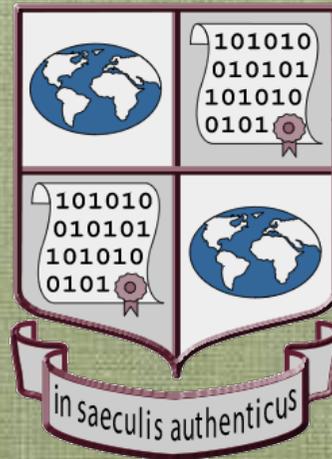
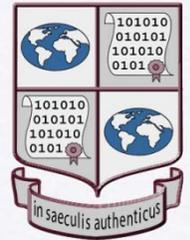


The InterPARES Project: An Overview



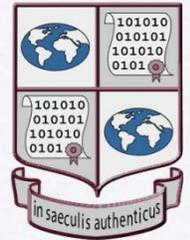
Dr. Luciana Duranti
InterPARES Project Director

InterPARES 1 Goal



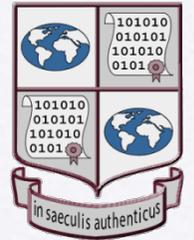
To develop the theoretical and methodological knowledge essential to the permanent preservation of authentic records generated and/or maintained electronically, and, on the basis of this knowledge, to formulate model policies, strategies and standards capable of ensuring that preservation.

InterPARES 1: 1999-2001



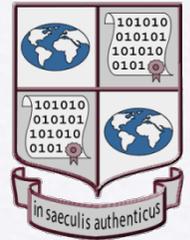
- Major Funding from SSHRC, NHPRC, UBC, NARA, NAC
- 13 countries in 4 continents, 60 researchers
- Public and private sector
- Professionals and academics (ratio 80% to 20%)
- Archival science, diplomatics and records management, plus computer science and engineering, jurisprudence, research methods

What Did We Study?



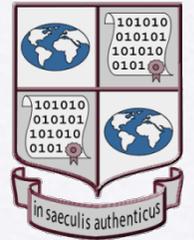
- Mostly textual records born digital in the course of administrative and legal activities
- Databases and document management systems
- Records in inactive or non-current status
- Characteristics of e-records and concept of authenticity
- Activities of appraisal and preservation from the preserver point of view

Methodology



- Theory and methods of diplomatics, archival science and law for the definition of concepts and development of requirements and methods
- Grounded theory for selection of case studies
- Comparative analysis for the study of appraisal and preservation reports
- IDEF0 modeling for the representation and definition of the activities involved in appraisal and preservation
- Chemistry for the study of storage media and computer engineering for the study of digital preservation technology and technological methods of authentication

InterPARES 1 Products



A body of concepts (e.g. record vs. document and data) and principles (e.g. trusted custodian), and a series of analytical instruments (e.g. a template) for studying new types of digital documents and developing new requirements and method as needed

Examples of Concepts



Reliability: the trustworthiness and accuracy of content

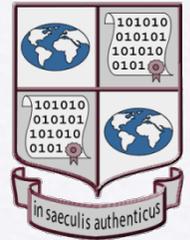
Authenticity: the trustworthiness of an entity as such. An authentic record is one that has not been tampered with or otherwise corrupted. Authenticity comprises identity and integrity.

Identity refers to the attributes of a record that uniquely characterize it and distinguish it from other records.

Integrity refers to the wholeness and soundness of a record. A record has integrity if it is intact and uncorrupted

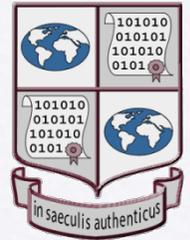
Authentication: a means of declaring authenticity at a point in time

Other InterPARES 1 Products



- Authenticity Requirements for those who generate and keep records and for those who preserve them (e.g. metadata for identity and integrity, access privileges, etc.)
- Selection and preservation methods and procedures (e.g. models representing activities and responsibilities)
- A framework for the development of policies, strategies and standards related to the proper creation, maintenance and preservation of digital records that can be proven authentic over time

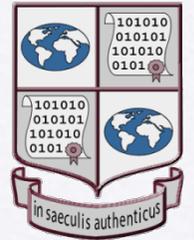
Most Important Findings



An electronic record must have:

- fixed form and unchangeable content;
- identifiable administrative and documentary contexts, and explicit linkages to other records within or outside the digital system; and
- five identifiable persons involved in its creation; and it
- must participate in or support an action either procedurally or as part of the decision making process

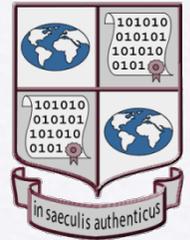
Most Important Findings (cont.)



Most systems that should contain records do not, because the entities in them lack fixed form and stable content. The systems that do, contain bad records, primarily because of lack of identifiable contexts and relationships

Inactive records that are no longer kept in active systems often cannot be preserved because either they were not created and/or maintained in preservable formats or they are obsolete

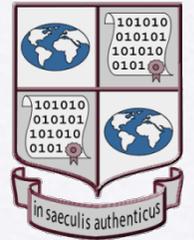
Most Important Conclusions



The preservation of authentic electronic records

- is a continuous process that begins with the records creation and whose purpose is to transmit authentic records across time and space
- must be predicated on the concepts of trusted record-keeping system and the role of the preserver as a trusted custodian
- must incorporate records appraisal and archival description

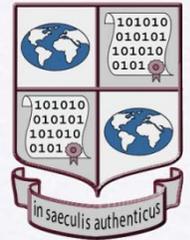
Most Important Conclusions (cont.)



The only way of preserving an inactive electronic record is to make an authentic copy of its last instantiation as an authentic record of the creator

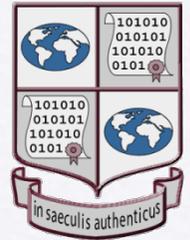
The preserver must be concerned with both the assessment and the maintenance of the authenticity of electronic records throughout their life-cycle

Lessons Learned



- Solutions to the preservation problem are inherently dynamic due to technological change and the increasing complexity of its products
- Technology cannot determine the solution to the long-term preservation of electronic records
- Archival needs define the problem and archival principles must establish the correctness and adequacy of each technical solution, or...must they?

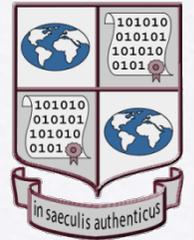
Usefulness of Theory



Archival and Diplomatic Theory

- demonstrated when systems were designed to contain fluid data rather than records
- showed what attributes of a record's identity are implicit in the system and need to be made explicit
- supported the identification of the requirements for a presumption of authenticity

However...



- The classic concept of record limited our capacity to understand electronic systems containing a variety of complex entities that do not correspond to it
- That which is known is not always very useful to understand that which is unknown
- Theory, which decontextualizes the record, is not useful to deal with the variety and complexity of systems: a complementary inductive approach is necessary

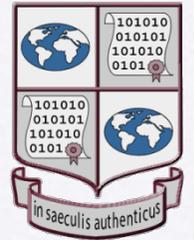
Therefore...



We developed a new research project that

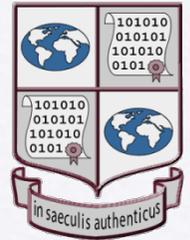
- examines all kinds of digital entities in complex systems and then develops a concept of record;
- is concerned with the entire life-cycle of the record and develops a chain of preservation model beginning with the design of the system and capable of ensuring accuracy and reliability in addition to authenticity;
- studies the digital entities created in the course of activities that have used complex systems well before government and business, that is, e-art and e-science; and
- uses the concepts and methodology of all the fields touched by our study

InterPARES 2 Goal



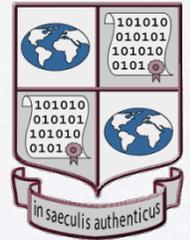
To ensure that the portion of society's recorded memory digitally produced in dynamic, experiential and interactive systems in the course of artistic, scientific and e-government activities can be created in accurate and reliable form and maintained and preserved in authentic form, both in the long and the short term, for the use of those who created it and of society at large, regardless of digital technology obsolescence and media fragility.

InterPARES 2: 2002-2006



- Major funding from SSHRC, NHPRC, NSF, UBC, and UNESCO (for last 18 months)
- 21 countries in 5 continents, 100 researchers
- Public and private sectors
- Academics and professionals (80% to 20%)
- Archival science, diplomatics and records management; music theory, composition, performance; film theory, production, description; dance and theatre theory; a variety of hard and social sciences; jurisprudence; computer science and engineering

What Do We Study?



Dynamic entities: they depend for their content upon data extracted from a variety of systems which may have variable instantiations.

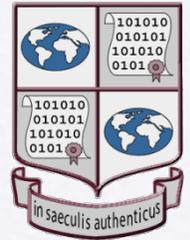
Experiential entities: objects whose essence goes beyond the bits constituting it to incorporate the behaviour of the rendering system and the effects of subjective user's interactions.

Interactive entities: objects to which each user intervention or input from another system causes a change of content and/or form.

Live, active materials, followed from creation to preservation, and obsolete inaccessible materials

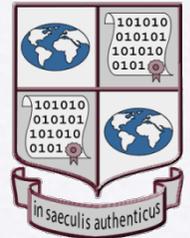
Approaches taken by both creators and preservers who have concerned themselves with accuracy, reliability and authenticity

Methodological Principles



- Interdisciplinarity
- Transferability
- Open Inquiry
- Layered Knowledge Environment
- Multi-method design: surveys, case studies, modeling, prototyping, diplomatic and archival analysis, and text analysis, etc. to address domain and cross-domain research questions

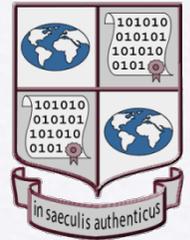
Team Structure



	FOCUS 1 Artistic activities	FOCUS 2 Scientific activities	FOCUS 3 Governmental activities
DOMAIN 1 Records creation & maintenance	Working Group 1.1	Working Group 1.2	Working Group 1.3
DOMAIN 2 Authenticity, accuracy & reliability	Working Group 2.1	Working Group 2.2	Working Group 2.3
DOMAIN 3 Methods of appraisal & preservation	Working Group 3.1	Working Group 3.2	Working Group 3.3
Terminology			
Policy			
Description			
Modeling			

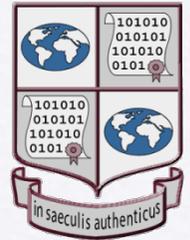
Matrix depicting Task Forces, Working Groups and Cross-domain Research Teams

Work Accomplished by Focuses



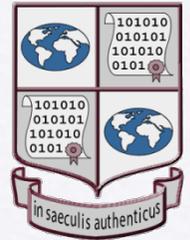
- 15 of the 21 case studies have been completed and represented in activity and entity models and analysed according to diplomatic and archival principles
- surveys of governments websites, of digital photographers, composers, and film makers, of the practice of preservation of interactive music, of file formats and encoding languages used for non-textual materials, and the analysis of a prototype for a persistent archives based on data grids
- annotated bibliographies and literature reviews, conceptual analyses of the findings of the reviews, and bibliographic databases for the management of references

Work Accomplished by Cross-Domains



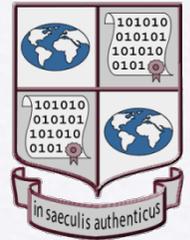
- **Terminology:** a terminology database including 4 lexicographic instruments, a Register, a Dictionary, a Glossary and a Thesaurus
- **Modeling:** Manage the Chain of Preservation model depicting all the activities involved in the management of electronic records throughout their lifecycle, from creation to permanent preservation

Work Accomplished by Cross-Domains (cont.)



- **Description:** Metadata Schema Registry, which is a centralized repository of schemas that will aid to identify metadata sets, or the combinations of elements from several sets which are appropriate to serve various recordkeeping needs; identification of the relationship between metadata and archival description; collaborations with modeling and policy cross-domains
- **Policy:** the identification of barriers to preservation which currently exist in laws, regulations, policies and standards concerning copyright and intellectual rights, privacy and freedom of information, authenticity and authentication, open standards and open source, and records and archival management

Domains' Work in Course



Development of

conceptual responses to the original research questions, primarily concerning the identification of the 'record' and of the preservable entities, and

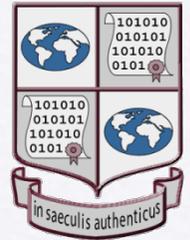
Production of:

guidelines for records creators and preservers; appraisal and preservation methodologies and procedures;

frameworks and models for the development of policies, strategies and legislation; and

descriptive schemas for digital entities

Findings to Date: Methodological



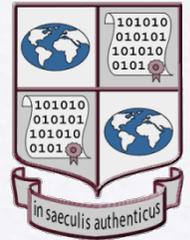
Surprisingly successful:

- looking at the arts and sciences for inspiration and guidance
- interaction with the creators both as co-investigators and as the interested party in the case studies and in the surveys
- continuous dissemination for feedback to the various stakeholders (e.g. scientists, arts conservators)

Expected:

The archival and diplomatic analyses of the preliminary results of the various research activities are fruitful only if one keeps in mind that these methods are retroactive by definition while we are examining live entities (not just active entities)

Findings to Date: Conceptual



- Fixed form and stable content are essential to the existence of a document, let alone a record
- Some systems are conceptually one record at any given moment in time, but they also disappear as such at any given moment in time
- Most entities examined in the case studies exist as potential records: they may become at any time 'interim records', kept records in a recordkeeping system, or records within the system creating them
- Creators want to have records and maintain functionality
- The 'one size fits all' solution cannot work

InterPARES Web Site



www.interpares.org