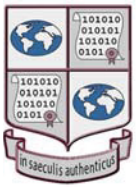


# InterPARES 3 Project

**International Research on Permanent Authentic Records in Electronic Systems**

**TEAM Canada**

# Intellectual Framework

## Version 2.0

## September 2008

**Introduction**

The InterPARES 3 Project is an international, multidisciplinary collaboration that is applying a multi-method approach to the development of practices, processes and tools that will help in the securing of a protected and lasting environment for the digital records produced by low-resources archives.

**Research Goal**

The goal of InterPARES 3 is to enable small and medium-sized, and low-resources public and private archival organizations and programs, 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.

**Research Objectives**

1. To promote an environment supportive of the research goal by demonstrating to regulatory and auditing bodies and to policy makers that it is essential to integrate 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 scalable policies, strategies, procedures, and/or action plans that they can implement in order to preserve 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 InterPARES and other 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 statements about the relationship between preservers and the records creators;
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 determine the measures that are required to ensure the preservation of digital records;
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 those that could be specific to their situation;
8. to formulate models that, for each choice of preservation methods and of digital objects to be preserved, identify the ethical consequences for individuals and society;
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 archives;
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 knowledge generated by this research—including actual examples and success stories—to appropriate local, national and international stakeholders; and
13. to establish a strong 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.

**Stakeholders**

The beneficiaries of the new knowledge produced by this Project are:

- small and medium-sized archival organizations and units/programs (hereinafter, archives);
- archives endowed with low resources;
- records creators for whom such archives represent the designated preserver;
- archivists, and any other professionals whose primary responsibility is to ensure the permanent and authentic preservation of records;
- the international digital preservation community;
- the legal and law enforcement communities;
- researchers of all disciplines;
- educators in all disciplines involved in this research;
- the computer and information technology sector; and
- the citizenry at large.

**Guiding Concepts**

Several guiding concepts of InterPARES 3 derive from the work of InterPARES 1 and 2.

*Record*: A document made or received in the course of a practical activity as an instrument or a by-product of such activity and set aside for action or reference.

*Accuracy*: The degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter.

*Reliability*: The trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about and is established by examining the completeness of the record's

form and the amount of control exercised on the process of its creation.

*Authenticity*: The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption. Authentic records are records that have maintained their identity and integrity over time.

*Authentication*: A means of declaring authenticity at a point in time by a person entitled to do so (e.g., public officer, professional, certifying authority).

*Trusted Record-making System:* A set of rules governing the making of records, and a set of tools and mechanisms used to implement these rules. To generate reliable records, every record-making system should include in its design records capture, identification, declaration, execution and transfer systems incorporating integrated business and documentary procedures, record metadata schemes, records forms, a record profile scheme, access privileges and record-making technological requirements, along with a rigorous monitoring function to oversee the activities of the entire system.

*Trusted Recordkeeping System*: The whole of the rules that control the creation, maintenance use and disposition of the records of the creator and provide a circumstantial probability of the authenticity of the records, and the tools and mechanisms used to implement those rules. To maintain authentic records, every recordkeeping system should include in its design records information, indexing, storage, retrieval, access and disposition systems incorporating a recordkeeping metadata scheme, a classification scheme, a retention schedule, a registration system, a controlled recordkeeping vocabulary, access privileges and procedures for maintaining accurate and authentic records, along with a rigorous monitoring function to oversee the activities of the entire system.

*Trusted Record Preservation System*: The whole of the rules that control the preservation and use of the records of the creator and provide a circumstantial probability of the authenticity of the records, and the tools and mechanisms used to implement those rules. To ensure the creation and maintenance of authentic copies of a creator's records, every record preservation system should include in its design records information, selection (appraisal), acquisition, description, storage, retrieval and access systems incorporating a preservation metadata scheme, a records transfer registration scheme, an accession registration scheme, a controlled preservation vocabulary, access privileges and procedures for assessing and maintaining the authenticity of records, along with a rigorous monitoring function to oversee the activities of the entire system.

*Designated Records Preserver*: The entity responsible for taking physical and legal custody of and preserving (i.e., protecting and ensuring continuous access to) authentic copies of a creator's inactive records.[1] Be it an outside organization or an in-house unit, the role of the designated records preserver should be that of a *trusted custodian* for a creator's records. To be considered as a trusted custodian, the preserver must:

- act as a neutral third party, that is, demonstrate that it has no stake in the content of the records and no reason to alter records under its custody, and that it will not allow anybody to alter the records either accidentally or on purpose;

[1] In the InterPARES *Terminology Database,* the term "preservation" is defined as "The whole of the principles, policies, rules and strategies aimed at prolonging the existence of an object by maintaining it in a condition suitable for use, either in its original format or in a more persistent format, while leaving intact the object's intellectual form." See InterPARES *Terminology Database* online at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

- be equipped with the knowledge and skills necessary to fulfill its responsibilities, which should be acquired through formal education in records and archives administration; and
- establish a trusted record preservation system that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved.

InterPARES 1 identified the necessary characteristics of digital records as follows:

*Medium*: The physical material or substance upon which information can be or is recorded or stored. The medium is the material support carrying the record; it is necessary to the existence of the record, which must be affixed to a support, but is a component of the technological context, not of the record.

*Documentary Form*: The rules of representations according to which the content of a record, its administrative and documentary context, and its authority are communicated. Form possesses both extrinsic and intrinsic elements.

*Archival Bond*: The network of relationships that each record has with the records belonging in the same records aggregation, whether inside or outside the system.

*Five Necessary Persons*: Author (the person competent for issuing the record), writer (the person competent for the articulation of the content), originator (the person responsible for the electronic space from which the record is sent), addressee (the person for whom the record is intended) and creator (the person in whose fonds the record exists)**.**

*Act*: An action in which the record participates or which the record supports that is intended to create, maintain, change or extinguish a situation.

*Five Necessary Contexts*: Juridical-administrative (the legal-administrative environment in which the record is created, that is, made or received and set aside for action or reference), provenancial (the creator, its mandate or mission, functions, and organization), procedural (the procedure in the course of which the record is created), documentary (the fonds to which the record belong and its internal structure), technological (the technological environment in which the record was made or received, maintained and preserved)

*Content*: The message contained in the body of the record or which the record is meant to convey.

*Element of Form*: A constituent part of the record's documentary form, visible on the face of the record. It may be either extrinsic, like a seal, or intrinsic, like a subscription.

*Record Metadata*: The attributes of the records that demonstrate their identity and integrity (authenticity).

InterPARES 2 has further clarified:

*Digital Component*: A digital object that is part of one or more digital documents, and the metadata necessary to order, structure or manifest its content and form, requiring a given preservation action.

InterPARES 2 has further distinguished:

*Digital Encoding*: The use of discrete numeric values (such as the binary values 0 and 1) rather than a continuous spectrum of values (such as those generated by an analogue system).

*Digital Data*: The smallest meaningful units of information, expressed as binary bits that are digitally encoded and affixed to a digital medium.

*Bitstream*: Digital data encoded in an unstructured sequence of binary bits that are transmitted, stored or received as a unit.

*Digital Object*: A discrete aggregation of one or more bitstreams and the metadata about the properties of the object and, if applicable, methods of performing operations on the object.

*Digital Document*: A digital component, or group of digital components, that is saved and is treated and managed as a document.

*Digital Record*: A digital document that is treated and managed as a record.

*Stored Digital Document*: A digital document that is placed in a storage system on a digital medium and is treated and managed as a document.

*Stored Digital Record*: A stored digital document that is treated and managed as a record.

*Manifested Digital Record*: A digital record that is visualized or rendered from a stored digital record and/or stored digital component(s) in a form suitable for presentation either to a person (i.e., in human readable form) or to a computer system (i.e., in machine language). Sometimes it does not have a corresponding stored digital record, but is (re)created from fixed content data when a user's or system's actions associates them with specific form data and composition data following fixed rules (e.g., a record produced from a relational database)

*Prospective Record*: A document that guides what to do and/or how to do it and is managed as a record. It differs from a *retrospective record*, which constitutes either the enactment or the memorial of an action, and includes dispositive, probative, supporting and narrative records.[2] A prospective record enables or at least informs interactions, experiences or dynamic processes. It can be an instructive record or an enabling record.

*Instructive Record*: A subclass of prospective record that contains instructions about executing an action or process. With dispositive, enabling, narrative, probative and supporting, one of six functional categories of records. Examples of instructive records are musical scores, regulations, manuals of procedures, and instructions for filling out forms.

*Enabling Record*: A subclass of prospective record encoded in machine language that is actively involved in carrying out an action or process. With dispositive, instructive, narrative, probative and supporting, one of six functional categories of records. Examples of enabling records include software patches that enable a musical instrument to interact with a computer, software in online marketing sites that interprets data about a visitor's actions on the site to determine what elements of content should be presented next to that visitor, and software agents that enable interacting business applications to execute transactions autonomously.

[2] Dispositive Record: A record constituting written evidence of a juridical act; Probative Record: A record for which the juridical system requires a written form as evidence of an action that came into existence and was complete before being manifested in writing; Supporting Record: A record constituting written evidence of an activity that does not result in a juridical act; Narrative Record: A record constituting written evidence of an activity that is juridically irrelevant.

InterPARES 2 has also identified two types of digital objects:

*Static Digital Objects*: Digital objects that do not provide possibilities for changing their manifest content or form beyond opening, closing and navigating. They are records when they present all the characteristics of a record listed above. Examples are: e-mail, reports, sound recordings, video, snapshots of Web pages.

*Interactive Digital Objects*: Digital objects that present variable content, form, or both and the rules governing the context and form of presentation may be either fixed or variable. They are records only if they have a fixed form and a stable content, or can be considered having a bounded variability.

*Fixed Form*:

- The 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 but same documentary presentation).
- If the same content can be presented on the screen in several different ways in a limited series of possibilities, we have a different manifested view of the same stored digital record having stable content and fixed form (different documentary presentations but same digital presentation).

*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.

*Bounded Variability*: When 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 stable content, due to the intention of the author or to different operating systems or applications.

InterPARES 2 also developed Ontologies presenting the basic concepts in relationship with each other ("Ontology A-Concept of a Record," "Ontology B-Concept of the Status of Transmission of a Record," "Ontology C-Concept of the Trustworthiness of a Record;" see: http://www.interpares.org/ip2/display_file.cfm?doc=ip2_ontology.pdf).

## Guiding Methodological Principles

*Interdisciplinarity, Multidisciplinarity and Transdisciplinarity*
The Project is interdisciplinary in the measure in which its goal and objectives can only be achieved through the contribution of several disciplines, integrating methodologies, concepts, principles and techniques from a variety of fields as needed.

This process is very different from the multidisciplinary one, which tends to examine the same problem in the context of each separate discipline and to solve it within such discipline, without any integration of theory or methods, after which the results are compared and the best solutions adopted.

In contrast, transdisciplinarity is multi-referential and multi-dimensional. Whereas interdisciplinarity involves the transfer of one or more methods or ideas from a discipline to another, and

multidisciplinarity involves the analysis of the same object by many disciplines, transdisciplinarity, as the prefix "trans" indicates, involves thinking at the same time within, across and outside each discipline and beyond all disciplines. Its purpose is to gain an understanding of present reality, one imperative of which is the unity of knowledge. "*Rigor*, *openness*, and *tolerance* are the fundamental characteristics of the transdisciplinary attitude and vision. *Rigor* in argument, taking into account all existing data, is the best defense against possible distortions. *Openness* involves an acceptance of the unknown, the unexpected and the unforeseeable. *Tolerance* implies acknowledging the right to ideas and truths opposed to our own."[3]

*Transferability*
The ultimate goal of the Project is archival in nature, in that it is concerned with the development of trusted record-making and recordkeeping systems and of preservation systems that ensure the authenticity of the records under examination over the long term. This implies that the work carried out throughout the Project in the various disciplinary areas must be constantly translated in archival terms and linked to archival concepts, which are the foundation upon which the systems intended to protect the records are designed. However, upon completion of the research, the archival systems need to be made accessible and comprehensible to records creators, organizations and institutions and disciplinary researchers. In other words, the research outcomes must be translated back into the language and concepts of each discipline that need to make use of them.

*Open Inquiry*
InterPARES 3 espouses no epistemological perspective or intellectual definitions a priori. Instead, researchers in each case or general study will identify the perspective(s), research design, and methods that they believe to be most appropriate to their inquiry. In fact, InterPARES 3 is conceived to work as a "layered knowledge" environment, in the sense that some of the research work will build upon knowledge developed in the course of InterPARES; some will take knowledge of similar issues developed in other research projects and in other areas of endeavour and bring it to bear on records creation and preservation; and some will explore new issues and study entities never examined before and develop entirely new knowledge.

*Multi-method Design*
Although the overall methodological approach is action research, each case study and general study will be carried out using the method and the tools that the dedicated investigating team considers the most appropriate for it. The methods used are surveys, case studies, modeling, prototyping, diplomatic and archival analysis, and text analysis. The research is to be guided by the research questions and by instruments developed by the International Team.

[3] See *Charter of Transdisciplinarity* online at http://nicol.club.fr/ciret/english/charten.htm (accessed on 01/02/2007). Every transdisciplinary project is by definition also disciplinary, interdisciplinary and multidisciplinary.