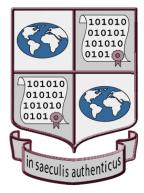
# **InterPARES Project**

International Research on Permanent Authentic Records in Electronic Systems



#### Continuity and Transformation in the Role of the Archivist

Luciana Duranti Beijing, 16 April 2010



## The Goal of InterPARES 1 & 2

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.



#### **Policy Framework**

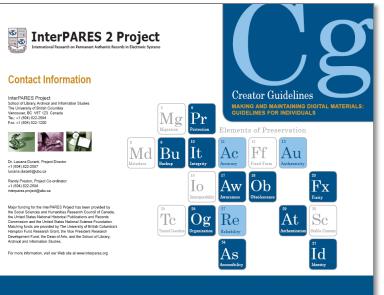
A framework of principles guiding the development of policies for records creating and preserving organizations

	erPARES 2 Project		opuction
Charles Internet		STR	ICTURE OF THE PRINCIPLES
		Dem	CIPLES FOR RECORDS CREATORS
			Digital objects must have a stable content and a fixed documentary form to be considered records
			and to be capable of being preserved over time. (P5)
Tieles	A Framework of Principles for the	(C2)	Record creation procedures should ensure that digital components of records can be separately maintained and reassembled over time. (P4)
Title.		(C3)	Record creation and maintenance requirements should be formulated in terms of the purposes the
	Development of Policies, Strategies and		records are to fulfil, rather than in terms of the available or chosen record-making or recordskeeping technologies. (P6)
	Standards for the Long-term Preservation of	(C4)	
	Digital Records	(05)	record reliability, accuracy and authenticity expressly and separately. (P2). A trusted record-making system should be used to generate records that can be presumed reliable.
	-		A trusted recorditeeping system should be used to maintain records that can be presumed accurate
		(07)	and authentic. (P11, P12) Preservation considerations should be embedded in all activities involved in record creation and
		(07)	maintenance if a creator wishes to maintain and preserve accurate and authentic records beyond its operational business needs. (P7)
		(C8)	A trusted custodian should be designated as the preserver of the creator's records. (P1)
Status:	Final (public)	(C9)	All business processes that contribute to the creation and/or use of the same records should be explicitly documented. (P10)
Version:		(C10	Third-party intellectual property rights attached to the creator's records should be explicitly identified
		(011	and managed in the record-making and recordkeeping systems. (P8).
Submission Date:	June 2005		protected in the record-making and recordkeeping systems. (P9)
Release Date:	March 2008	(012	Procedures for sharing records across different jurisdictions should be established on the basis of the legal requirements under which the records are created. (P13)
Author:	The InterPARES 2 Project	(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 record section activities, have the same effects as the first
Writer(s):	Luciana Duranti, Jim Suderman and Malcolm Todd		manifestation, and each is to be considered at any given time the record of the creator. (P3)
Project Unit:	Policy Cross-domain	PRIN	CIPLES FOR RECORDS PRESERVERS
		(P1) (P2)	A designated records preserver fulfits the role of trusted custodian. (C8) Records preservation policies, strategies and standards should address the issues of record
URL:	ip2(pub)policy_framework_document.pdf		accuracy and authenticity expressly and separately. (C4)
	ipa(pab)pointy_inamework_docament.par	(P3)	Reproductions of a creator's records made for purposes of preservation by their trusted custodian are to be considered authentic copies of the creator's records. (C13)
		(P4)	Records preservation procedures should ensure that the digital components of records can be separately preserved and reassembled over time. (C2)
			separately preserved and reassembled over time. (C2) Authentic copies should be made for preservation purposes only from the creator's records; that is.
		(PO)	from digital objects that have a stable content and a fixed documentary form. (C1)
		(P6)	Preservation requirements should be articulated in terms of the purpose or desired outcome of preservation, rather than in terms of the specific technologies available. (C3)
		(P7)	Preservation considerations should be embedded in all activities involved in each phase of the
		(P8)	records iffecycle if their continuing authentic existence over the long term is to be ensured. (C7) Third-party intellectual property rights attached to the creator's records should be explicitly identified
		(P9)	and managed in the preservation system. (C10) Privacy rights and obligations attached to the creator's records should be explicitly identified and
			protected in the preservation system. (C11)
		(P10	Archival appraisal should identify and analyze all the business processes that contribute to the creation and/or use of the same records. (C9).
			Archival appraisal should assess the authenticity of the records. (C6)
		(P12	<ul> <li>Archival description should be used as a collective authentication of the records in an archival fonds. (C6)</li> </ul>
		(P13	
		(P13	



#### **Creator Guidelines**

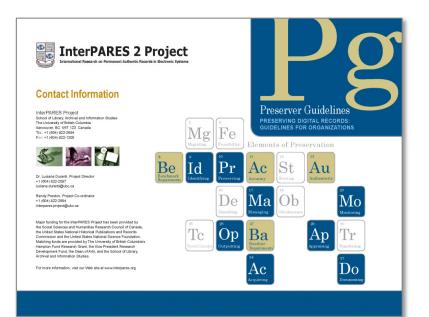
Recommendations for making and maintaining digital materials for individuals and small communities of practice





#### **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<sup>a</sup>
  - name of writer<sup>b</sup> (if different from the author)
     name of originator<sup>c</sup> (if different from name of auth
- name of addressee<sup>d</sup>
   A 1 a ii Name of action or matter
- A.1.a.iii Date(s) of creation and transmission, that is: - chronological date<sup>e</sup> - received date<sup>f</sup>
  - archival date<sup>9</sup>
     transmission date(s)<sup>h</sup>
- A.1.a.iv Expression of archival bond<sup>1</sup> (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<sup>k</sup> (if different t
  - A.1.b.iii Indication of types of annotations added to the recor A.1.b.iv Indication of technical modifications<sup>m</sup>

#### REQUIREMENT A.2: Access Privileges The creator has defined and effectively implemented access privilege

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

#### << REQUIREMENT SET A (cont) >>

REQUIREMENT A.3: Protective Procedures: Loss and Corruption of Records The creator has established and effectively implemented procedures to prevent, disc correct loss or corruption of records.

REQUIREMENT A4: Protective Procedures: Media and Technology The creator has established and effectively implemented procedures to guarantee the or identity and integrity of records against media deterioration and across technological c

REQUIREMENT A5: 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 w 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 As: Removal and Transfer of Relevant Documentation If there is a transition of records from active status to semi-active and inactive statu involves the removal of records from the electonic system, the creator has establis effectively implemented procedures determining what documentation has to be remo transferred 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**

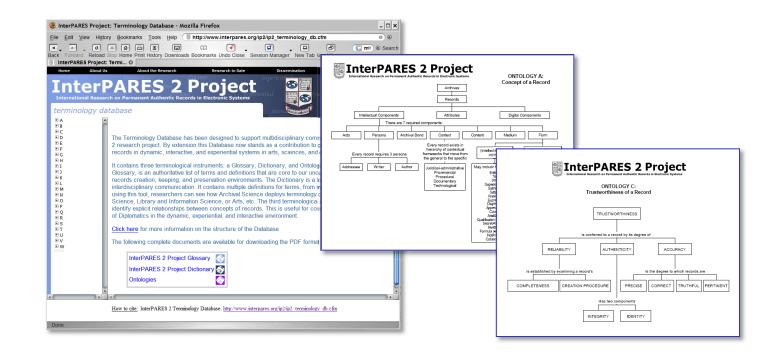
# Principles and criteria for adoption of file formats, wrappers and encoding schemes

Int	erPARES 2 Project	Table of Contents
	ul Resarch on Permanent Authentic Records in Biotronic Systems General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation	Interduction         .1           1. Terminology         .1           1.1 What is a fe format?         .1           1.2 "Depart file formats         .3           1.3 "Standard file formats         .6           1.5 "Standard file formats         .6           2.5 election orders         .6           2. Selection orders         .6           2.1 Mon-proprietary ordgin         .7           2.2 Mon-proprietary ordgin         .7           2.4 Pattorm independence (intercontrolliny)         .9
Status: Version	Final (public)	2.5 Compression     Compr
	March 2007	
	The InterPARES 2 Project	Appendix A: list of repositories reviewed
Writer(s):		
Project Unit:		Bibliography
	http://www.interpares.org/display_file.cfm?doc= ip2_file_formats(complete).pdf [English]	
	http://www.interpares.org/display_file.ofm?doo= ip2_file_formats_fichiers_numériques.pdf [French]	



#### **Terminology Database**

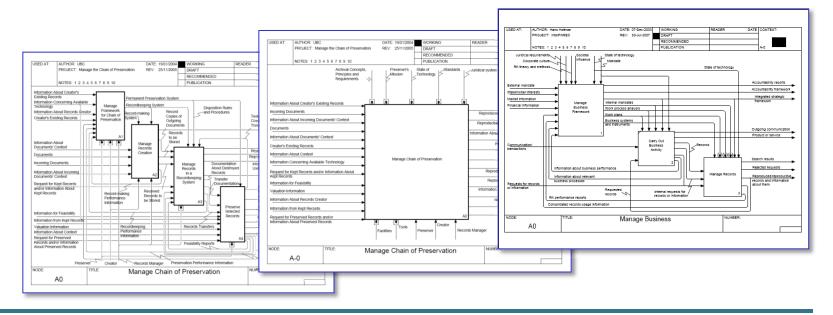
#### Including a glossary, a dictionary and ontologies





#### **Two Records Management Models**

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





#### 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



### **Most Important Findings**

#### Conceptual

- The Concept of Record
- The Concept of Trustworthiness
- The Concept of Life Cycle
- Methodological
  - Appraisal
  - Preservation Concept and Procedure
- Strategic
  - Relationship Creator-Preserver
  - The Role of the Archivist



# New Concept of Records Life-Cycle

We cannot maintain or preserve digital records, but only the ability to re-produce or re-create them, therefore

- re-productions of digital records, if made by the creator in the course of and for the purposes of its business, are records of the creator, while
- if made by the preserver in the course and for the purposes of archival functions, are authentic copies of the records of the creator



## **Digital Records Life-Cycle**

It comprises **two phases** related to who reproduces the record: whether the creator's or the preserver's

Phase 1: Records of the creator: the re-productions made by the creator in the usual and ordinary course of its activities behave and have to be treated as originals every time they are used and acted upon
Phase 2: Authentic copies of the records of the creator: the re-productions made by the preserver cannot be treated as originals because the creator has never used or acted upon them after re-production, which is made for preservation purposes rather than for

the purposes of the creator's activities.



### As a Consequence...

- The creator can decide at any time to stabilize its own fluid entities and to give them the most useful, accessible, interoperable form, or the form that best serves its present and projected needs
- The preserver can only preserve what it receives from the creator by making an authentic copy of it
- Whether the stabilized record of the creator and its authentic copy made by the preserver are to be considered trustworthy depends on the context in which they are created and used



#### **Preservation**

Involves the creation by the preserver of authentic copies of the records of the creator. Their authenticity is guaranteed by:

- a controlled process of migration of the acquired records to the archives technological environment (always keeping the records also in the format in which they were acquired)
- the accurate documentation of any change that the records undergo during such process and every time that the archives technological environment is upgraded
- the implementation and monitoring of privileges concerning the access, use and reproduction of the records within the archives



# **Preservation (cont.)**

- the establishment of procedures to prevent, discover, and correct loss or corruption of records, as well as
- procedures to guarantee the continuing identity and integrity (i.e. authenticity) of the records against media deterioration and across technological changes; and
- if **authentication** of individual records is required, by the existence of rules determining responsibility for and means of authentication.



### **Archival Description**

Archival description acquires a primary authentication function

- The authentication function of archival description is a collective attestation of the authenticity of the records of a fonds and of all their interrelationships as made explicit by their administrative, custodial and technological history, the illustration of their scope and content, and the hierarchical representation of the records aggregates
- The unique function of archival description is to provide an **historical view of the records and of their transformations** while presenting them as a universe in which each part is subject to the bond of a common provenance and destination



### **Archivist as Trusted Custodian**

The trusted custodian is a person who

- 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;
- But, mostly...



# The Archivist's New Role

- Positions him/herself at the beginning of the record life-cycle, taking the role of "designated" trusted custodian
- 2. Assesses the **authenticity of the records** and **monitors it** throughout their existence
- 3. Identifies the records to be preserved at the moment of their creation and monitors their transformation through time
- 4. Determines the **feasibility of preservation** on the basis of the archives technological capacity



# The Archivist's New Role (cont.)

- 5. Determines a **preservation strategy** independently of technological trends (tries to influence the industry through the adoption of standards, but not vice versa) and maintains a focus on interoperability
- 6. Controls the accuracy of the records after each conversion or migration
- 7. Develops procedures that address issues of intellectual rights and privacy
- 8. Recognizes to archival description a primary authentication function



#### The Archivist's New Functions (cont.)

- 9. Is constantly **involved in research and development projects** similar to those carried out by the industry, addressing questions like the following:
  - What entity constitutes the record in each information system
  - What manifestation of such entity can be regarded as the record
  - How to keep such entities accurate and authentic through time
  - How to enable users to verify such authenticity over time

#### We must abandon the old ways!



### The Old Ways

- Academics conduct research which very few read and even less try to implement, usually unsuccessfully (as demonstrated by research)
- Archival associations establish committees which issue guidelines, usually expressing the minimum common denominator shared by the experiences of the members, rather then research findings
- ISO issues standards under the pressure of groups who need basic guidance and either develop their own (see OAIS) or are ready to adopt the ways of the most involved parties (see RM standard)
- Legislators issue laws too often based on the expertise of IT professionals and without serious consultation with archivists (see the European Directive on digital signatures)
- Governments make technological choices without consulting with archivists
- Archives have to respect often unreasonable laws, implement far too generic standards, and preserve unidentifiable and non preservable material



# Why It Does Not Work

**Technology changes** very rapidly while national and international consensus of any kind is very slow General standards and laws need much adaptation to specific contexts to be implemented **Research results must be translated** in concrete terms to be understood by professionals Research has demonstrated that solutions to digital records preservation are dynamic and specific The financial, technological, and knowledge resources of archives are very different



#### **A Better Way**

Each **archives becomes a locus of research** by establishing a partnership with academics involved in international research, professionals involved in standards development, experts in law and information technology and, most importantly, with the creators of the records that fall under their jurisdiction.

- Each archival association promotes an environment supportive of the archives goal by demonstrating to regulatory and auditing bodies, and policy makers that they ought to embed digital records preservation requirements (not rules) in any activity that they regulate, audit or control.
- This results in 1) the generation of new knowledge, 2) the achievement of action-oriented outcomes, 3) the education of all participants, 4) results that are relevant to the local setting, 5) a sound and appropriate research and development methodology, and 6) and the empowerment of the archives.



### **The Power of Archives**

Each archives will then be able to establish a **policy** for the institution, **strategies** for implementing it, **plans of action** for specific aggregations or types of records, and **detailed procedures**, and to **update** all of the above continuously according to changes in available technologies, records produced, and resource availability

This is what we are doing in InterPARES 3 (2007-2012) in an effort to place the archives at the center of society as a point of reference for any institution, organization, community or person who needs 1) guidance in the creation, maintenance and preservation of its records, 2) a neutral third party to take care of the digital evidence of its activities, or 3) an expert witness who attests the authenticity of digital records presented as evidence in legal proceedings.



#### **Other Reasons for a Third Phase of InterPARES**

A study of the effectiveness of workshop and seminar 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

Wendy M. Duff, M., Amy Marshall, Carrie Limkilde and Marlene van Ballegooie (2006) "Digital Preservation Education: Educating or Networking?" *The American Archivist* 69(1): 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.



#### **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. 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.



# **Expected Products (cont.)**

- 4. Guidelines for addressing digital preservation requirements that apply to specific types of records, but not to other materials.
- **5. Evaluation models for assessing the degree of success**, if any, of the chosen preservation action.
- 6. Cost-benefit models for various types of archival organizations or programs and for various kinds of records and/or systems.
- 7. Ethical models that identify and make explicit the consequences for individuals and society of types of preservation measures or lack thereof.



# **Expected Products (cont.)**

- 8. Training and education modules for preservers, 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.
- **9. 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.



# **3 Primary Components**

#### 1. Research component

(short-term and long-term projects, including case studies related to policy, records or systems, 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)



# **General Methodology**

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



# General Methodology (cont.)

- <u>Ethnography</u>: 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



### **Research Activities To Date**

- **General Studies**: Studies of general issues affecting a multiplicity of organizations and supporting the research of many case studies, such as literature reviews, review of other research projects websites, etc. Each TEAM can take the initiative for a general study considered important only in its context or can propose a study useful to everybody but not needing everybody's contribution, such as the e-mail general study conducted by TEAM Italy and the organizational cultures study conducted by TEAM Canada. The Director or the International TEAM can propose a general study to be carried by several or all TEAMs and relevant to several or all, such as terminology, web 2.0, cloud computing, open source softwares, or the registry system.
- **Case Studies**: Studies of specific cases in the context of each TEAM. They can be of three types: policy, records, systems.



# **Case Study Methodology**

- Identifying the Problem Initially, each test-bed will identify a body of digital material for which a preservation plan has to be developed, be it already in its custody or not. Alternatively, the test-bed identifies a policy need, or a system to be designed and implemented.
- Data Collection Using archival methods, data are collected about the context and limitations of each test-bed. Subsequently, using also interviews, diplomatics, modelling, and text analyses, data are collected either about the specific body of material, its documentary forms, technological constraints, functional or cultural meaning, etc., or about a system requirement or policy needs and constraints. In order to be consistent across the TEAMs, a case study workflow has been jointly prepared, templates have been created for contextual analysis and diplomatic analysis, and three sets of questions related to records, policy, and systems have been developed to guide the interviews.
- **Development of Solutions** -- In the course of plenary workshops
- Testing solutions and adjusting them according to feedback



## **InterPARES 3 Composition**

#### International Alliance

15 regional, national & multinational TEAMs:

TEAM (Theoretical Elaboration into Archival Management) Canada (including US); Africa; Brazil; Catalonia; China; Colombia; Italy; Korea; Malaysia; Mexico; Netherlands & Belgium; Norway; Singapore; Turkey; and UK & Ireland

#### Director: Luciana Duranti

Headquarters: UBC - SLAIS (facilities provided by UBC)

Summits: Twice a year, each time hosted by a different country

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



#### InterPARES 3 Web Site



#### www.interpares.org

æ

B

O objetivo geral do InterPARES 3 é capacitar as organizações arquivísticas públicas e privadas de médio e pequeno porte, bem como programas, que são responsáveis por documentos arquivísticos digitais resultantes de atividades de governo, dos negócios, da pesquisa, da arte e do entretenimento, sociais e/ou comunitárias, a preservar ao longo do tempo documentos arquivísticos autênticos, que satisfaçam as necessidades da instituição e da sociedade de ter um registro confiável de seu passado.

- 1. propiciar um ambiente que apóie o objetivo geral da pesquisa, demonstrando aos organismos reguladores e de auditoria e aos responsáveis pela elaboração de políticas que é essencial integrar os requisitos de preservação digital em todas as atividades que são por eles reguladas, auditadas ou
- 2. colaborar com arquivos de pequeno e médio porte no desenvolvimento de políticas, estratégias, procedimentos e/ou planos de ação escaláveis, que possam ser implementados com o objetivo de preservar o material digital sob sua guarda, ou que serão recebidos, por eles, utilizando recomendações



- 🗆 × 00

G Ø

ers' Area

#### InterPARES Web Site

#### www.interpares.org

