



# Archives 4.0.

## Artificial Intelligence for Trust in Records and Archives

Luciana Duranti, Project Director and Professor  
School of Information  
The University of British Columbia  
Costa Rica 19 February 2020

# 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**,  
**learning** and **adapting** responses to changing situations,  
**recognizing** and **classifying** objects.

Can we develop systems carrying out competently and efficiently all records and archives functions all the while maintaining the trustworthiness of the records?



# AIS Identified Issues

They provide

- **Inconclusive** Evidence (based on probabilities)
- **Inscrutable** Evidence (no interpretability and transparency)
- **Misguided** Evidence (as good as the data provided)
- **Unfair** Outcomes (disproportionate impact on 1 group of people)
- **Transformative** Effects (challenges for autonomy and privacy)
- **Non Traceability** (hard to assign responsibility)

*Mittelstadt and colleagues (2016).*

Plus

- The decisions AIS make are based on past decisions, and
- when it comes to human affairs, tomorrow rarely resembles today, and data and numbers can't say what has a moral value, nor what is socially desirable



# The Montreal Declaration on Responsible Development of AIS (2018)

- It offers an **ethical framework** that allows the promotion of internationally recognized human rights in the fields affected by the rollout of artificial intelligence.
- Taken as a whole, the articulated principles lay the foundation of social trust towards artificially intelligent systems.



# Montreal Principles

- **Well-being** principle
- Respect for **Autonomy** Principle
- Protection of **Privacy** Principle
- **Solidarity** Principle
- Democratic **Participation** Principle
- **Equity** Principle
- **Diversity** and **Inclusion** Principle
- **Caution** Principle
- **Responsibility** Principle
- **Sustainable** Development Principle



# The Archival Challenge

- Overwhelming abundance, complexity, and lack of control of public records
- We have struggled to develop reproducible strategies and methods to both screen digital records for sensitive, confidential or legally restricted information, and provide effective access to them
- Metadata – the data about data that allows us to prove the authenticity (i.e. identity and integrity) of a record – is beyond our ability to manage manually
- Current classification methods are unable to achieve the levels of precision, accuracy, and recall that match or exceed the levels generated by human classifiers



# Archives 4:00. The Goal

The **goal** of this project is to

- generate original research,
- train highly qualified personnel (HQP), and
- establish a virtuous circle (i.e. a feedback loop that reinforces the knowledge and capabilities of each party) among academic institutions, records offices, archival institutions, and industry to
- leverage AI technologies and tools to support the ongoing availability and accessibility of trustworthy public records.



# Archives 4.0. The Objectives

- **Identify specific AI technologies** that can address critical archival challenges (i.e, the identification and organization, selection, description, and provision of access to records and archives while ensuring human rights, privacy, transparency, and accountability; protecting trustworthy evidence; and preserving cultural heritage);
- **Determine the risks and benefits** of using AI technologies on records and archives;
- Establish how **archival concepts and principles can inform the development of responsible AI**; and
- **Validate outcomes** from objective 3 through case studies and demonstrations.



# Archives 4.0. The Approaches

- Our short-term approach focuses on **identifying high impact problems and limitations in records and archival functions, and applying AI to improve the situation**. Our interdisciplinary team of archival scientists and AI experts will do so in consultation with archival practitioners, who will provide their perspective on the problem domain.
- Our long-term approach focuses on **identifying and developing tools with which the records and archives community can address requirements for trustworthy records** in an ever-changing socio-technical environment. This includes decision-making support and, following that, rapid application of AI-based solutions to specific needs that arise over time.



# Archives 4.0. The Focus

## Identifying

- the **holistic impact of AI solutions** in the long-term context of public records and archives, on a global basis
- **how to support decision-making** at the global, national, and local level when confronted with a larger challenge linked to the use of AI: results leading to incorrect, inequitable, or dangerous consequence
- the broader **ethical, legal, and social implications** of any solution implemented to create, manage, use, preserve, and provide access to them.

InterPARES  
Trust



# Interdisciplinarity

- Cooperation between scholars and professionals in the fields of archival science, artificial intelligence, law, cybersecurity, digital forensics and ethics
- The theoretical framework will be based on **archival and diplomatic theory**
- This theoretical foundation will be complemented and enhanced by the theory of **Computational Archival Science**, a “transdisciplinary field grounded in archival, information, and computation science that is concerned with the application of computational methods and resources, design patterns, socio-technical constructs, and human-technology interaction to large-scale (big data) records/archives processing, analysis, storage, long-term preservation, and access problems” (Payne)
- As one of the primary objectives of this project is to understand the risks and benefits of AI adoption to carry out archival functions, we will also apply the theory of **risk management**, an area of study that informs decision making in the context of an uncertain environment



# Methodology

- **Soft systems thinking methodology**, characterised as an attempt to find common principles that apply at different levels of scale and across different types of phenomena.
- The project will proceed through **five phases**, which will utilize **qualitative methods** to meet the objectives. Each phase builds upon the knowledge gained in the previous phase.



# Methods: Phase 1

- Identify critical archival challenges that might be addressed by AI
  - **Surveys and interviews** with practitioners and experts within the archival community
- Identify within each critical challenge the specific factors to be addressed and how AI might address them
  - **Expert interviews and mapping**
- Identify candidate AI technologies
  - **Use cases**
- Create initial **evaluation criteria** for AI solutions for archival challenges



# Methods: Phase 2

- Determine the requirements of public records compared to the capabilities of AI technologies
  - Conduct **doctrinal legal research**
  - Develop a **value structure for risks and benefits**
- Identify the limitations of each potential AI solution
  - **Policy analysis**
  - **Expert interviews**
  - **Environmental Scans**
- Develop list of threats and vulnerabilities
  - **SWOT/PESTLE Analysis**
  - **Theoretical Analysis**
  - **Stakeholder Interviews**
  - **Expert Assessment**
- Augment the validation criteria



# Methods: Phase 3

- Establish archival principles to be used to inform AI
- Identify and confirm bias
  - **Consistency Analysis**
- Examine if AI informed by archival principles is more aligned with archival needs
  - **Quasi-experimental methods**



# Methods: Phase 4

- Measure AI solutions against the validation criteria developed in Phases 1 and 2
  - **Experimental and quasi-experimental methods**
- Examine feasibility, sustainability, bias, transparency, generalizability, and preservation of context in AI solutions
  - **Case studies**
  - **Use cases**
- Develop and validate tools, including framework for evaluation and checklists for institutions considering AI implementation
  - **Use Cases**

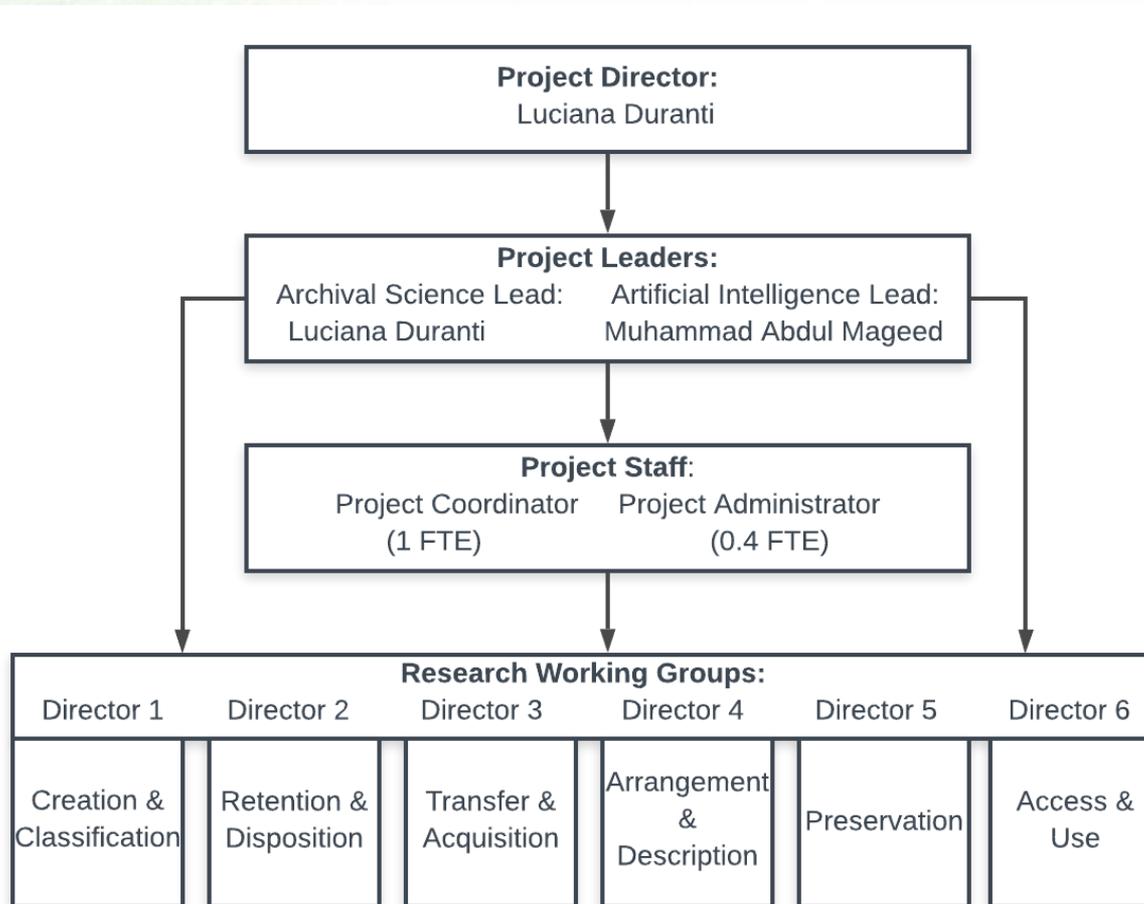


# Methods: Phase 5

- Research Dissemination and Output
  - Interdisciplinary knowledge flow
  - Interactive online content
  - Public in-person presentations
  - Curriculum development and teaching
- Publications



# Governance



# Partners to date

- Archives, City of Victoria
- Archives, City of Vancouver
- Artefactual, Canada
- Bar Ilan University, Israel
- Beijing Union University, China
- Carleton University, Canada
- Center for Information Assurance and Cybersecurity, U. of Washington, USA
- City Council, Girona, Spain
- City of Toronto, Canada
- Landesarchiv Baden-Wuerttemberg, Germany
- Library and Archives Canada
- National Archives of Brazil
- National Archives of Iran
- National Archives of Slovenia
- National Archives of the UK
- State Archives of Hawaii, USA
- Universiti Kebangsaan, Malaysia
- Israel State Archives, Israel
- McGill University, Canada
- Mid-Sweden University, Sweden
- North Vancouver Museum and Archives, Canada
- Regione Emilia-Romagna, Italy
- San Jose State University, USA
- UNESCO Archives
- ICCROM Archives
- University of Macerata, Italy
- University of Zagreb, Croatia
- Uppsala University, Sweden

InterPARES  
Trust



# Individual Researchers

Canada

USA

UK

Brazil

China

Croatia

Germany

Israel

Italy

Malaysia

Russia

Slovenia

South Africa

Sudan

Sweden

InterPARES  
Trust



# Timeline

If we get funded...

- Second stage of the application deadline: November 1<sup>st</sup>
- Results: April 2021
- Start date: June 2021
- End date: May 2026



**THANK YOU!**

**LUCIANA.DURANTI@UBC.CA**

**InterPARES  
Trust**

