

InterPARES 2023 SUMMER SCHOOL

Day 1 – *Friday 7 July*
Symposium on I Trust AI

9:00 - 16:30

Day 1 – Friday 7 July – Symposium on I Trust AI

9:00 - 16:30

Artificial Intelligence to support the ongoing availability and accessibility of trustworthy public records - an overview of InterPARES Trust AI

9:00 *Gather and welcome*

9:30 Luciana Duranti & Muhammad Abdul-Mageed, University of British Columbia
The Whys and Hows in “I Trust AI”: Objectives, methods, expected outcomes

10:00 Jason R. Baron, University of Maryland
AI and Freedom of Information Laws: Using AI To Filter For Exempt Materials in Documents

10:30 Jim Suderman, City of Toronto (ret.)
Clear values, murky responsibilities: considering the ethical pipeline of archival information and AI implementation

11:00 Umi Mohktar, Universiti Kebangsaan Malaysia
AI for Records Classification

11:30 Pat Franks, San Jose State University
Paradata: Documenting the AI Process for Transparency and Accountability

12:00 Peter Sullivan, University of British Columbia
Applying AI tools in Archival Functions

12:30 *Lunch*

14:00 Hugolin Bergier, Regis University
Intensional Logic in RiC-O

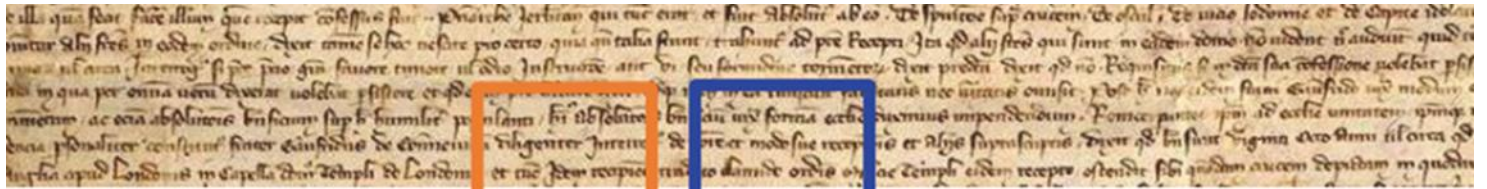
14:30 Emanuele Frontoni & Pierluigi Feliciati
AI for Multimedia understanding: from computer vision to deep learning

15:00 Sanja Seljan, University of Zagreb
Data Acquisition and Corpus Creation for Security-Related Domain

15:30 Moises Rockembach
AI literacy and the future of Records Management and Archives

16:00 Round-table discussion and questions

16:30 Adjourn



InterPARES 2023 SUMMER SCHOOL

Day 2-5 – *July 8-11*
Summer School

9:00 - 16:30



Day 2 - Saturday 8 July

9:00-9:30 - Luciana Duranti, *An overview of the InterPARES research project and its products*

9:30-10:30 - Luciana Duranti, *The Concept of Digital Record*

InterPARES tested the traditional concept of record in the digital environment, determining the characteristics of digital records as well as the necessary and sufficient attributes of a record that must be captured and preserved to ensure that its nature remains intact. This session will discuss the findings of InterPARES with respect to what is a record in the digital environment.

10:30-10:45 - Break

10:45-12:00 - Corinne Rogers, *Authenticity and Authentication*

According to archival science, record authenticity is determined on the basis of provenance and documentary context, and is linked to the immutability of documentary evidence affixed to a medium. However, in the digital environment, content, structure and form are no longer inextricably linked, and a presumption of authenticity is made by investigation of the record's elements of identity and integrity. This session will present the findings of InterPARES on authenticity and authentication in the digital environment, and the ontology of trustworthiness to which authenticity belongs.

12:00-13:30 - Lunch

13:30-16:30 - Hrvoje Stancic, *Technological Authentication*

(includes 15 minute break at 14:45)

The digital era brought new challenges to the archives. The underlying archival theory and concepts are the same for analogue and digital records. However, the technical manifestation of digital records, their variety, speed of creation, volume, and volatility require new, computational approaches. The digital era brought new challenges in authentication of records and the concept of (one) original. The digital signature further complicated long-term preservation of digital records, because their certification expires much sooner than the records' retention period. The blockchain and distributed ledger technologies (DLT) can support records' integrity, confirm their sequence, enhance non-repudiation, but also help preserve digitally signed records. The TrustChain model, resulting from the InterPARES Trust project research, establishes trust among voluntarily interconnected institutions (i.e. network nodes) and enables confirmation of digital signatures' validity and records' integrity even after the expiration of the signing certificates. The use case of the system for authenticating analogue university diplomas by connection to the blockchain will be shown.

Day 3 - Sunday 9 July

9:00-12:00 - Basma Makhoulf Shabou, *Information governance maturity assessment: concepts, tools and recent developments*

(includes 15 minute break at 10:30)

12:00-13:30 - Lunch

13:30-16:30 - Tracey P. Lauriault, *Data as artifacts and as records*

(includes 15 minute break at 14:45)

In this session participants will learn about the relationship between metrology and data, to think of data beyond normalized technological understandings, about different kinds of data and their ontological characteristics, including data in administrations, as national and historical artifacts and about data infrastructures. The session will frame concepts within contemporary systems such as smart cities, digital twins, and how data are often the root cause of many AI ethical and social justice issues. The session will end with a discussion of whether data are cultural and social artifacts, simply the outputs of scientific or administrative endeavours, and whether they are records.

Day 4 - Monday 10 July

9:00 – 12:00 - Erik Borglund, *Records in the Cloud*

(includes 15 minute break at 10:30)

Modern cloud service as we see it today was established by Amazon in 2002 where they used it as an internal service. In 2006, they offered the service outside of the company. The cloud as technology and the cloud where you find service providers have been a challenge for archival practitioners, records managers, and others. In this lecture the focus is upon the findings from the InterPARES “records in the cloud” project, but also a focus will be upon cloud technology challenges that you find today in 2023.

12:00-13:30 - Lunch

13:30 - 16:30 - Ken Thibodeau, *Proteus Bound: Trustworthy Digital Preservation*

(includes 15 minute break at 14:45)

Trustworthy digital preservation looks like an oxymoron, a contradiction. How can you keep something intact when it depends entirely on a network of loosely connected and very diverse things that change at different rates and in unpredictable ways? This is a question that InterPARES has addressed from the beginning. The first two InterPARES projects formulated and elaborated the concept of the chain of preservation, which is not a way to preserve digital records, but a way of documenting how they have been preserved. It provides a basis for judging whether their preservation is trustworthy. InterPARES provides a comprehensive model for integrating what is required in digital preservation with technical approaches for meeting such requirements. This session will review InterPARES research on digital preservation, its products and also looked more deeply into the problem, examining digital records in interactive, dynamic and experiential systems; that is, electronic records that have no precedents in hard copy. This research led to a deeper understanding of what it means to preserve digital records. These different threads came together in the articulation of functional and data requirements for digital preservation under the rubric of Preservation as a Service for Trust (PaaST). PaaST its impacts.

Day 5 - Tuesday 11 July

9:00-10:30 - Joe Tennis, *Metadata*

The findings of InterPARES have established Benchmark and Baseline requirements for preserving authentic records in digital systems. These, along with the Chain of Preservation Model, have been used to create the InterPARES Authenticity Metadata (IPAM). This session introduces these findings from InterPARES and discusses their relevance in the emerging technological environment.

10:30-10:45 - Break

10:45-12:00 - Pierluigi Feliciati, *Trusted and easy access to records and archives*

This session will present and discuss the main concepts and issues related to Archival Reference and Access resulting from InterPARES. It will focus on access as presented in the main archival conceptual models and standards, the role of records managers and archivists in mediating between authentic records and users in digital environments, and on how the related activities could be successfully managed. One of the topics will be a user-centred approach in the phases of conception and development, considering the organization of user studies to conceive, build and maintain good archival digital services. The presentation will include the first results of the InterPARES Trust AI study on “Users approaches and behaviors in accessing records and archives in the perspective of AI: a global user study”. Part of the educational activity will be discussing and evaluating some archival access services to better focus the notion of “quality” and the primary metrics for its evaluation.

12:00-13:30 - Lunch

13:30-16:30 - Jessica Bushey, *Managing & Preserving Digital Images Collections*

(includes 15 minute break at 14:45)

In this session participants will learn about managing and preserving aggregations of digital images based on InterPARES research into digital image creation and recordkeeping practices. The role of metadata to capture information contributing to the authenticity and reliability of an image will be explored, along with standards and best practices for metadata and image formats for access and preservation. Images held in social media collections and the challenges these platforms present to access and preservation will be also discussed. The session will end with an exploratory discussion about the opportunities and potential obstacles in using Artificial Intelligence to manage and preserve digital image collections.