# The Preservation of Trustworthy Digital Records: 25 Years of InterPARES Research.

Luciana Duranti, InterPARES Principal Investigator
Professor Emerita, School of Information
The University of British Columbia, Vancouver, Canada

Symposium CARLI, Northern Illinois University
15 April 225

# When digital records happened…

- **The Original Question**: is the theory of the record still **valid in the digital environment**?
- In 1993 I received a grant from the Social Sciences and Humanities Research Council of Canada to find out.
- Soon after, the US Department of Defense (DOD) sought me out to guide them in the design of a record-making and record–keeping system that they could trust: "we offer you the opportunity to demonstrate that theory works in practice."
- I accepted. The outcome of the shared research project was the **DOD 5015.2 standard** for recordkeeping, issued in 1997 and reaffirmed in 2007. It became the foundation of the ISO standard for recordkeeping and later of MoReq.
- **Then, the next Question was**: now that we can keep reliable and authentic records, **how do we maintain those characteristics over the long term?**

InterPARES TrustAI

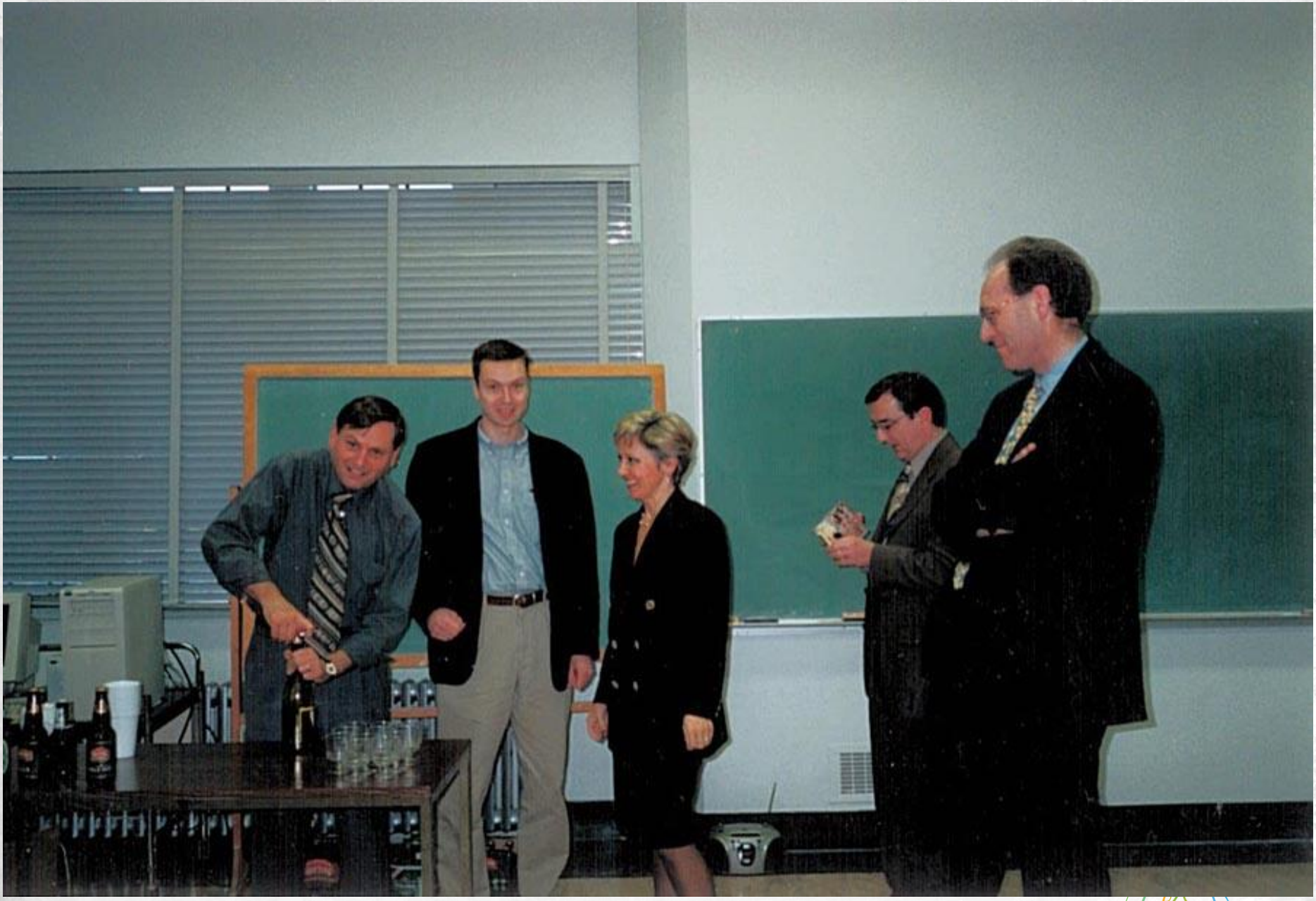
# InterPARES (1998 – 2026)

- To find out we needed **Original Archival Research** conducted **by archivists** and grounded on **archival diplomatics theory**.
- Thus, I created an archival research network where none existed before and called it **inter pares = among peers** to refer to the fact that all members of the network had equal input in the development of new knowledge.
- To express the purpose of the research, I made of *inter pares* an acronym: **International research on Permanent Authentic Records in Electronic Systems**: **InterPARES**.
- The general goal of all InterPARES projects has remained through the years to ensure that digital records be created reliable and accurate and preserved authentic, for as long as needed by their creator and society at large, across technological changes.

**InterPARES has been funded by SSHRC grants and matching funds from UBC and all the partners (in cash and/or in kind) since 1998**

InterPARES Trust AI

InterPARES TrustAI



# The Goal of InterPARES 1 and 2 (1998-2006)

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.

InterPARES TrustAI

# Goal of InterPARES 3 (2007-2012)

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

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

# InterPARES Concepts

- **Manifested record** (what you see) and **stored records** (the encoding: content data, form data, composition data)
- **Static records** (do not provide possibilities for changing their manifest content or form beyond opening, closing and navigating: e-mail, reports, sound recordings, motion video, snapshots of web pages) and **Interactive records** (present variable content, form, or both, and the rules governing the content and form of presentation may be either fixed or variable
- **It is not possible to preserve a digital record**. It is only possible to preserve our ability to reproduce it or recreate it, preserving its trustworthiness
- **Trustworthiness**: reliability, accuracy, authenticity

InterPARES TrustAI

# Concepts: Trustworthiness

## Reliability

The trustworthiness of a record as a statement of fact,

*based on:*

- the competence of its author (the person who has the authority and capacity to issue it)
- the controls on its creation

## Accuracy

The correctness and precision of a record's content

*based on:*

- the competence of its author
- the controls on content recording and transmission

## Authenticity

The trustworthiness of a record that is what it purports to be, untampered with and uncorrupted

*based on:*

- identity
- Integrity
- reliability and integrity of the system

InterPARES TrustAI

# Concepts: Identity

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

- the **names** of the persons concurring in its creation (i.e., author, addressee, writer, originator, creator);
- its **date(s)** of creation (i.e. making, receipt, filing) and transmission;
- the matter or **action** in which it participates;
- the expression of its **relationships** with other records (e.g. classification code); and
- an indication of any **attachment(s)**

On this basis InterPARES identified the necessary **identity metadata**

InterPARES TrustAI

# Concepts: Integrity

*Integrity refers* to the quality of being complete and unaltered in all essential respects.

We need information that helps infer that the record is the same as when created (**integrity metadata** or properties):

- name(s) of **handling persons over time**
- name of person primarily **responsible for keeping** the record
- indication of additions (**annotations**) made to the record
- indication of **technical changes** (e.g. format, encoding, upgrading, changes to digital components, migration)
- indication of presence or removal of a **digital signature**
- **planned removal from the system**, by destruction or transfer to a custodian
- existence and location of **duplicates** outside the system

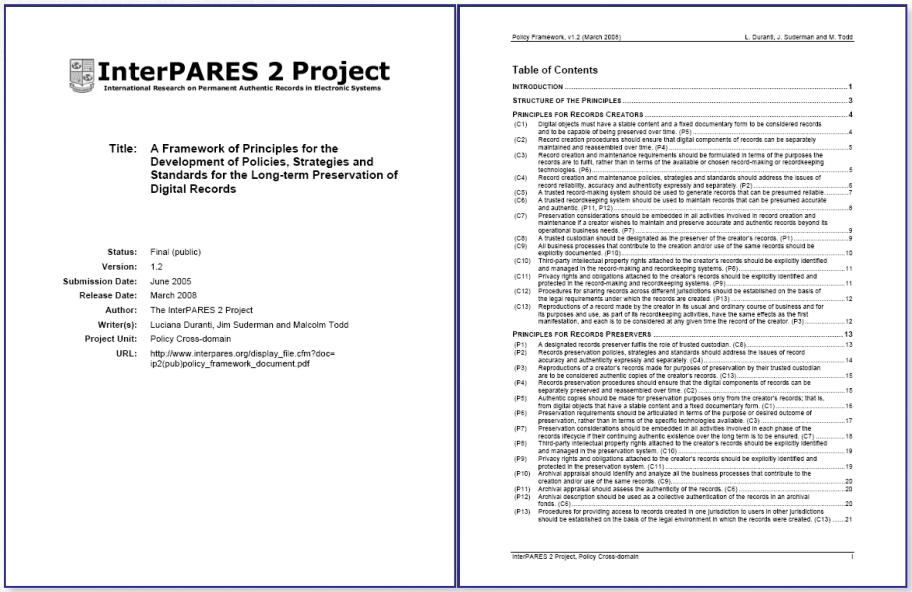
The last two bullets establish a proper **chain of custody**

InterPARES TrustAI

# Key IP 1 & 2 Products

## Policy Framework

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



InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

**Title:** A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records

| | |
|---|---|
| **Status:** | Final (public) |
| **Version:** | 1.2 |
| **Submission Date:** | June 2005 |
| **Release Date:** | March 2008 |
| **Author:** | The InterPARES 2 Project |
| **Writer(s):** | Luciana Duranti, Jim Suderman and Malcolm Todd |
| **Project Unit:** | Policy Cross-domain |
| **URL:** | http://www.interpares.org/display_file.cfm?doc=ip2(pub)policy_framework_document.pdf |



Table of Contents


INTRODUCTION ..... 1
STRUCTURE OF THE PRINCIPLES ..... 3
PRINCIPLES FOR RECORDS CREATORS ..... 4
(C1) 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) ..... 4
(C2) Record creation procedures should ensure that digital components of records can be separately maintained and reassembled over time. (P4) ..... 5
(C3) Record creation and maintenance requirements should be formulated in terms of the purposes the records are to fulfil, rather than in terms of the available or chosen record-making or recordkeeping technologies. (P6) ..... 5
(C4) Record creation and maintenance policies, strategies and standards should address the issues of record reliability, accuracy and authenticity expressly and separately. (P2) ..... 6
(C5) A trusted record-making system should be used to generate records that can be presumed reliable. ..... 7
(C6) A trusted recordkeeping system should be used to maintain records that can be presumed accurate and authentic. (P11, P12) ..... 8
(C7) Preservation considerations should be embedded in all activities involved in record creation and maintenance if a creator wishes to maintain and preserve accurate and authentic records beyond its operational business needs. (P7) ..... 9
(C8) A trusted custodian should be designated as the preserver of the creator's records. (P1) ..... 9
(C9) All business processes that contribute to the creation and/or use of the same records should be explicitly documented. (P10) ..... 10
(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) ..... 11
(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) ..... 11
(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) ..... 12
(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 as the first manifestation, and each is to be considered at any given time the record of the creator. (P3) ..... 12
PRINCIPLES FOR RECORDS PRESERVERS ..... 13
(P1) A designated records preserver fulfils the role of trusted custodian. (C8) ..... 13
(P2) Records preservation policies, strategies and standards should address the issues of record accuracy and authenticity expressly and separately. (C4) ..... 14
(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) ..... 15
(P4) Records preservation procedures should ensure that the digital components of records can be separately preserved and reassembled over time. (C2) ..... 15
(P5) Authentic copies should be made for preservation purposes only from the creator's records; that is, from digital objects that have a stable content and a fixed documentary form. (C1) ..... 16
(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) ..... 17
(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. (C7) ..... 18
(P8) Third-party intellectual property rights attached to the creator's records should be explicitly identified and managed in the preservation system. (C10) ..... 19
(P9) Privacy rights and obligations attached to the creator's records should be explicitly identified and protected in the preservation system. (C11) ..... 19
(P10) Archival appraisal should identify and analyze all the business processes that contribute to the creation and/or use of the same records. (C9) ..... 20
(P11) Archival appraisal should assess the authenticity of the records. (C6) ..... 20
(P12) Archival description should be used as a collective authentication of the records in an archival fonds. (C6) ..... 20
(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. (C13) ..... 21

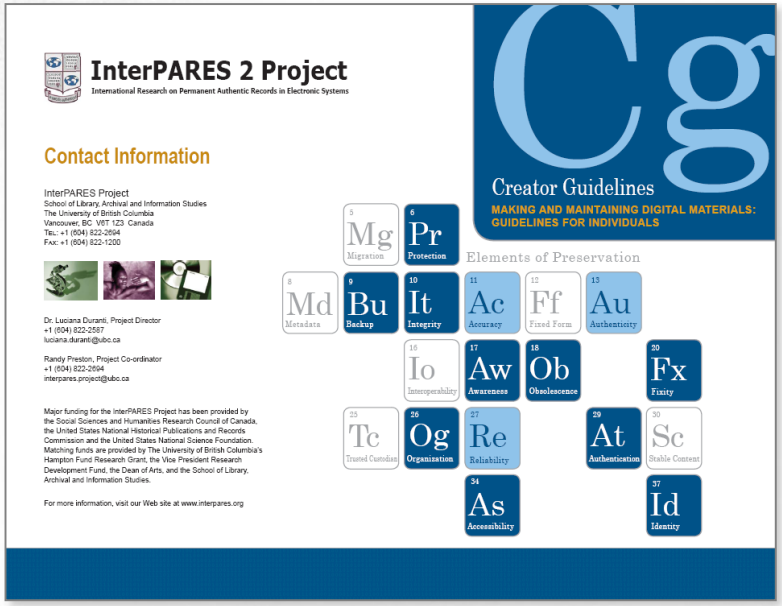



InterPARES TrustAI

# IP 1 & 2 Products

## Creator Guidelines

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

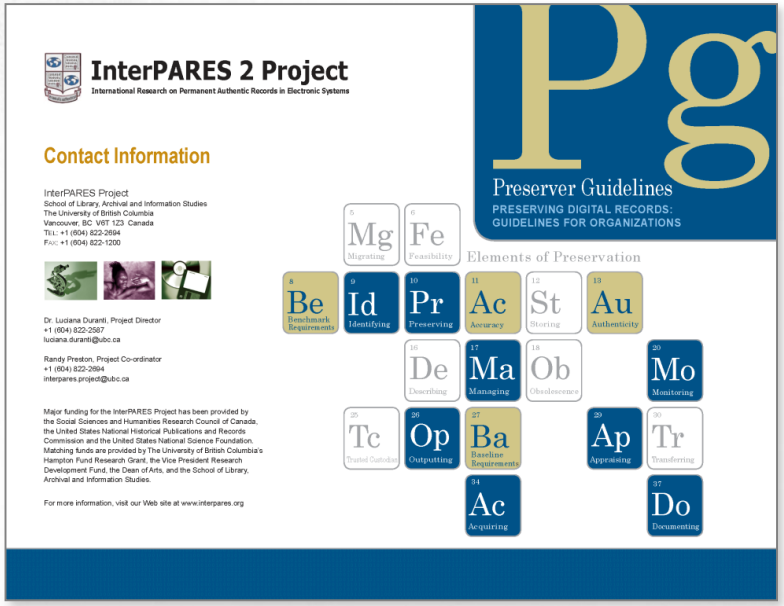



InterPARES TrustAI

# IP 1 & 2 Products

## Preserver Guidelines

Recommendations for digital preservation for archival institutions

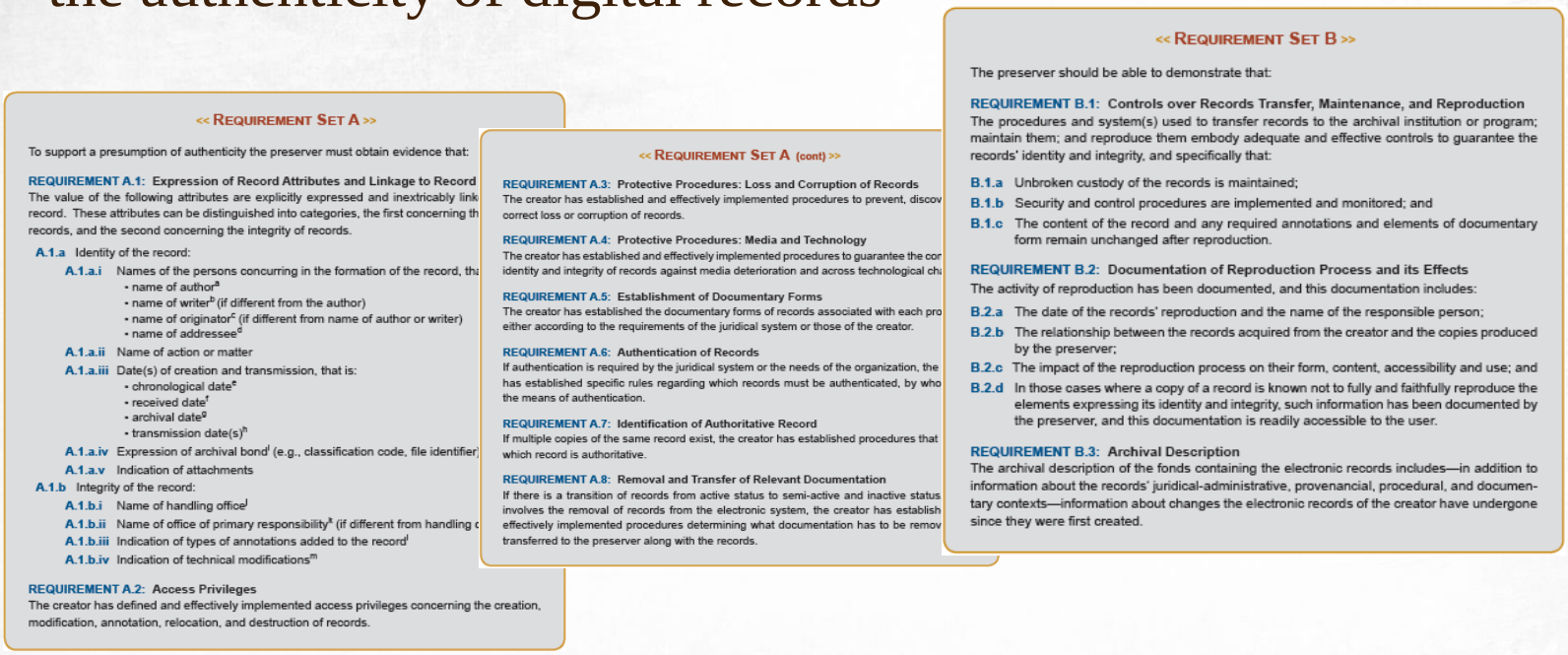



InterPARES TrustAI

# IP 1 & 2 Products

## 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 link record. These attributes can be distinguished into categories, the first concerning th 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 the record, th
  - name of author[a]
  - name of writer[b] (if different from the author)
  - name of originator[c] (if different from name of author or writer)
  - name of addressee[d]
- A.1.a.ii Name of action or matter
- A.1.a.iii Date(s) of creation and transmission, that is:
  - chronological date[e]
  - received date[f]
  - archival date[g]
  - transmission date(s)[h]
- A.1.a.iv Expression of archival bond[i] (e.g., classification code, file identifier)
- A.1.a.v Indication of attachments

A.1.b Integrity of the record:
- A.1.b.i Name of handling office[j]
- A.1.b.ii Name of office of primary responsibility[k] (if different from handling c
- A.1.b.iii Indication of types of annotations added to the record[l]
- A.1.b.iv Indication of technical modifications[m]

**REQUIREMENT A.2: Access Privileges**
The creator has defined and effectively implemented access privileges concerning the creation, 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, discov correct loss or corruption of records.

**REQUIREMENT A.4: Protective Procedures: Media and Technology**
The creator has established and effectively implemented procedures to guarantee the cor identity and integrity of records against media deterioration and across technological ch

**REQUIREMENT A.5: Establishment of Documentary Forms**
The creator has established the documentary forms of records associated with each pro 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, the has established specific rules regarding which records must be authenticated, by who the means of authentication.

**REQUIREMENT A.7: Identification of Authoritative Record**
If multiple copies of the same record exist, the creator has established procedures that which record is authoritative.

**REQUIREMENT A.8: Removal and Transfer of Relevant Documentation**
If there is a transition of records from active status to semi-active and inactive status involves the removal of records from the electronic system, the creator has establish effectively implemented procedures determining what documentation has to be remov 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.

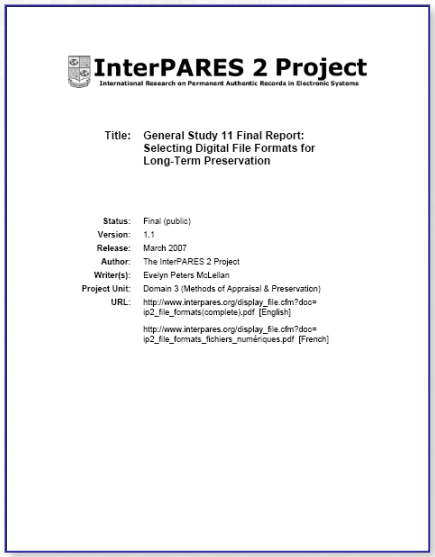
**REQUIREMENT B.3: Archival Description**
The archival description of the fonds containing the electronic records includes—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the electronic records of the creator have undergone since they were first created.

InterPARES TrustAI

# IP 1 & 2 Products

## File Format Selection Guidelines

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



InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

Title: General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation

Status: Final (public)
Version: 1.1
Release: March 2007
Author: The InterPARES 2 Project
Writer(s): Evelyn Peters McLellan
Project Unit: Domain 3 (Methods of Appraisal & Preservation)
URL: http://www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf [English]

http://www.interpares.org/display_file.cfm?doc=ip2_file_formats_fichiers_numériques.pdf [French]

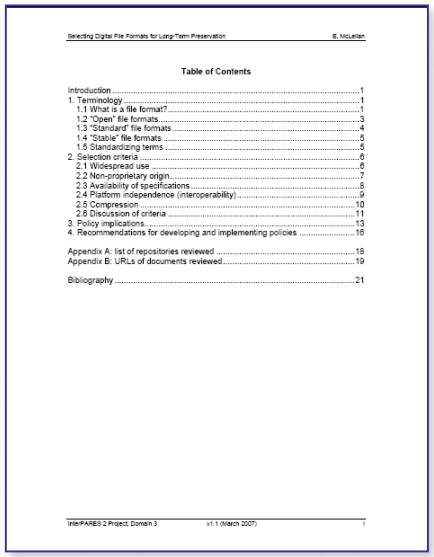




Table of Contents

Introduction ..... 1
1. Terminology ..... 1
1.1 What is a file format? ..... 1
1.2 "Open" file formats ..... 3
1.3 "Standard" file formats ..... 4
1.4 "Stable" file formats ..... 5
1.5 Standardizing terms ..... 5
2. Selection criteria ..... 6
2.1 Widespread use ..... 6
2.2 Non-proprietary origin ..... 7
2.3 Availability of specifications ..... 8
2.4 Platform independence (interoperability) ..... 9
2.5 Compression ..... 10
2.6 Discussion of criteria ..... 11
3. Policy implications ..... 13
4. Recommendations for developing and implementing policies ..... 16

Appendix A: list of repositories reviewed ..... 18
Appendix B: URLs of documents reviewed ..... 19

Bibliography ..... 21

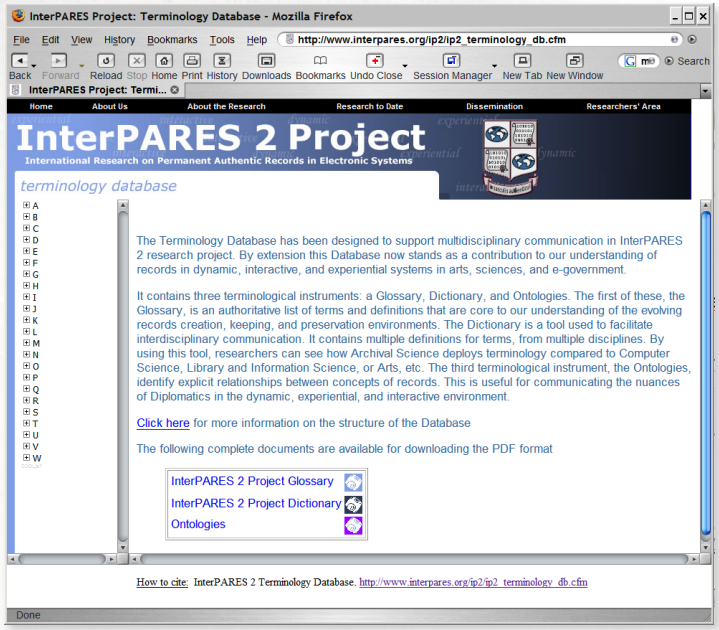


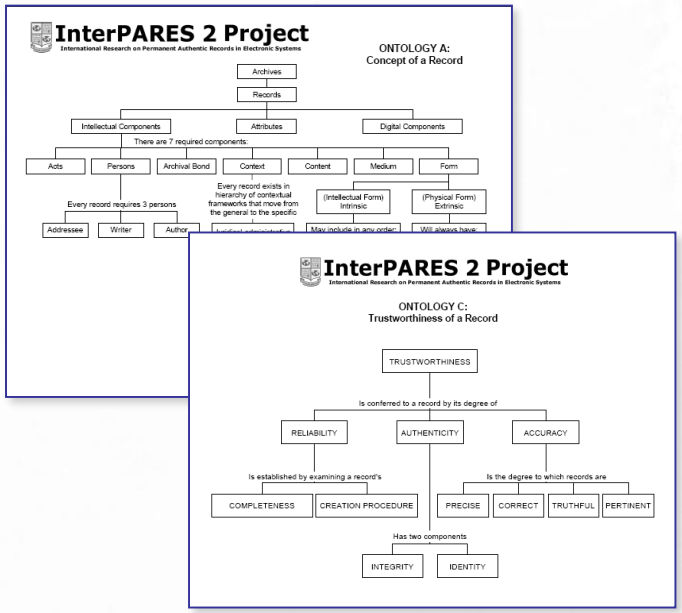
InterPARES TrustAI

# IP 1 & 2 Products

## Terminology Database

Including a glossary, a dictionary and ontologies



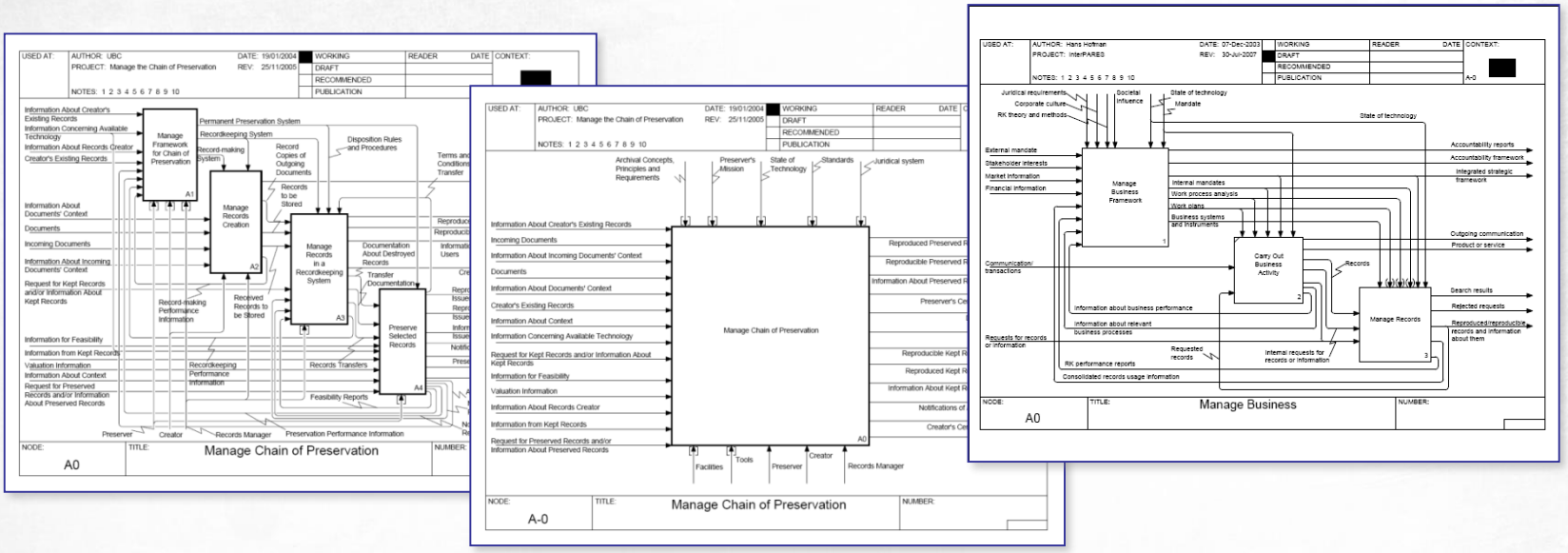

InterPARES TrustAI

# IP 1 & 2 Products

## Two Records Preservation Models

Chain of Preservation (COP) Model (lifecycle)

Business-driven Recordkeeping (BDR) Model (continuum)



# IP 1 & 2 Final Products

**Two books:**

Luciana Duranti, ed. *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (San Miniato: Archilab, 2005). Available on line at http://www.interpares.org/book/index.cfm

Luciana Duranti and Randy Preston, eds. *InterPARES 2: Interactive, Dynamic and Experiential Records* (Roma: ANAI, 2008). Available on line at http://www.interpares.org/ip2/book.cfm.

# InterPARES 3 General Studies

- National Standards Relevant to IP3
- Community Archives e-Records Assessment
- Public Sector Audit Report for Digital Recordkeeping
- Records Management Policies and Procedures Template
- **Cost-benefit Models**
- **Ethical Models**
- File Viewers
- **Education Modules (ICA Education Modules)**
- Open Source Records Management Software
- **Metadata Applications Profiles**
- Organizational Culture & Risk Assessment

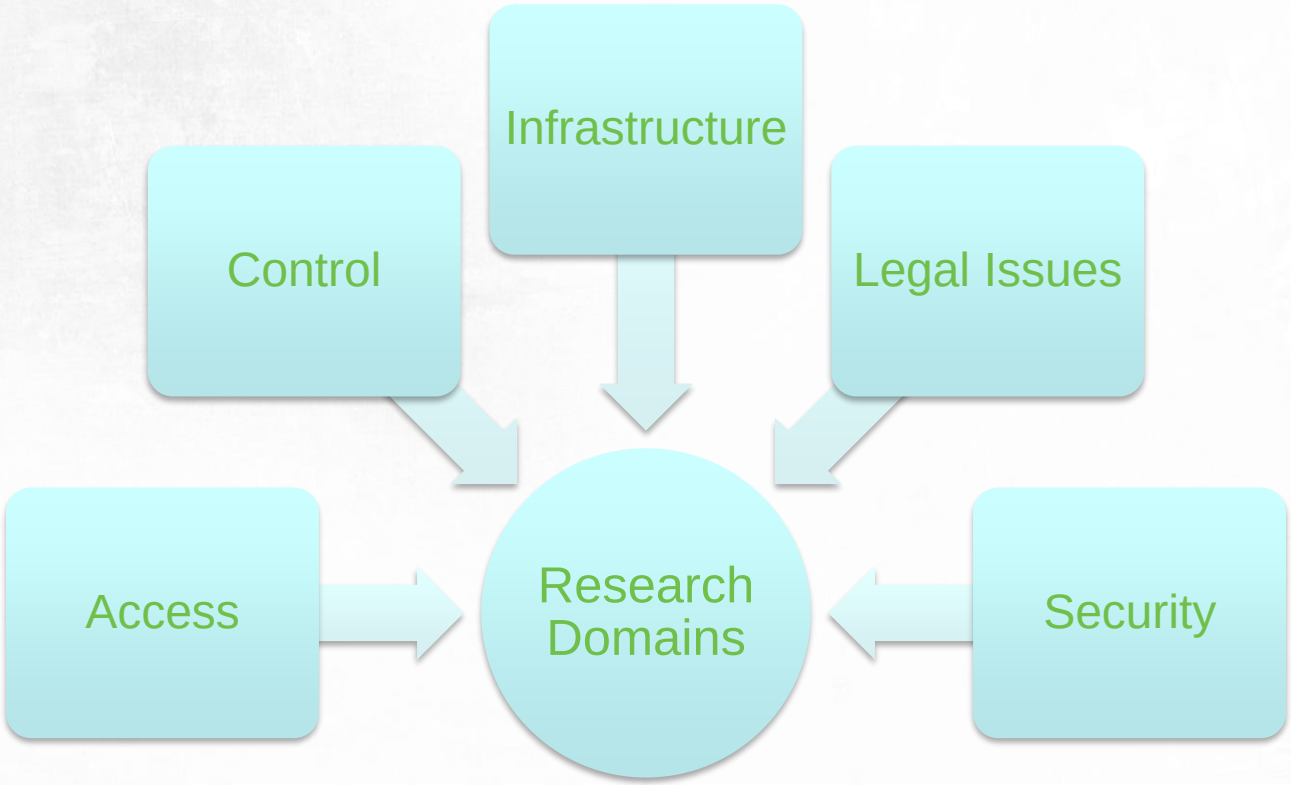
# InterPARES 3 General Studies

- Web 2.0/Social Media
- **Terminology Database (ICA Multilingual Archival Database – now it includes 26 languages)**
- Analysis of Other Digital Preservation Projects
- International Standards Relevant to IP3
- Annotated Bibliography Database
- **E-mail Preservation**
- Preservation of Registries

# InterPARES Trust (2013-2020)

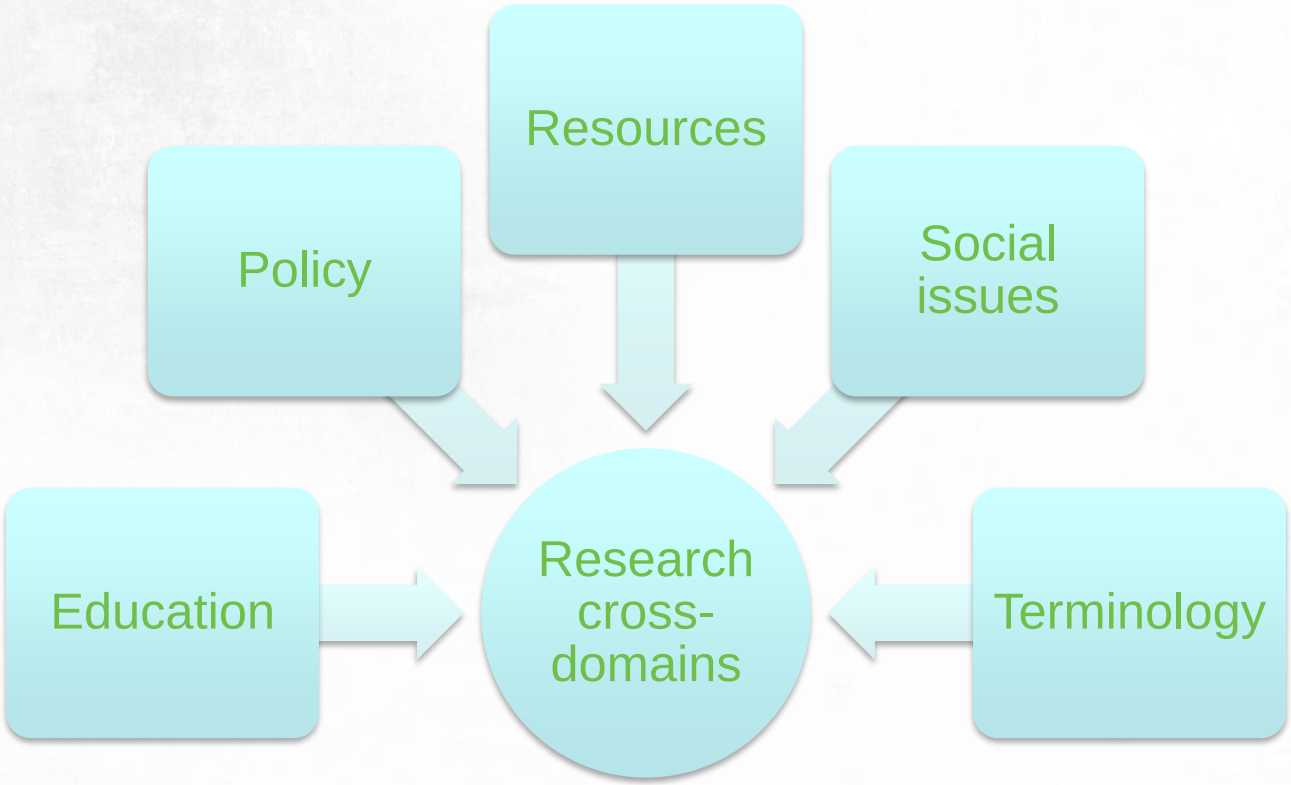
The **goal of InterPARES Trust** was to generate the theoretical and methodological **frameworks** that should support the development of integrated and consistent local, national and international **networks of policies, procedures, regulations, standards and legislation concerning digital records online**, to ensure public trust.

# Research Domains




InterPARES TrustAI

# Research Cross-Domains



# Final I Trust Final Products

- Luciana Duranti and Corinne Rogers eds. ***Trusting Records in the Cloud.*** Facet Publishers and the Society of American Archivists. London, UK and Chicago, Ill. 2019
- Jens Boel and Eng Sengsavang eds., ***Recordkeeping in International Organizations Archives in Transition in Digital, Networked Environments***. London, UK: Routledge, Taylor and Francis, 2021
- Hrvoje Stancic ed., ***Trust and Records in an Open Digital Environment.*** London, UK: Routledge, Taylor and Francis, 2021
- Alicia Barnard ed., ***InterPARES en Latinoamérica y el Caribe 2005-2019.*** Alcaldía Coyoacán CD MX: Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales, 2020.
- Mpho Ngoepe ed., ***Managing Digital Records in Africa.*** London, UK: Routledge, Taylor and Francis, 2022.
- InterPARES and Canadian Institute for Information and Privacy Studies ***Duty to Document***, a documentary accessible on www.interparestrust.org

InterPARES TrustAI

# InterPARES Trust AI (2021-2026)

The goal of the fifth phase of InterPARES, I Trust AI, is to design, develop, and leverage Artificial Intelligence to support the ongoing availability and accessibility of trustworthy public records. We aim to

- Identify specific AI technologies that can address critical records challenges;
- Determine the benefits and risks of using AI technologies on records; and
- Ensure that records concepts and principles inform the development of responsible AI

InterPARES TrustAI

# Participants

- 101 partner organizations in 42 countries (in 5 continents)
- 131 co-applicants  (academics)
- 129 collaborators (professionals)
- 3 postdocs
- 60 Graduate Academic Assistants in any given year

InterPARES TrustAI

# Approach

- The fact that the *I Trust AI* project is a **multinational interdisciplinary endeavour** means that our first effort had to be to **understand each other, starting with the language we use**. For example, archival professionals talk about **records**, while computer scientists and AI professionals talk about **data**. To archivists, data are the smallest meaningful unit of information in a record. To an AI specialist, data is (note: singular) organized information (possibly in a database), be it facts or not, regardless of size, nature and form.
- Thus, key to our work have been **AI tutorials and workshops** for non-AI researchers, and **archival and diplomatics theory tutorials** for non archival researchers. These educational endeavours are supported by the **Terminology Database** which is developed in collaboration by a multidisciplinary team.

InterPARES TrustAI

# Indirect Outcomes

- New Professionals: by the end of the project, there will be well over 100 professionals who will have worked as student research assistants on case studies with test-bed organizations and who will spread the acquired knowledge, without counting all the future professionals taught such knowledge during their course of study
- Students from other disciplines: computer scientists, lawyers, etc. will understand and value the archival perspective in their work and the impact of records and recordkeeping on the broader society
- Knowledge co-creation: the project will enrich research in archival science, records management, AI, cybersecurity, information science, law, and ethics, through knowledge exchange and uptake between scholars and practitioners within and among those disciplines.
- Sensitizing AI developers, scholars, and other members of that community to the role of AI in record keeping and archival preservation and to the role of archival concepts and principles in AI design and development.

InterPARES TrustAI

# Studies

- Studies are **all international and interdisciplinary**
- Focus on all aspects of archival functions
  1. Creation and use of trustworthy records
  2. Appraisal and acquisition of archival material
  3. Arrangement and description
  4. Retention and preservation
  5. Management and administration of records and archives
  6. Reference and access

InterPARES TrustAI

# Case Study: Deep Learning for parchment documents

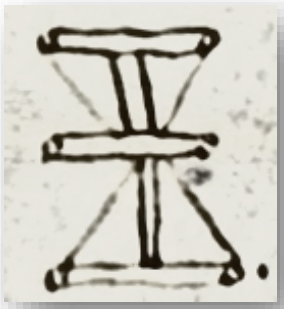
**Purpose**: development of a tool to identify the "identity attributes" of thousands of digitized parchments issued by city notaries in medieval times. Study led by Prof. *Emanuele Frontoni and his team Vision computing and Robotic for AI*

- It uses **computer vision**, a field of AI that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs, and take actions or make recommendations based on that information
- The basic feature chosen to be identified for the research is **the signum, an authentication element** affixed by notaries.

InterPARES TrustAI

# THE SIGNUM: A USEFUL MARKER

- The **signum or notarial sign** is a specific and personally drawn mark used by a single notary in the top part of the record and before his signature at the bottom of the record.
- Identifying the signum means that every notary can be **recognised and tracked** in a virtually infinite series of documents.
- The AI will contribute in creating both a **register of signa—a virtual registry of notaries,** and the basis for investigating the **less visible features of the parchments**

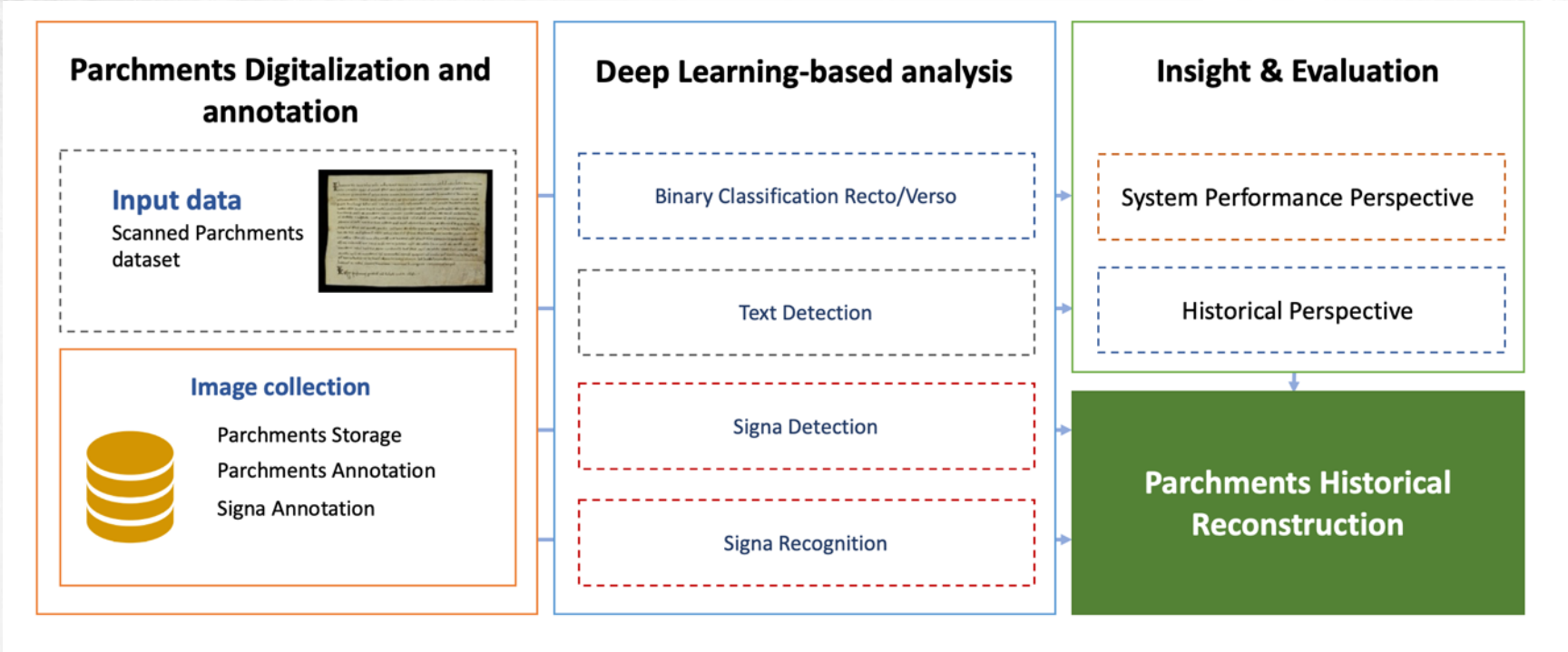


InterPARES TrustAI

# PERGANET

## A Deep Learning Framework for Automatic Appearance-Based Analysis

https://link.springer.com/chapter/10.1007/978-3-031-13324-4_25




InterPARES TrustAI

# OTHER ARCHIVAL APPLICATIONS

Once refined and fully developed, this Deep Learning tool, Perganet, could be used in a **wide range of applications**:

- Recognize the **peculiar system of writing** of individual authors.
- Analyze **archival annotations on the back of the documents** and retrace previous archival arrangements or uses of groups of documents.
- Recognize **recurring images or other features** in huge series of documents.
- Identify **common patterns in manuscript maps or drawings**.
- Make publicly available **original & relevant AI datasets.**
- And many others…

InterPARES TrustAI

# UNESCO Audio Archives

Interviews and reports 1950s-1980s.

~6,500 available on digital platform (16,000 total)

But only ~*800* described so far



InterPARES TrustAI

# UNESCO Archives Languages

70+ recognized languages

French

English

Spanish

*Multilingual (4%)*

InterPARES TrustAI

# Metadata Scheme

57 elements total

**Vital for discoverability**

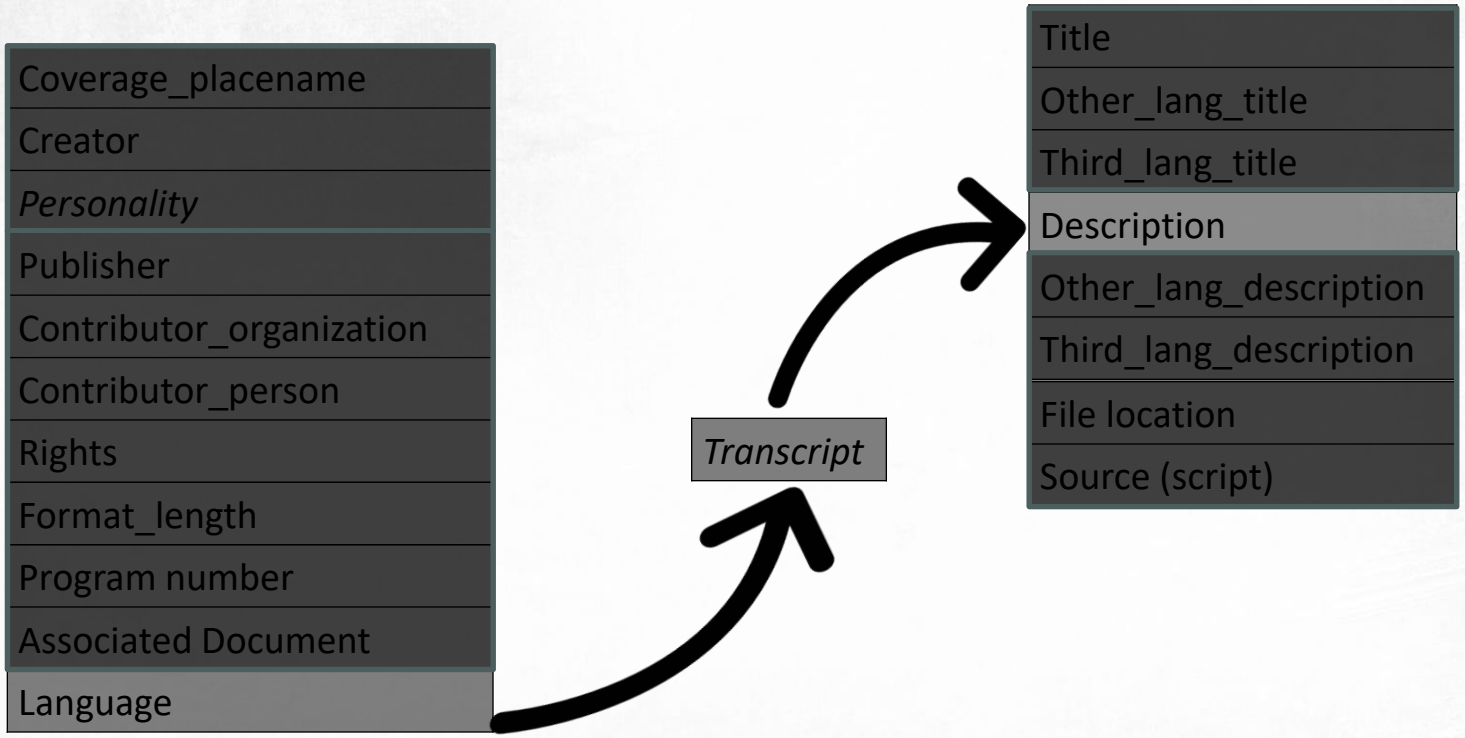
| |
|---|
| Title |
| Other_lang_title |
| Third_lang_title |
| Description |
| Other_lang_description |
| Third_lang_description |
| File location |
| Source (script) |

| |
|---|
| Coverage_placename |
| Creator |
| Personality |
| Publisher |
| Contributor_organization |
| Contributor_person |
| Rights |
| Program number |
| Associated Document |
| Format_length |
| Language |
| Access_category |
| Rightsholder |



InterPARES TrustAI

# UNESCO Archives Metadata Enrichment Plan






InterPARES TrustAI

# Text vs. Speech Language ID

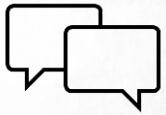
*Text Based*

- *Orthographic clues*
- *Full sentence used as input*

*Models:*

*Transformer Model (BERT etc…)*

*Speech Based*



- *Potential Confounders:*
  - *Speaker*
  - *Recording setup*
- *Short audio clip*

*Models:*

*Convolutional Neural Network (x-vector)*

*Fine-tuned Audio Transformer*

InterPARES TrustAI

# Steps in Metadata Creation

**Speech Transcription**

**Identification of genre** (e.g. interview)

**Diplomatic Labeling** according to: protocol (names of persons, dates, topic in interviews) or eschatocol (in reports) depending on genre of audio.

InterPARES TrustAI

# I Trust AI Paradata Study

**A General Study is one that concerns all archival functions. Purpose** of the Paradata study:

Developing an approach for **documenting the AI process** to fulfill archival accountability in the use of AI, thereby also **supporting the authenticity of the outcome**.

*Researchers: Pat Franks, Babak Hamidzadeh, Scott Cameron, Norman Mooradian, Alex Richmond, Mario Beauchamp*

*The slides that follow are extracted from several of their presentations.

InterPARES TrustAI

# XAI vs. Accountable AI

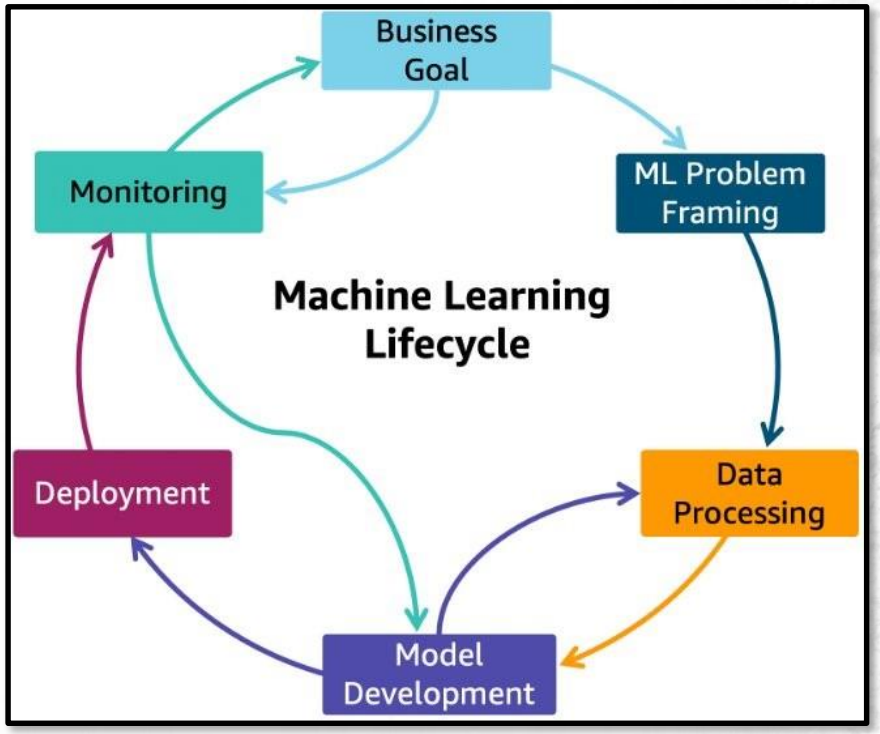
- Explainable AI (XAI) has received a lot of attention. XAI focuses on **why a given tool produced a given output from a given set of inputs**.
- But building accountable AI **must also consider the individuals, organizations, and environment in which the AI tool operates**.
- **Paradata is necessary to explain why, how, by whom, and to what effect a given tool was used in a particular context.**

InterPARES TrustAI

# PARADATA & AI Process

**Paradata** is **information about the procedure(s) and tools** used to create and process information resources, along with **information** about the operation of the tools and the execution of the procedures, and **about the persons** carrying out those procedures and using the tools.

*~ITrustAI working definition*




InterPARES TrustAI

# *Metadata vs Paradata*

**Metadata** is formalized data about a record needed to search for, display, and analyze that record

**Paradata** is formalized data on methodologies, processes, and persons associated with the production and assembly of records.

InterPARES TrustAI

# Examples of Paradata

## Technical Paradata

- AI Model (tested & selected)
- Evaluation & performance metrics
- Logs generated
- Model training data set
- Training parameters for model
- Vendor documentation
- Versioning information

## Organizational Paradata

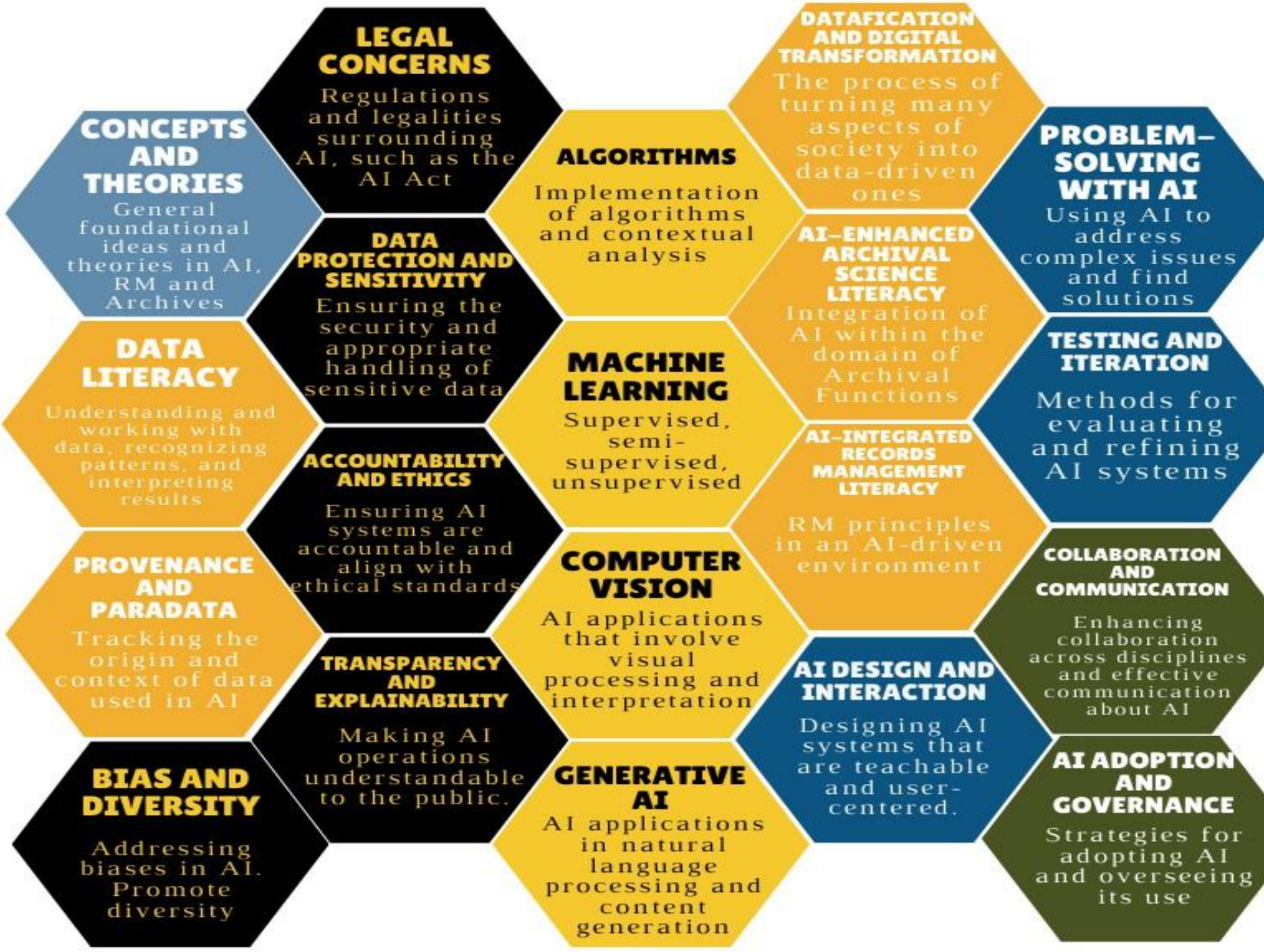
- AI policy
- Design plans
- Employee training
- Ethical consideration
- Impact assessments
- Implementing process
- Regulatory requirements

InterPARES TrustAI

# Other General Studies

- Terminology Database
- Identification of **Personal Private Information** (through diplomatics labeling)
- Providing access without providing access: Using **Privacy Enhancing Technologies** (e.g. distant reading) – Clio X
- **Veracity of images**
- Development of methods to enhance the reliability of LLMs and minimize the occurrence of hallucinations (a promising solution is the use of **retrieval augmented generation** RAG).
- Ensuring Archivists' AI Literacy (Hernandez and Rochenbach)

InterPARES TrustAI


LEGAL CONCERNS
Regulations and legalities surrounding AI, such as the AI Act
DATAFICATION AND DIGITAL TRANSFORMATION
The process of turning many aspects of society into data-driven ones
CONCEPTS AND THEORIES
General foundational ideas and theories in AI, RM and Archives
ALGORITHMS
Implementation of algorithms and contextual analysis
PROBLEM–SOLVING WITH AI
Using AI to address complex issues and find solutions
DATA PROTECTION AND SENSITIVITY
Ensuring the security and appropriate handling of sensitive data
AI–ENHANCED ARCHIVAL SCIENCE LITERACY
Integration of AI within the domain of Archival Functions
DATA LITERACY
Understanding and working with data, recognizing patterns, and interpreting results
MACHINE LEARNING
Supervised, semi-supervised, unsupervised
TESTING AND ITERATION
Methods for evaluating and refining AI systems
ACCOUNTABILITY AND ETHICS
Ensuring AI systems are accountable and align with ethical standards
AI–INTEGRATED RECORDS MANAGEMENT LITERACY
RM principles in an AI-driven environment
PROVENANCE AND PARADATA
Tracking the origin and context of data used in AI
COMPUTER VISION
AI applications that involve visual processing and interpretation
COLLABORATION AND COMMUNICATION
Enhancing collaboration across disciplines and effective communication about AI
TRANSPARENCY AND EXPLAINABILITY
Making AI operations understandable to the public.
AI DESIGN AND INTERACTION
Designing AI systems that are teachable and user-centered.
BIAS AND DIVERSITY
Addressing biases in AI. Promote diversity
GENERATIVE AI
AI applications in natural language processing and content generation
AI ADOPTION AND GOVERNANCE
Strategies for adopting AI and overseeing its use

# Stay tuned!
# Thank you!

www.interpares.org

www.interparestrustai.org

@itrustai

www.facebook.com/interparestrust

InterPARES TrustAI