

InterPARES 3 Project

International Research on Permanent Authentic Records in Electronic Systems

TEAM Canada

From Us To You: Overview of Research Findings and Products

Luciana Duranti InterPARES Project Director

International Symposium
Barcelona, Spain
5 October 2011

Goal of InterPARES 3 (2007-2012)

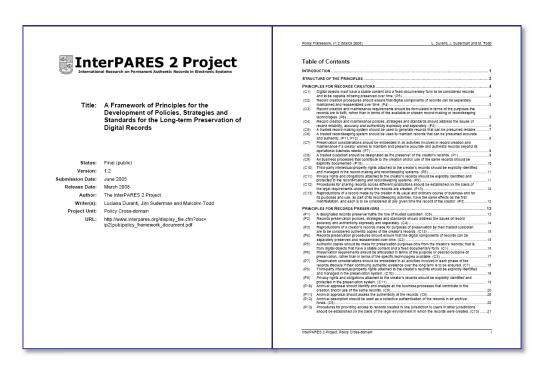
To enable public and private archival organizations and programs with limited resources to preserve over the long term authentic records that satisfy the requirements of their stakeholders and society's needs for an adequate record of its past.

It did so by building on the products of the first two phases of InterPARES (1998-2006)

Key IP 1 & 2 Products

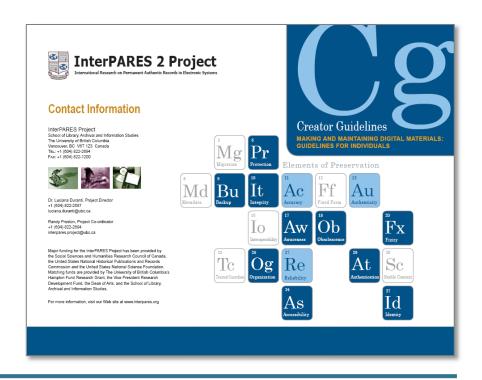
Policy Framework

A framework of principles guiding the development of policies for records creating and preserving organizations



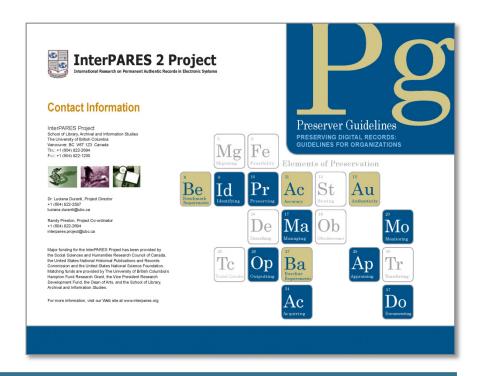
Creator Guidelines

Recommendations for making and maintaining digital materials for individuals and small communities of practice



Preserver Guidelines

Recommendations for digital preservation for archival institutions



Benchmark and Baseline Requirements

Authenticity requirements for assessing and maintaining the authenticity of digital records

« REQUIREMENT SET A »

To support a presumption of authenticity the preserver must obtain evidence that:

REQUIREMENT A.1: Expression of Record Attributes and Linkage to Record
The value of the following attributes are explicitly expressed and inextricably linked to every
record. These attributes can be distinguished into categories, the first concerning the identity of
records, and the second concerning the integrity of records.

A.1.a Identity of the record:

- A.1.a.i Names of the persons concurring in the formation of
 - name of author
 - name of writer^b (if different from the author)
 - name of originator^c (if different from name of auth
 name of addressee^d

A.1.a.ii Name of action or matter

A.1.a.iii Date(s) of creation and transmission, that is:

- chronological date^e
- received date
- archival date^g
 transmission date(s)^h

A.1.a.iv Expression of archival bond (e.g., classification code

A.1.a.v Indication of attachments

A.1.b Integrity of the record:

A.1.b.i Name of handling office

A.1.b.ii Name of office of primary responsibility^k (if different t

A.1.b.iii Indication of types of annotations added to the recor

A.1.b.iv Indication of technical modifications^m

REQUIREMENT A.2: Access Privileges

The creator has defined and effectively implemented access privilege modification, annotation, relocation, and destruction of records.

<< REQUIREMENT SET A (cont) >>

REQUIREMENT A3: Protective Procedures: Loss and Corruption of Records
The Creator has established and effectively implemented procedures to prevent, disconrect loss or corruption of records.

REQUIREMENT A.4: Protective Procedures: Media and Technology

The creator has established and effectively implemented procedures to guarantee the oridentity and integrity of records against media deterioration and across technological or

REQUIREMENT A.5: Establishment of Documentary Forms

The creator has established the documentary forms of records associated with each p either according to the requirements of the juridical system or those of the creator.

REQUIREMENT A.6: Authentication of Records

If authentication is required by the juridical system or the needs of the organization, th has established specific rules regarding which records must be authenticated, by wife the means of authentication.

REQUIREMENT A.7: Identification of Authoritative Record

If multiple copies of the same record exist, the creator has established procedures the which record is authoritative.

REQUIREMENT A.8: Removal and Transfer of Relevant Documentation

If there is a transition of records from active status to semi-active and inactive status involves the removal of records from the electronic system, the creator has establist effectively implemented procedures determining what documentation has to be removed to the preserver along with the records.

«REQUIREMENT SET B»

The preserver should be able to demonstrate that:

REQUIREMENT B.1: Controls over Records Transfer, Maintenance, and Reproduction The procedures and system(s) used to transfer records to the archival institution or program; maintain them; and reproduce them embody adequate and effective controls to guarantee the records' identity and integrity, and specifically that:

- B.1.a Unbroken custody of the records is maintained;
- B.1.b Security and control procedures are implemented and monitored; and
- B.1.c The content of the record and any required annotations and elements of documentary form remain unchanged after reproduction.

REQUIREMENT B.2: Documentation of Reproduction Process and its Effects

The activity of reproduction has been documented, and this documentation includes:

- B.2.a The date of the records' reproduction and the name of the responsible person:
- B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver;
- B.2.c The impact of the reproduction process on their form, content, accessibility and use; and
- B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user.

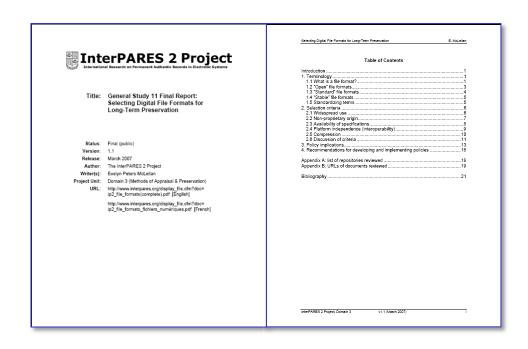
REQUIREMENT B.3: Archival Description

The archival description of the fonds containing the electronic records includes—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the electronic records of the creator have undergone since they were first created.



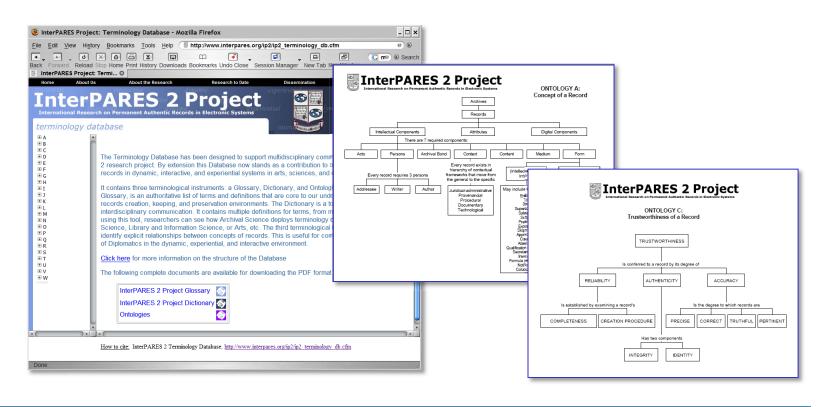
File Format Selection Guidelines

Principles and criteria for adoption of file formats, wrappers and encoding schemes



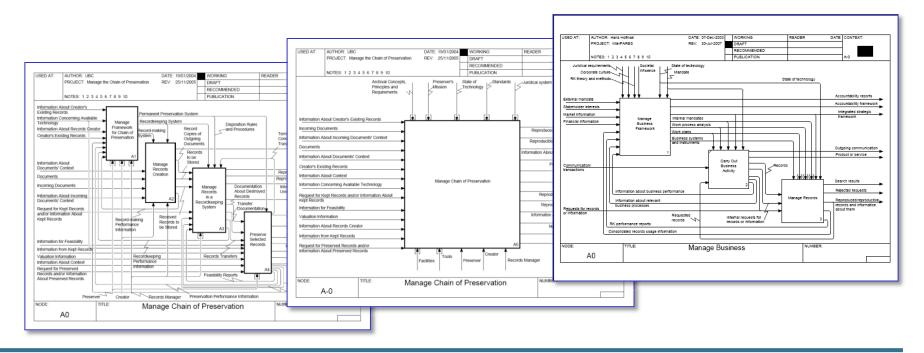
Terminology Database

Including a glossary, a dictionary and ontologies



Two Records Management Models

Chain of Preservation (COP) Model (lifecycle)
Business-driven Recordkeeping (BDR) Model (continuum)



IP 1 & 2 Final Products

Two books:

Luciana Duranti, ed. *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (San Miniato: Archilab, 2005). Available on line at

http://www.interpares.org/book/index.cfm

Luciana Duranti and Randy Preston, eds. *InterPARES 2: Interactive, Dynamic and Experiential Records* (Roma: ANAI, 2008). Available on line at

http://www.interpares.org/ip2/book.cfm.

InterPARES 3 International Alliance

- Teams: TEAM (Theoretical Elaboration into Archival Management) Canada (including US); Africa; Brazil; Catalonia; China; Colombia; Italy; Korea; Malaysia; Mexico; Norway; Singapore; and Turkey.
- **Director:** Luciana Duranti
- **Headquarters:** UBC SLAIS (facilities provided by UBC)
- Funding: SSHRC, and various sources from each country

3 Primary Components

1. Research component

(short-term and long-term projects, including case studies and general studies)

- 2. Education and training component
 - (in the context of research projects, apprenticeships, activities credited as part of coursework, etc.)
- 3. Knowledge-mobilization component (workshops, seminars, colloquia, policy manuals and other publications, public lectures, etc., that meet the needs of both academic and community partners)

Methodology (general)

Action Research

- Practical, collaborative, pragmatic research directed toward producing solutions that are directly useful to a group of people
- Research subjects are co-participants and stakeholders in the process
- We **jointly** define research objectives and goals, coconstruct research questions, pool knowledge and **develop solutions** and performance tests that implement specific strategies

Methodology (general - cont.)

Action Research

Two distinct methods of research:

- 1. Prototype development research
- 2. Ethnographic research

Methodology (general - cont.)

Prototype development research

- User-centered, collaborative prototyping approach that explores the interplay between theory and practice
- Proof-by-demonstration
- Comprises three major iterative stages:
 - 1. concept building
 - 2. system building
 - 3. system evaluation

Methodology (general - cont.)

Ethnographic research

- Creators of records, their users and archivists form a community of practice—the archival environment—for which social interaction creates meaning and defines values
- Researchers place themselves within an archival environment to gain the cultural perspective of those responsible for records
- Observation of the environment with detailed description,
 extensive interviewing and analysis of the documents

Outcomes

- Policies, strategies and procedures for archival organizations or programs, and guidelines for the records creators whose records fall under their responsibility.
- Action plans for the specific case studies carried out in the course of the Project.
- Criteria to determine "most-at-risk" materials
 e.g., checklist of age (date created, date last accessed), physical
 carrier, operating system, software used, equipment required
 and its availability, etc.

Outcomes (cont.)

- Evaluation models for assessing the degree of risk, if any, involved in the chosen preservation action.
- Cost-benefit models for archival organizations or programs.
- Ethical models that identify and make explicit the consequences for individuals and society of types of preservation measures or lack thereof.

Outcomes (cont.)

- Training and education modules for preservers, professional associations and university programs; and a strategy for delivering them.
- **Position papers** directed to key regulating, auditing and policymaking bodies, advocating the vital need of embedding planned digital preservation in the requirements they issue for the activities they regulate, audit or control.

Case Studies

Case studies (3 types):

- 1. Dealing with specific groups or types of **records**.
- 2. Dealing with record/information systems.
- 3. Dealing with **policies**.

Case Studies (cont.)

- The analysis of the testbeds by all researchers produced action items which were implemented and the outcome of which was reported back to the TEAM.
- The process continued in an **iterative** way until the archival environment and all TEAM researchers were satisfied with the solutions
- The entire process was guided by a case study flowchart, which ensured that all steps were followed in the correct order, and was concluded by a final report.
- Each document produced in the course of the case study was structured as to form and content on the basis of a template used for all case studies.

General Studies

General studies (3 types):

- 1. Studies carried out by one TEAM or a group of TEAMs for the benefit of all TEAMs (e.g., Annotated Bibliography of International Standards, E-mail Preservation)
- 2. Studies in which all **TEAMs** take part (e.g., Terminology Database).
- 3. Studies conducted by a TEAM that will prove **useful for that TEAM only** (e.g., Annotated Canadian Standards)
- 4. Studies conducted by a group of TEAMS for its own benefit (e.g. Protocol Register)

General Studies (collaborations)

- Web 2.0/Social Media (Canada and Turkey)
- Terminology (International Alliance)
- Digital Preservation Projects (International Alliance)
- International Standards Relevant to IP3 (International Alliance)
- Bibliographic Database
- E-mail Preservation (Italy)
- Protocol Registry (Catalonia, Italy, Brazil)

General Studies (TEAM Canada)

- Canadian Standards Relevant to IP3
- Community Archives e-Records Assessment
- Public Sector Audit Report for Digital Recordkeeping
- University IR/IRK Survey
- Records Management Policies and Procedures Template
- Cost-benefit Models
- Ethical Models
- File Viewers
- Education Modules
- Open Source Records Management Software
- Metadata Applications Profiles
- Organizational Culture & Risk Assessment

Education Modules

- Address the identified gap in educational resources in digital preservation strategies for small and medium sized organizations
- Produce in-house training kits for staff of small and medium sized archival organizations plan digital preservation program
- Assist professional associations in providing training and career development for members
- Provide university programs with content and structure for courses in digital preservation
- Provide non-archivists with necessary tools to create, manage and preserve authentic records
- Select effective delivery and dissemination tools

Education Modules-list

Introductory Module

Introduces the set of modules and explains how to use them, outlines objectives, and summarizes the contents of each module. It includes resources for institutional readiness and self-assessment tools, lays out the logic for the chain of preservation model and summarizes the findings of InterPARES.

Module 1: Development of a Digital Preservation Policy

Module 2: Management and Preservation of E-mails

Module 3: Management and Preservation of Records in Web Environments

Module 4: Selection and Appraisal of Digital Records



Education Modules-list

Module 5: Organizational Culture and its Effects on Records Management

Module 6: Planning a Shared Drive Migration to a Controlled Records Environment

Module 7: Management and Preservation of Records in Specialized Environments

Module 8: Cloud Computing Primer

Module 9: Metadata

Open Source Records Management Software

- Numerous proprietary electronic records management systems (ERMS) are currently on the market, but their high cost place them out of reach of small and medium-sized organizations
- Several open-source ERMS have emerged, introducing the possibility of implementing electronic records management without paying heavy software licensing costs
- Focus on Alfresco Records Management, the most widely used product and the only one certified to comply with the *Design Criteria Standard for Electronic Records Management Software Applications* (DoD 5015.2)
- Purpose: determining whether an organization would be likely to deploy it successfully
- Concluded that Alfresco does not provide an open-source electronic records management tool that is feasible for use in small to medium-sized organizations

Metadata Applications Profiles

- Metadata used to aid in the presumption of authenticity of digital records.
- This would allow others who wanted to use the same metadata to understand what functionality they would gain from adopting this system.
- Interoperable metadata if it is clearly defined both for humans and machines: synchronic interoperability because it is happening at one point in time (now); diachronic interoperability systems working through time.
- Intentional interoperability is when it is linked to an application profile

Metadata Applications Profiles (cont.)

There are two sets of three things we need to examine in the social world of metadata:

- 1. Meaning: Semantics, Syntax, and Pragmatics
- 2. Interoperability: Synchronic, Diachronic, and Intentional

We are currently working through many issues in our work with an IP AP. Here are two:

- 1. Units of analysis (record / aggregation)
- 2. Attestations (explicit identity metadata) vs. Contextualization (can we from context infer and later describe these attributes of the body of records?)

We must continue to reaffirm **purpose**, **function**, **and context** for metadata vs. description in digital preservation.

Organizational Culture

Findings of case studies

- "institutional culture" was an obstacle to completion of a case study
- organizational culture was "uncommitted to good records management practices"
- similar E-mail case studies yielding drastically different outcomes
 - need of an e-mail general study
- similar recordkeeping case studies requiring drastically different development procedures

Organizational Culture

Definition:

The specific collection of values and norms that are shared by people and groups in an organization and that control the way they interact with each other and with stakeholders outside the organization

Research Questions

- How does organizational culture affect the selection and implementation of recordkeeping and/or digital records preservation systems?
- To what extent does organizational culture **affect the ability of InterPARES 3 in carrying out its research in the test-bed sites?** What conditions would be necessary with respect to organizational culture for InterPARES 3 to carry out its research in the test-bed?
- What are the **fundamental similarities and differences** in organizational culture with respect to recordkeeping and/or preservation practices of different organizations within and across similar industries?

Research Questions (cont.)

- Within the same organization, what are the variations of sub- or professional cultures that shape recordkeeping and/or preservation practices?
- What are the **varying levels of expectations** stakeholders have in terms of their roles and responsibilities in recordkeeping/and or preservation as well as their expectations of other stakeholders?
- What are the **methodologies** for facilitating the selection and implementation of recordkeeping and/or digital records preservation systems?

Summary of Findings

- Developing, learning and teaching **how to use** the structural features of an application or a system **is important**
- More important is learning the spirit behind those features
- Users who are not acquainted with archival principles and methodologies may – intentionally or unintentionally – appropriate an application or a system "unfaithfully" more easily than records professionals
- With digital tools, which are mostly developed by IT experts outside the organization that will use them and often without consulting archival professionals, unfaithful appropriations are likely to happen more frequently
- Interpretive flexibility or lack of it may determine acceptance or rejection

Summary of Findings (cont.)

- Training for users of applications and systems "emphasizes details of use rather than general philosophy."
- **Time is important**: the moment of the launch of a new system is very critical for its success
- Managers favor the explicit knowledge that is incorporated in organizational artifacts like processes, structures, documents, and technology
- Thus, it has been common to design systems primarily focused on the codified, explicit organizational knowledge
- Management reporting systems, decision support systems, and ERMS, are all focused on the identification, collection, and dissemination of this knowledge type

Summary of Findings (cont.)

- We must pay more attention to **knowledge management** literature
- A core competency for implementers requires *know-how*, i.e. "the particular ability to put know-what into practice"
- Fostering this more complex form of organizational capital should be the focus of our case studies
- The outcome of our efforts will be successful only if
 - we are able to make the archival environment understand the spirit of what we recommend
 - we will be able to incorporate into our recommendations the outlook and way of working of those whom it intends to serve.

InterPARES 3 Web Site

