hosting, and delivery constraints). They can be adapted to a variety of different delivery mechanisms.

The current presentation of the modules assumes a traditional teacher-student learning format (whether the content is taught by an instructor or self-administered), however, the affordances of digital communications technologies offer the possibility of many creative and evolving delivery formats. Haythornthwaite has identified today's learner as not merely "an empowered user [of educational content], but a participant in the production, direction, commentary, discussion, and reproduction of information and knowledge" (Haythornthwaite 2009). It is hoped that these modules will provide not only a useful knowledge base in and of themselves, but also the potential for creative interaction and knowledge production through the addition of comments, case studies, examples, exercises, and research findings, and through aggregation and reformatting – a resource that remains current, in a state of "perpetual beta" (Haythornthwaite 2009).

Bibliography

- CCSDS. 2002. "Reference Model for an Open Archival Information System (OAIS)". Consultative Committee for Space Data Systems. <http://public.ccsds.org/publications/archive/650x0b1.pdf>.
- Duranti, Luciana. 2005a. *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*. San Miniato: Archilab.
- . 2005b. "InterPARES 3 Grant Proposal." www.interpares.org.
- Duranti, Luciana, and Sandie Bradley. 2010. "General Study Research Proposal-Education Modules". InterPARES 3. www.interpares.org.
- Duranti, Luciana, and Randy Preston. 2008. *Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*. Padova: Associazione Nazionale Archivistica Italiana.
- Haythornthwaite, Caroline. 2009. "Learning in the Age of Web 2.0" presented at the Leverhulme Trust Public Lectures, London, England. <http://newdoctorates.blogspot.ca/2009/10/leverhulme-trust-public-lectures.html>.
- NIST. 2010. "Definition of Cloud Computing V15". Online. <http://csrc.nist.gov/groups/SNS/cloud-computing/>.

Luciana Duranti is Chair of the Archival Studies Programs at the University of British Columbia, Canada, and Professor of archival theory, diplomatics, and management of digital records. She is Director of the Centre for the International Study of Contemporary Records and Archives (CISCRA) and of several research projects about digital records, among which InterPARES, Digital Records Forensics, and Records in the Cloud.

Corinne Rogers is a doctoral candidate in the School of Library, Archives, and Information Studies at the University of British Columbia, Canada. She teaches digital diplomatics and digital records forensics, and her doctoral research investigates concepts of authenticity of digital records, documents, and data, and the assessment of authenticity of digital documentary evidence in the legal system.