

VSA-AAS

100

Jahre
ans
anni



Verein Schweizerischer Archivarinnen und Archivare

InterPARES
Trust AI



Archivists & AI: What are the expectations?

An introduction to InterPARES Trust AI

Corinne Rogers, PhD
Project Coordinator, InterPARES Trust AI
September 16, 2022



Artificial Intelligence Systems

Artificial Intelligence Systems (AIS) are computing systems using algorithms capable of carrying out complex tasks that were once believed to be the sole domain of natural intelligence:

- processing large quantities of information
- calculating and predicting outcomes
- learning and adapting responses to changing situations
- recognizing and classifying objects

AI challenges for archivists



Products of AI may provide:

- Inconclusive Evidence (based on probabilities)
- Inscrutable Evidence (no interpretability or transparency)
- Misguided Evidence (as good as the data provided)
- Unfair Outcomes (disproportionate impact on one group of people)
- Transformative Effects (challenges for autonomy and privacy)
- Non Traceability (hard to assign responsibility)



A question to consider

Can we use or develop AI systems to conduct records management and archival processes that will maintain the nature and trustworthiness of the records?



What do Archivists know about AI?

Using Artificial Intelligence to support archival work is not a new idea:

- 1969: J. B. Rhoads advocated for the “cybernetic archivist” in the (then) National Archives and Records Service, US
- 1987: P. B. Hirtle discussed recent developments in AI, expert systems, and automation in archives
- 1991: F. J. Stielow included AI in consideration of a paradigm shift in archival theory and practice
- Interest in AI for archival purposes continues, but research tends to focus on testing off-the-shelf tools for archival processes



AI challenges for archivists

- Relying on existing off the shelf tools limits what challenges can be met, as it makes the needs of archives subservient to the larger field of machine learning
- Many tangible instances of bias have been found in modern machine learning models, often driven by *laissez faire* data collection practices
- This raises the questions of a) whether off the shelf tools are the best solution for the archival field and b) what AI could look like if this power relationship between AI and archives were reversed, with archival theory informing the creation of AI tools

The InterPARES research network 1998-2026 and ITrust AI



Social Sciences and Humanities
Research Council of Canada

Conseil de recherches en
sciences humaines du Canada





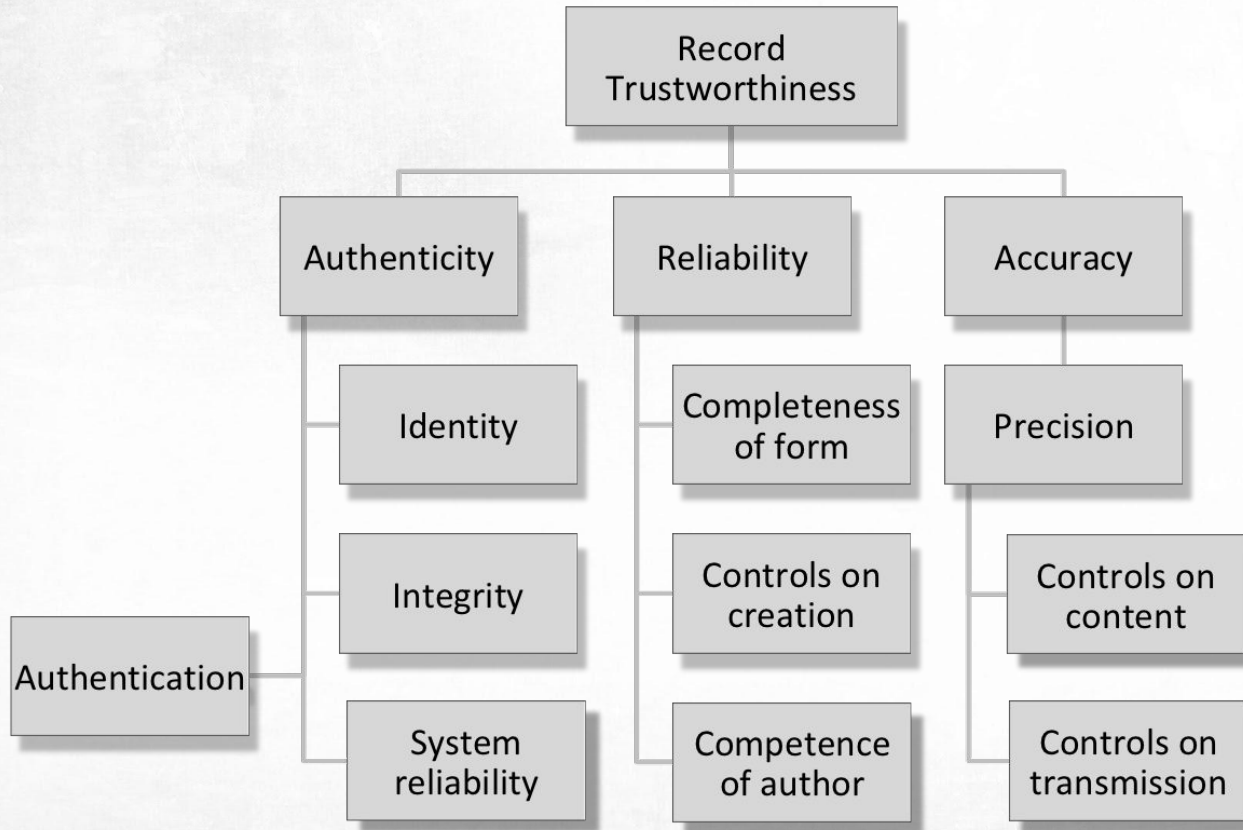
InterPARES Trust AI

- ITrust AI is the 5th phase of the InterPARES project, directed by Luciana Duranti and Muhammad Abdul-Mageed, and funded by the Social Sciences and Humanities Research Council of Canada (SSHRC).
- As in the previous phases, I Trust AI focuses on
 - maintaining the trustworthiness of digital records over time
 - trustworthy access to and preservation of digital records in all media and form
- I Trust AI focuses on Artificial Intelligence Systems applied to archival functions

Trust AI relies on fundamental archival concepts

- A **record** (archival document) is any document (i.e. information affixed to a medium, with stable content and fixed form) made or received in the course of activity, and kept for further action or reference (i.e. created)
- Because of the circumstances of its creation a record is **natural**, (a by-product of activity), **interrelated** (linked by an archival bond to all the records participating to the same activity), **impartial** (not created to answer questions researchers may ask of it in the future), and **authentic** (with respect to the creator, if used as an instrument of activity)
- Any tool we use on the records must ensure that these characteristics are protected, that is that the records remain trustworthy

Underpinning ITrust AI is the ontology of record trustworthiness





ITrust AI Project Goal

- Design, develop, and leverage Artificial Intelligence to support the ongoing availability and accessibility of trustworthy public records by forming a sustainable, ongoing partnership producing original research, training students and other highly qualified personnel, and generating a feedback loop between academia, archival institutions, government records professionals, and industry, to reinforce the knowledge and capabilities of each party



Objectives

- Determine the benefits and risks of using AI technologies on records and archives
- Identify and develop specific AI technologies that can address critical records and archives challenges
- Ensure that archival concepts and principles inform the development of responsible AI
- Validate outcomes through case studies and demonstrations



Studies

- Studies are international and interdisciplinary
- Focus on all aspects of archival functions
 - Creation and use of trustworthy records
 - Appraisal and acquisition of archival material
 - Arrangement and description
 - Retention and preservation
 - Management and administration of records and archives
 - Reference and access



Studies

- Studies address common concerns across domains (cultural heritage, government, education, private sector)
 - Privacy, personal and sensitive information
 - Ethical issues about records and about use of AI
 - Education, teaching, and AI literacy
 - Cultural heritage: digitization, access and retrieval, analysis and data mining
 - Records management and metadata
 - Municipal, regional, and national digital strategies



Examples

- Teaching curricula and AI literacy
 - Creating lesson plans and educational materials for records professionals to leverage and/or design AI to support access to trustworthy public records (Canada)
 - Identify competencies for critical AI evaluation (Brazil)
- Cultural heritage
 - Testing tools for image data analysis in medieval charters (Italy)
 - Modeling key archival functions that may benefit from the use of AI in digitization projects (EU-UNESCO)
 - Improve online access and retrieval of archival material using gamification as a component of machine learning (Turkey)



Examples

- Records management, privacy, ethical issues
 - Creating a functional model for records classification using AI tools (Indonesia)
 - Proof of concept using the Legal Entity Identifier standard to create enterprise master data sets, and best practice guidelines in the application of metadata for accessibility and preservation for financial institutions (Canada)
 - Exploring the meaning of “trustworthy AI” and the role of records management in national and international AI strategies (China, EU)
 - Testing/developing tools to identify and extract personal and sensitive information from records to support accessibility (Mexico, Chile)
 - Developing guidelines for ethical application of AI to records, and conducting ethical reviews (Canada)



Examples

- National and international digital strategies
 - Exploring the use of AI to manage records and data from critical systems surrounding incidents such as natural disasters or malicious actions (USA, Canada)
 - Understanding the records and their privacy and preservation challenges arising from smart grid communication systems and the role of AI (Canada)
 - Creating digital twins capable of trustworthy preservation (Canada)
 - Preserving AI tools and techniques as paradata - contextual materials/information in support of preserving the records they are applied to (USA)



Join us!

- InterPARES Trust AI is a partnership of organizations (archives and cultural heritage institutions, universities, technology companies, government departments...)
- We welcome new partners at any time
- Contact corinne.rogers@ubc.ca for more information
- <https://interparestrustai.org>
- @itrustai
- <https://www.facebook.com/interparestrust>
- <http://interpares.org> & <http://interparestrust.org>

THANK YOU!