

Towards InterPARES 3



Dr. Luciana Duranti
InterPARES Project Director

ENGLISH

The Goal of InterPARES 1 and 2 1999-2007



To develop the body of theory and methods necessary to ensure that digital records produced in databases and office systems as well as in dynamic, experiential and interactive systems in the course of artistic, scientific and e-government activities can be created in accurate and reliable form and maintained and preserved in authentic form, both in the long and the short term, for the use of those who created them and of society at large, regardless of technology obsolescence and media fragility.



Object of Inquiry: Records

A record is any document created (i.e., made or received and set aside for action or reference) by a physical or organizational person in the course of a practical activity as an instrument and by-product of it. A record is more than:

- **document** is information affixed to a medium in a determined form
- **information** is a message intended for communication across space or time
- **data** is the smallest meaningful piece of information

A **digital record** is a record created (i.e., made or received and set aside for action or reference) in digital form

Characteristics of a Digital Record



A digital record must have:

- **stable content** and **fixed form**;
- **identifiable administrative and documentary contexts**, through explicit linkages to other records within or outside the digital system; and
- **five persons** (author, writer, originator, addressee, creator) involved in its creation.
- It must **participate in or support an action** either procedurally or as part of the decision making process; and
- It presents **formal elements, attributes, and digital components**.



Qualities We Want to Protect

- **Reliability**: the trustworthiness of a record as a statement of fact. A reliable record is complete and generated according to a controlled procedure
- **Accuracy**: the exactness and correctness of a record content, dependent on the competence of the author and the controls on the process by which data are recorded and transmitted through space (i.e., between persons, systems or applications) and time (i.e., when stored off line, or when the hardware or software used to process, communicate or maintain it is upgraded or replaced)
- **Authenticity**: the trustworthiness of a record as a record. An authentic record is one that has not been tampered with or otherwise corrupted. Authenticity is maintained by protecting a record identity and integrity. It differs from

Authentication: a means of declaring authenticity at a point in time

The Digital Records Challenge



- They do not exist as physical entities, but are constituted of linked digital components (the “manifested” record differs from the “stored” record)
- Their original manifestation disappears when they are saved: we cannot maintain or preserve digital records, but only the ability to re-produce or re-create them
- The facility of reproduction and manipulation makes it difficult to identify the final, official, reliable or accurate version
- Technological obsolescence makes digital records inaccessible in a very short time span
- Intellectual property and privacy rights are hard to protect



The Digital Records Challenge (cont.)

Most systems that should contain records do not, because the entities in them **lack fixed form and stable content**.

- In dynamic systems they depend for their content upon data extracted from a variety of other systems which may have variable instantiations.
- In experiential systems they are intended to incorporate the behaviour of the rendering system and the effects of subjective user's interactions.
- In interactive systems, each user intervention or input from another system causes a change of content and/or form.

The systems that do contain records, contain bad records, primarily because of **lack of identifiable contexts and relationships** among themselves and with records outside the system

InterPARES Principles



- Technology cannot determine the solution to the reliable and accurate creation of digital records or to their authentic preservation over the long term: **organizational needs** define the problem and **archival principles** must establish the correctness and adequacy of each technical solution
- Solutions to the digital records challenges are inherently **dynamic** and **specific** to the cultural, disciplinary, administrative and legal situations
- Preservation is a **continuous process that begins with records creation**
- We must be able to **presume records trustworthiness**, till proof to the contrary is established
- We must be able to **infer authenticity** on the basis of the circumstances of records creation, maintenance and preservation

InterPARES Research Activities



- Identification of **what constitutes a record** in each type of system and in each context, and **what record has the force of an original**, based on archival and diplomatics theory and six thousand years of record making and recordkeeping
- Definition of **what a reliable, accurate and authentic record** is in the arts, science, law and administration on the basis of analysis of literature, surveys, interviews
- Development of the requirements for the design of **a trusted record making system, a trusted recordkeeping system and a trusted record preservation system**, on the basis of 42 case studies, and the modeling and diplomatic analysis of the results
- Development of **methods and procedures for the creation, maintenance, appraisal, selection and disposition, and long-term preservation of digital records**, based on archival theory, law, and issues related to organizational culture

InterPARES Final Products



- A **framework of principles guiding the development of policies** for records creating and preserving organizations
- **Guidelines for making and maintaining digital records** for individuals and small communities of practice
- **Guidelines for digital preservation** for archival institutions
- **Authenticity requirements** for records systems
- A **metadata registry** for the registration and analysis of metadata schemas
- Principles and criteria for adoption of **file formats, wrappers, and encoding**
- A **terminology database** including glossary, dictionary and antologies

Intellectual Framework for Policy Development: Purpose



- To establish a flexible, consistent and solid basis for the development of policies, strategies and standards
- To provide a core of concepts capable of balancing different cultural, social and juridical perspectives on a variety of issues, such as records trustworthiness, access to information, data privacy, and intellectual property
- To demonstrate the need for a strong continuing relationship between records creators and records preservers throughout the lifecycle of the records
- To show the nature of such relationship and the ways in which it should play out

Record: Fixed Form & Stable Content



[C1] Digital entities must have stable content and a fixed documentary form to be considered records, to serve a memorial function, and to be capable of being preserved over time

[P5] Authentic copies should be made for preservation purposes only from the creator's records, that is from digital entities that have a stable content and a fixed documentary form

Stable Content and Fixed Form



- **Stable Content:** the fact that the data and the message in the record are unchanged and unchangeable, meaning that data cannot be overwritten, altered, deleted or added to
- **Fixed Form:** its binary content is stored so that the message it conveys can be rendered with the same presentation it had on the screen when first saved (different digital presentation)



Stable Content and Fixed Form (cont.)

- But, if the same content can be presented on the screen in several different ways in a limited series of possibilities, we may have either a different view of **the same stored record** having stable content and fix form (different documentary presentations e.g. statistical data as a pie chart, a bar chart, or a table) or **several manifested records** with stable content and fixed form derived from the same stored record
- **Bounded Variability**: when there is no stored record but content data, form data and composition data that are quite separate and only connected by a query, and changes to the form are limited and controlled by fixed rules, so that the same query or interaction always generates the same result, and we have different views of different subsets of content, due to the intention of the author or to different operating systems or applications

Digital Components



[C2] Records creation procedures should ensure that the digital components of records can be separately maintained and reassembled over time

[P4] Records preservation procedures should ensure that digital components of records can be separately preserved and reassembled over time.

Elements, Attributes, Components



- **Formal Elements:** a constituent part of the record documentary form as shown on its face (intrinsic or extrinsic)—in MoReq 2 called intellectual components
- **Attributes:** the unique characteristics of the records that demonstrate its identity (expressed in identity and integrity metadata)
- **Digital Components:** objects that either contain one or more records or are contained in the record and require a specific preservation measure

Creation, Maintenance, and Preservation Requirements



[C3] Record creation and maintenance requirements should be formulated in terms of **the purposes the records are to fulfill**, rather than in terms of the available or chosen record-making and record-keeping technologies available

[P6] Preservation requirements should be formulated in terms of **the purpose or desired outcome of preservation**, rather than in terms of the available or chosen technologies available

Records Trustworthiness



[C4] Records creation and maintenance policies, strategies and standards should address the issues of record **reliability, accuracy, and authenticity** expressly and separately

[P2] Records preservation policies, strategies and standards should address the issues of record accuracy and authenticity expressly and separately

Trusted Record-making System



[C5] A trusted record making system should be used to generate records that can be presumed reliable and accurate

No corresponding requirement for the Preserver other than as a Record Creator itself

Trusted Recordkeeping and Record Preservation Systems



[C6] A trusted recordkeeping system should be used to maintain records that can be presumed accurate and authentic

[P11] Archival appraisal should assess the authenticity of the records.

[P12] Archival description should be used as a collective authentication of the records in a fonds

Preservation Begins at Creation



[C7] Preservation considerations should be embedded in all activities involved in record creation and maintenance if a creator wishes to maintain and preserve authentic records beyond its operational business need

[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

Trusted Custodian



[C8] A trusted custodian should be designated as the preserver of the creator's records

[P1] A designated record preserver fulfills the role of trusted custodian

Designated Preserver



Trusted custodian:

- acts as a **neutral third party**, i.e., demonstrates that he/she has no stake in the content of the records and no reason to alter records under his/her custody, and that he/she will not allow anybody to alter the records either accidentally or on purpose,
- is equipped with the **knowledge and skills** necessary to fulfil its responsibilities, which should be acquired through formal education, and
- establishes a **trusted preservation system** that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved

Business Processes



[C9] All business processes that contribute to the creation and/or use of the same records should be explicitly documented

[P10] Archival appraisal should identify and analyze all the business processes that contribute to the creation and/or use of the same records

Intellectual Property Rights



[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] Third-party property rights attached to the creator's records should be explicitly identified and managed in the preservation system

Privacy Rights



[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 rights and obligations attached to the creator's records should be explicitly identified and protected in the preservation system

Records Sharing Across Jurisdictions



[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] 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

Reproduction of Records



[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 of its first created manifestation and each is to be considered at any given time the record of the creator

[P3] Reproductions of a creator's records made for purposes of preservation by their trusted custodian are authentic copies of the creator's records

Why a Third Phase?



- A study of the effectiveness of workshops and seminars experiences for increasing archivists' skills in digital preservation and their ability to implement these skills in their repositories has shown that very few participants were able to implement the skills once they returned to their work environments (Duff, Wendy M., Amy Marshall, Carrie Limkilde, and Marlene van Ballegooie. "Digital Preservation Education: Educating or Networking?" *The American Archivist*, 69, 1 (2006): 188-212. In the context of ERPANET.)
- Feedback on the outcomes of the two phases of InterPARES from archivists working in institutions smaller than national archives has consistently shown concern about their downward-scalability and their relevance to small and medium sized organizations.

Goal of InterPARES 3



To enable small and medium sized public and private archival organizations and programs (units within records creating organizations), which are responsible for the digital records resulting from government, business, research, art and entertainment, social and/or community activities, 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.

Objectives



1. To **promote an environment supportive** of the research goal by demonstrating to regulatory and auditing bodies, and policy makers that they ought to embed digital records preservation requirements in any activity that they regulate, audit or control.
2. To collaborate with small and medium sized archival organizations and programs in the **development of policies, strategies, procedures, and/or plans of actions** for the preservation of the digital materials that they expect to acquire or have already acquired, using the recommendations and products of leading edge research projects.
3. To **assess the applicability of** the recommendations of such projects about **trusted record-making and recordkeeping** to the situations of the small and medium sized archival organizations or programs selected as test-beds, and in particular the validity of InterPARES statements about the relationship between preservers and the records creators.

Objectives (cont.)



4. To assess the **applicability of these projects' preservation solutions** to the concrete cases identified by the test-bed partners as needing immediate attention, both when the records in question are already in their custody and when they still reside with their creator.
5. To refine and further **elaborate the theory and methods**, concepts and principles developed by these research projects on the basis of the results of the above activities.
6. To establish when such theory and methods, concepts and principles **apply across jurisdictions**, regardless of legal/administrative, social and cultural environment; and, in the situation where they do not apply, to identify why, and to determine the measures that are required to ensure the preservation of digital records.

Objectives (cont.)



7. To assist small and medium sized archival organizations or programs in addressing **the legal issues** that have been identified by the relevant research projects as providing obstacles to long term digital preservation, and additional issues that could be specific to the partner archival organizations and programs.
8. To formulate models that put into relation the choice of methods and objects of preservation with the **ethical consequences of each choice** for individuals and society, both in general and specifically.
9. To create **evaluation models** capable of measuring the success of the preservation solutions that have been proposed and implemented.
10. To develop **models of preservation costs** for various types of records and archival organizations and programs.

Objectives (cont.)



11. To **develop awareness and educational materials** that can a) enable the staff of small archival organizations and programs to plan for and carry out digital preservation, b) assist professional associations in promoting career development of their members, and c) provide university programs with content and structure for university courses on digital preservation; and to identify effective delivery methods.
12. To **ensure transfer of the body of knowledge** generated by this research—including actual examples and success stories—to appropriate local, national and international stakeholders.
13. To **establish a strong international network of research and education** on digital preservation that is deeply rooted in the various communities served by each of its partners, and that integrates academic work with social and community action.



Expected Products

1. **Policies, strategies and procedures** for small archival organizations or programs, and **guidelines** for the records creators whose records fall under their responsibility.
2. **Action plans** for the specific case studies carried out in the course of the project.
3. **Analysis of the validity, applicability or adaptability of action plans** developed in the specific cases studied to different organizations, contexts or countries.
4. **Comparison among the action plans** developed for the preservation of records at different stages in their lifecycle (i.e. planning, creation, use, maintenance, modification, preservation)

Expected Products (cont.)



5. **Criteria to determine "most-at-risk" materials** - like a checklist of age (date created, date last accessed), physical carrier, operating system, software used, equipment required and its availability, etc.
6. **Guidelines for addressing digital preservation requirements** that apply to specific types of records, but not to other materials.
7. **Evaluation models for assessing the degree of success**, if any, of the chosen preservation action.
8. **Cost-benefit models** for various types of archival organizations or programs and for various kinds of records and/or systems.
9. **Ethical models** that identify and make explicit the consequences for individuals and society of types of preservation measures or lack thereof.

Expected Products (cont.)



10. A dedicated **web site** providing small and medium sized archival organizations world wide with good practical knowledge, internationally developed and shared, concrete examples of successful implementation, models of preservation costs, of programs evaluation, of ethical preservation behaviors, etc.
11. A refined body of **theoretical and methodological knowledge** on digital preservation, communicated in conference papers, symposia, and refereed publications.
12. **Training and education modules** for archival organizations or programs, professional associations and university programs; and awareness and education modules for non archivists, such as IT professionals, vendors, and service providers; human resources and financial managers; doctors, communities of practice, members of the general public, etc.; and a strategy for delivering them.
13. **Position papers** directed to key regulating, auditing and policy making bodies, advocating the vital need of embedding planned digital preservation in the requirements they issue for the activities they regulate, audit or control.

Methodology (general)



- **Action research**: a set of disciplined, material practices that involve collaborative dialogue, participatory decision making, inclusive democratic deliberation, and the maximal participation and representation of all relevant parties
- Research becomes practical, reflective, pragmatic action-directed toward solving problems in the world
- Research subjects become co-participants and stakeholders in the process of inquiry

Methodology (general - cont.)



- Together, stakeholders and action researchers create knowledge that is pragmatically useful. In the process, they **jointly** define research objectives and political goals, co-construct research questions, pool knowledge, hone shared research skills, and fashion interpretations and performance tests that implement specific strategies for social change
- For our purposes, we can include under the umbrella of action research two distinct methods of research: system development research method and ethnography

Methodology (general - cont.)



- **Prototype development research method:** a user-centered prototyping approach that allows for exploration of the interplay between theory and practice, advancing the practice, while also offering new insights into theoretical concepts
- Developing a system in a research context can serve as proof-by-demonstration of the underlying theory, while producing an artifact that can form the basis of ongoing and expanded research
- This method comprises three major iterative stages – concept building (which we have done), system building and system evaluation
- All stages of system development must reflect this focus on the concept that the system is to illustrate

Methodology (general - cont.)



- **Ethnography**: a form of inquiry characterized by the position of the researcher vis-à-vis the phenomena being studied
- The researchers place themselves within an archival environment to gain the cultural perspective of those responsible for records. The creators of records, their users, and archivists form a community of practice - the archival environment - for which social interaction creates meaning and defines values (an expert on organizational culture should be part of the research team)
- The process of ethnographic research includes observation of the environment with detailed description, extensive interviewing, and analysis of the documents produced or accumulated in the first two activities



Methodology (general - cont.)

- In conclusion, action research' s goals are: 1) the generation of **new knowledge**, 2) the achievement of **action-oriented outcomes**, 3) the **education** of both researchers and participants, 4) results that are **relevant to the local setting**, and 5) a **sound** and appropriate research **methodology**
- It is a collaborative, reflective, experiential, rigorous, interpretive and participatory mode of research that involves:
 - The **active engagement** of individuals traditionally known as subjects as participants and contributors in the research enterprise
 - The **integration of practical outcomes**
 - A **series of steps**, each including planning, action, and evaluation
- It targets mainly two tasks: to **produce knowledge that will be directly useful to a group of people**; and to **enlighten and empower the members of the group**, motivating each one to take up and use the information gathered in the research

InterPARES 3 International Team



Teams: TEAM (Theoretical Elaboration into Archival Management) Canada; Holland and Belgium; Italy; Brazil; Africa; Korea; China; Ireland and England; Norway; Sweden; Singapore and Malaysia; USA; **Mexico.**

Director: Luciana Duranti

Headquarters: UBC - SLAIS (facilities provided by UBC)

Staff: Project Coordinator, Technological Coordinator, Administrative Assistant (salaries paid by SSHRC: the direction and management of the project is a Canadian expense built into TEAM Canada grant application)

Meetings: Once a year, each time hosted by a different country

Symposia: Once a year, each time hosted by a different country

Networking: One common Web site with a common public area, and a restricted area including both shared and limited-access spaces, listservs, databases and research documents, working areas, etc.



Partners

The project partners must **jointly define** the research activities as well as the participatory arrangements under which individual researchers and research teams will carry out those activities. The partners should continue to develop and refine the research activities and, in addition to strengthening the original alliance, should also **recruit new partners** during the period of the project.

Composition of Each TEAM



- **Director**, principal investigator
- **Co-applicants** (individual academic and professional researchers)
- **Test-bed partners** (organizations that are the locus and subject of the research, the primary stakeholders)
- **Resource partners** (organizations that have an expertise in some part of the research content and share it providing regular input and feedback—they can be from the TEAM's country or foreign)
- **Collaborators** (individuals who have a special expertise in some parts of the research and will be called to contribute to the project as needed)
- **Graduate Research Assistants**

Primary Activities



- **a research component** (short-term and long-term projects, action research, etc.);
- **an education and training component** (in the context of research projects, apprenticeships, activities credited as part of coursework, etc.); and
- **a knowledge-mobilization component** (workshops, seminars, colloquia, policy manuals and other publications, public lectures, etc.) that meets the needs of both academic and community partners.

TEAM Canada Proposed Research



Title:

Theoretical Elaboration into Archival Management in Canada (TEAM Canada).

Subtitle:

Implementing the Theory of Preservation of Authentic Records in Electronic Systems in Small and Medium-Sized Archival Organizations and Programs.

Goal, Objectives, Products, Methods



- The same as those articulated for the overall project in general in the proposal and specifically by the International Team (directors of each TEAM)
- Inspired by the fundamental InterPARES principles of:
 - **inter/multi/transdisciplinarity**
 - **transferability**
 - **open inquiry**
 - **multi-method design**
 - **layered knowledge environment**

Methodology



- **Establishing the Context and the Research Cycle** -- Initially, each test-bed partner will identify a body of digital material for which a preservation plan will be developed, be it already in the custody of the partner or not. This is the context of problem solving with theory, and theory development through solving problems.
- **Data Collection** -- Data will be collected about the context and limitations of each test-bed, and the specific body of material, its documentary forms, technological constraints, functional or cultural meaning, etc.

Methodology (cont.)



- **First Iterations: Testing Different Solutions in Different Contexts** -- All (co-applicants, collaborators and) partners (i.e., test-bed and resource) will then reflect on these data and collectively articulate several possible solutions from which a single plan of action will emerge and be tested. This plan of action will include strategy, protocols, functional requirements, procedures, and expected outcome. The plan will be implemented and tested and test results will include performance assessment of this plan against benchmarks and baselines established in extant research.
- **Comparison of First Iterations** -- The results of the test will be shared among all researchers and analyzed. An assessment of these results will then allow us to reflect on this action, and refine our respective plans of action.

Methodology (cont.)



- **Second Iteration: Refining Solutions for Particular Contexts** -- After this assessment, the process will begin another cycle. This second iteration will account for anomalies in the test results, and benefit from the insight gained from a comparison across contexts. In so doing, it will refine our plans and performance measures. The second iteration will continue with small mini-iterations until a definitive plan of action is agreed upon for each context, implemented and tested again.

Methodology (cont.)



- **Comparisons of Second Iterations** -- The data will be compared among cohorts – the partner organizations of the same type (e.g., city archives, university archives, communities archives) – to establish what are the critical factors that determine the most appropriate solution for these contexts and whether they are linked to documentary forms, technology, organizational culture or function, or other environmental elements. This comparison will allow us to make statements of a general type. Furthermore, once a year, the results will be compared with parallel research results occurring abroad. This comparison of results will not happen in a vacuum because all international partners will share the same Web site and maintain ongoing communication.

Methodology (cont.)



- **Reflection, Analysis, and Synthesis** -- Throughout the research, the co-applicants and collaborators will reflect on issues and processes and make explicit their assumptions and biases, thereby giving rise to theoretical considerations. This reflexive and engaged scholarship will allow the co-applicants a chance to bind critical discourse with mission critical processes. Thus, while the project will start out with theory informing practice, as it will proceed, practice will refine theory, in a transformative cycle.

TEAM Canada Test-Bed Partners



- British Columbia Medical Association Archives
- Belkin Art Gallery Archives
- City of Vancouver Archives
- City of Victoria Archives
- Insurance Corporation of British Columbia Archives
- Museum of Anthropology Archives
- North Vancouver Museum and Archives
- Simon Fraser University Archives
- University of British Columbia Alma Mater Society Archives
- University of British Columbia Archives
- University of Victoria Archives

TEAM Canada Resource Partners



- Archival Association of British Columbia (AABC)
- Association of Canadian Archivists (ACA)
- British Columbia Corporate Information Management Branch (CIMB)
- Canadian Conservation Institute (CCI)
- Canadian Council of Archives (CCA)
- DOCAM Research Alliance (Langlois Foundation)
- Electronic Records Archives (ERA), NARA
- Irving K. Barber Learning Centre, UBC
- Library and Archives Canada (LAC)
- Royal British Columbia Museum (including the BC Archives)
- San Diego Supercomputer Center (SSC)



Examples of Case Studies

- **City of Vancouver:** ERDMS Examination & Implementation
- **City of Victoria:** Digitized and Born Digital Building Permits
- **Collaborative University Project (UBC, UVic, SFU):** E-mail Management, Preservation & Access
- **ICBC:** On-line Insurance Manual & Associated E-mail
- **ICBC:** Enterprise-wide Data Acquisition/Distribution System
- **UBC Alma Mater Society:** Web site Preservation & Access



Meetings

- **Steering Committee:** 4 times a year (twice in connection with the meeting of the full TEAM Canada)
- **Full TEAM Canada** (plenary research workshop): 2 times a year
- **International partners:** 1 time a year
- **Dissemination Committee:** 2 times a year at the time of the plenary research workshops



Timeline

- 1 July 2007—We will know whether the project will be funded
- **31 August 2007**—All TEAMS must be in place
- 1 September 2007—If funded, this is the starting date
- **October 2007**—First International Team meeting in Rome, Italy
- November 2007—We begin the case studies
- May 2012—End of project

International Partners Letter of Confirmation



- A statement that the partner (the Director of the national or multinational TEAM) has read the description of the project and agrees to the described role and level of participation
- A statement that the partner is willing and able to complete the activities described in the proposal
- A statement confirming the partner's financial and non-financial contribution (each international partner's contributions consist of hosting one meeting or a symposium in the course of the 5 years, and of supporting its director's participation in the yearly meeting and possibly in the symposium)

InterPARES Web Site



www.interpares.org