Authenticity and Authentication
From Assessing to Preserving it!

Luciana Duranti
The University of British Columbia
Webinar, Arsip Nasional Republik Indonesia (ANRI)
14 July 2021

Definitions: Records/Preservation

- A **record** – or **archival document** – is any document made or received in the course of activity, and kept for further action or reference (i.e. created)
  - Being a document (i.e. information affixed to a medium), a record has stable content and fixed form
- Because of the circumstances of its creation a record is **natural**, (a by-product of activity), **interrelated** (linked by an archival bond) to other records, and **impartial** (not created to answer questions researchers may ask of it in the future)
- To **preserve** a record means to ensure its physical and/or technological stabilization (for the purpose of extending its life indefinitely) and the protection of its nature, intellectual content, and relationships

Definitions: Digital Records/Digital Preservation

- The records **content, structure, and form are not inextricably linked**
- The record as a stored entity is distinct from its manifestation on a computer screen, and its **digital components** have to be considered as well as its **documentary form**
- Digital records are **vulnerable** (easy to destroy, lose, corrupt, tamper with, or become inaccessible if not protected) yet **persistent** (forever there, if not purposefully destroyed)
- When we save a record, we take it apart in its digital components. When we retrieve it, we create a copy: there are **no originals** in the digital environment
- **Hence, it is not possible to preserve digital records: we can only preserve the ability to re-produce or re-create it**
- **Digital preservation** is the process of maintaining **authentic copies** of digital materials and keeping them accessible during and across different generations of technology over time, irrespective of where they are stored

InterPARES Trust AI

Authenticity

- The canon goes: **a record is authentic when it is what it purports to be**.
- The question one might ask is how do we know what a record purports to be? Where do we find in the record such a claim? In the endorsement, the registration, the classification code, the subject line, the disposition? Or rather, not in what the record says directly but in what the record is materially made of? Or, on what body has the custody of it?
- It depends on the culture of a place and time, the discipline in the context of which authenticity has to be established, the technological context, or the law.
- **Civil law**, in determining record authenticity, has traditionally focused more on the issuer of the record than on the record itself, presuming authentic any record made by a sovereign authority or in its name (i.e. any public record) or by its delegate, such as a notary or lawyer attesting the identity of a private record.
- But, what if the person issuing the record is not a sovereign authority?

InterPARES Trust AI

Diplomatic Authenticity

- **Diplomatics** has long been concerned with the authenticity of records and, since first developed in 1681, it has aimed to establish a scientific methodology for determining the authenticity of any record.
- This methodology examined the **form** of the record, that is, the rules of representation used to convey a message (those characteristics of a record that can be separated from the determination of the particular subjects, persons, or places that the record is concerned with) and the records **degree of perfection** (draft, copy, original).
- Form is physical, i.e. the external make-up of a records (e.g. medium, ink), and intellectual, i.e. its internal articulation (e.g. salutation, preamble). If both correspond to the practice of the presumed or declared time, place, and author, then the record is authentic.
- The analytical approach of diplomatics aims to **establish on the record itself that the record is what it appears to be**, or what whoever submits it as evidence of a fact or an act claims it to be.

InterPARES Trust AI

Archival Authenticity

- Archival science includes authenticity among the qualities that characterize every record, together with naturalness, impartiality and interrelatedness, and links it to them.
- **All records are authentic with respect to their creator**, that is, the natural or juridical person who makes or receives them, and keeps them for further action or reference, that is, for its own legitimate purposes, even when, diplomatically, they are forgeries.
- Archives are authentic when they are created (made or received and kept) for the need to act through them and when they are preserved as faithful witness of facts and acts by the creator and its legitimate successors.
- Archival science, by **linking the record to its context of creation and preservation**, extended authenticity from being a property of the record itself to being a property of procedures and further tied it to unbroken custody

InterPARES Trust AI

Authenticity in the Digital Environment

- There was no question in archival science that the identity of a record, and therefore its authenticity, resided in the provenance and documentary context of the record, but **this fact turned out to be linked to the immutability of record affixed to a permanent medium, that is to its integrity.**
- In the 1990s, the InterPARES project understood that, **in the digital environment, authenticity could no longer reside only in the records context.**
- In fact, even if the relationships between and among the records established at creation remained intact throughout their use, maintenance and preservation, **the documentary component of the entity record could lose integrity** (a quality of the record that was never before in the equation for establishing authenticity at the side of identity), because its content, structure and form are not inextricably linked.
- Thus, InterPARES returned to diplomatic authenticity and looked separately to identity and integrity.

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)**

Integrity

Integrity refers to the quality of <u>being complete and unaltered in all essential respects</u>.

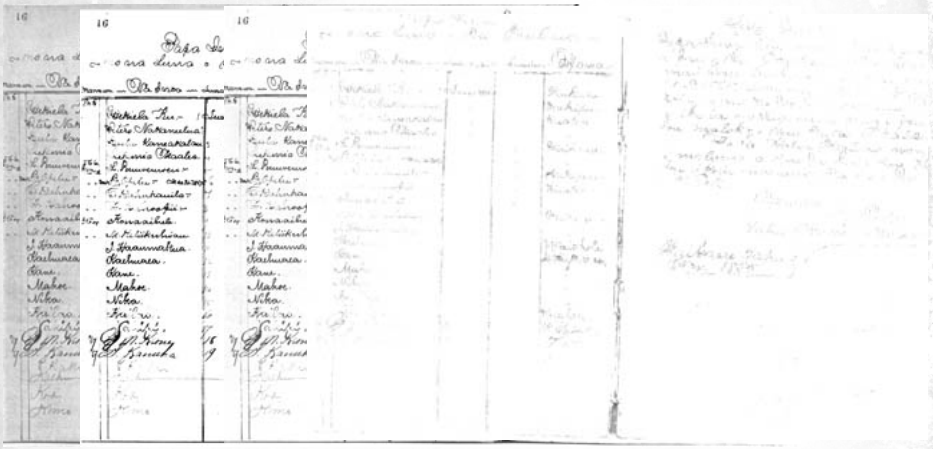
We were never fussy about it. What if a document had holes, was burned on a side or the ink passed through?

The same definition of integrity was used with respect to data, documents, records, copies, records systems

As long as it was good enough to understand it, ...but how good is good enough in the digital environment?

InterPARES TrustAI

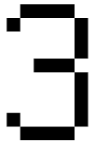
Loss of Integrity: Analog Document



InterPARES TrustAI

Loss of Integrity: Digital Document

- If Original Bits 101
- Change state to 110
- Continues to a 011



- Same bits, but Different value

Bitwise Integrity

Often identified with **Data Integrity**:

- The data in the document are not modified either intentionally or accidentally
- The original **bits are in a complete and unaltered state** from the time of capture, that is, they have the exact and same order and value

A small change in a bit means a very different value presented on the screen or action taken in a program or database.

Duplication Integrity

The process of creating a copy does not modify a record (either intentionally or accidentally) as the output is an exact bit copy of the original data set (form, content and composition data).

Duplication integrity is **linked to time** and one should consider the use of time stamps for that purpose.

But, in the digital environment, when we say duplication, we need to be explicit about what we mean, copy or image?

Duplication: Copy

Selective duplicate (e.g. PDF)

- You only copy what you see
- Rarely includes confirmation of completeness
- Provides incomplete picture of the digital environment

Duplication: Image

Forensic duplicate:

A bit by bit reproduction of the storage medium and its content, including ambient data (e.g. snapshots of each open file), swap space (virtual memory, with passwords and encryption keys) and slack space (with deleted material)

Duplication Process Integrity: Principles

Principle of Non-interference: the method used to re-produce or re-create a digital document does not change the digital entities

Principle of Identifiable interference: if the method used does alter the entities, the changes are identifiable and identified (including para-data)

Authentication

Definition: A declaration of authenticity based on either direct knowledge, material proof, inference, or deduction

Basis for authentication of digital records

- **A chain of legitimate custody** remains ground (an increasingly significant ground!) for inferring authenticity and authenticate a record (unbroken chain of custody).
- **Digital chain of information:** the information preserved about the record and its changes that shows specific data was in a particular state at a given date and time.
- A **declaration** made by an expert who bases it on the trustworthiness of the system hosting the record and procedures and processes controlling its preservation and use

Trustworthiness of the System

- Can we consider **system integrity** another component of authenticity? In addition to a record identity and integrity, do we now have the integrity of the system(s) in which the record resides at any given time in its lifecycle as a component of a record authenticity?
- **No**. The integrity of the system—as mentioned—supports authentication, which is a declaration of a record's authenticity.
- The InterPARES project determined that digital records cannot be preserved. *It is only possible to preserve our ability to re-produce or re-create them.*
- This is based on two considerations.
  - For records that are the counterpart of traditional paper records, the physical form in which they are stored is necessarily different from what is displayed on a screen or printed on paper, as well as from the physical form of the record in a computer processor.
  - For digital records of which there are no analogue precedents, such as records in dynamic, interactive and experiential systems, it is not possible to preserve, in digital form, a record intended for human use.

Stored Record

- InterPARES introduced the concept of a **stored record**.
- The record we keep is the stored digital encoding of a record.
- It is distinguished from the **manifested record**, which is either a copy of the record in a form suitable for human use or in a form suitable for use in an automated system designed to process such records.
- The general requirement for digital records, then, is that regardless of how a record is represented in bits in digital storage, it must be possible to generate a manifested record that has all the identifying attributes of the first effective version of that record (the original).
- This shifts the focus from the physical preservation of an original — the first complete records capable of reaching its effects, an amorphous concept in the digital realm — to the **making of authentic reproductions**.
- The practical issue extends all the way to feasibility. Thus, technology for a long time now has focused on technological authentication.

Technology Dependent Authentication

Digital signature:

- protects **bitwise integrity**
- verifies a record's origin (part of its **identity**), makes record indisputable and incontestable (**non-repudiation**)
- has been given legal value by legislative acts (e.g., European Directive on electronic signatures) or regulatory bodies (Security Exchange Commission on hash functions)
- is enabled through complex and costly public-key infrastructures (PKI)
- ensures authenticity of information **across space,** not **time**!
- is subject to **obsolescence**, and compounds the problem of preservation as it cannot be migrated with the record to which it is attached, and the certificates have an expiration date
- Theory tells us that it has the function of a **seal**, rather than of a signature, so it can be removed and substituted with metadata

Technology Dependent Authentication

Blockchain technology

- the underlying technology enabling Bitcoin
- a ledger, i.e. an information store which keeps a final and definitive (immutable) trace of transactions (their hash).

- relies upon a **distributed network** (all nodes—servers are equal) and **decentralized consensus** (no centre(s); no single point of control or attack)
- The confirmed and validated sets of transactions are held in blocks, which are linked (chained) in a chain that is tamper-resistant and append-only
- It starts with a genesis block and each block contains, in addition to the hash of a predetermined number of documents, a hash of the prior block in the chain.

How is Blockchain used?

Blockchain can be used to confirm

- the **integrity** of a record kept elsewhere
- that a record **existed** or **was created** at a certain point in time (i.e. not after being timestamped and registered in the blockchain)
- the **sequence** of uploading of records to the blockchain

Is it a **recordkeeping system**? No. It holds the hash of records, not records (smart contracts—agreements between parties directly written into lines of code—are not records). The records must still be stored and managed off chain. This is good, because, if they were in the blockchain, they would be **immutable.**

Immutability/Integrity

- It is the attraction of blockchain: it is what ensures integrity as nothing can be changed in a record or removed from a block
- It is the key problem of blockchain:
  - with **current records**, any **updating or correction** of wrong data; any form of **privacy protection**; any exercise of the **right to be forgotten**; any **disposition** of no longer needed records; any **record making system upgrade**; in short, any change in the record/s would invalidate the blockchain
  - with **records identified for continuing preservation**, any **transfer** to a preservation system; any **migration**; any **addition** to the records aggregation would invalidate the blockchain

InterPARES Trust AI

Identity

- The hash on the blockchain does not allow for links to
  - the hash of related records, hence **no archival bond**, the interrelationship among the records
  - the hash of metadata, hence **no context**
- If the metadata were embedded in the record at creation, the hash of such record would not allow for additions or changes

InterPARES Trust AI

Authenticity Problems with Blockchain

- Proving **authenticity at origin** is not possible
- Preserving the **contextual evidence** (the naturalness and interrelatedness of the records resulting from a process)
- Handling the **decentralized** (and thus trans-jurisdictional) nature of the blockchain (who is the creator? the author? the owner? What law applies?)
- Dealing with code in a situation where the necessary components of the transaction are controlled by different actors in different jurisdictions; and, with **smart contracts**, lacking both the equivalent of a signature and the date of the completion of an agreement.

A Limited Use?

- The Archangel Project team (TNA, U. of Surrey, and Tim Berners-Lee's Open Data Institute) states:

"Blockchain offers a shield which archives can use to defend the records as authentic. By enabling researchers to compare the content of evidence (including the checksum of the record) to that recorded on the blockchain, they can see proof that no changes (deliberate or accidental) have been made to the record since it was preserved in the archive. Further to this, the decentralized nature of blockchains removes the need for citizens to trust individual institutions as each is the guardian for the other guardians."

- Except that preservation activities, such as migration and conversion, change the bits in the records.

Decentralization Problems

- Information processing happens on a complex technological stack in which **different technical components may be in the custody of, and operated by, very different actors**.
- Some components may be under the control of a single organization, others under the control of business partners who are members of a blockchain consortium, and still others under control of unknown third-party actors.
- **An organization's records could be in the custody of thousands of independent actors over which records creators exercise little or no control.**

Decentralization (cont.)

- The **consensus mechanism** and other protocols or standards determining how the blockchain operates, may not be within the decision-making purview of the records creator (or the creator's designated records professional)
- Instead, these may be decided by remote (and even unknown) third party developers. In many cases, these protocols and standards are still unstable, and thus the reliability of the upload of organizational records to the blockchain could be very difficult to establish with any certainty

What About Partial Decentralization?

InterPARES TRUSTER Preservation Model

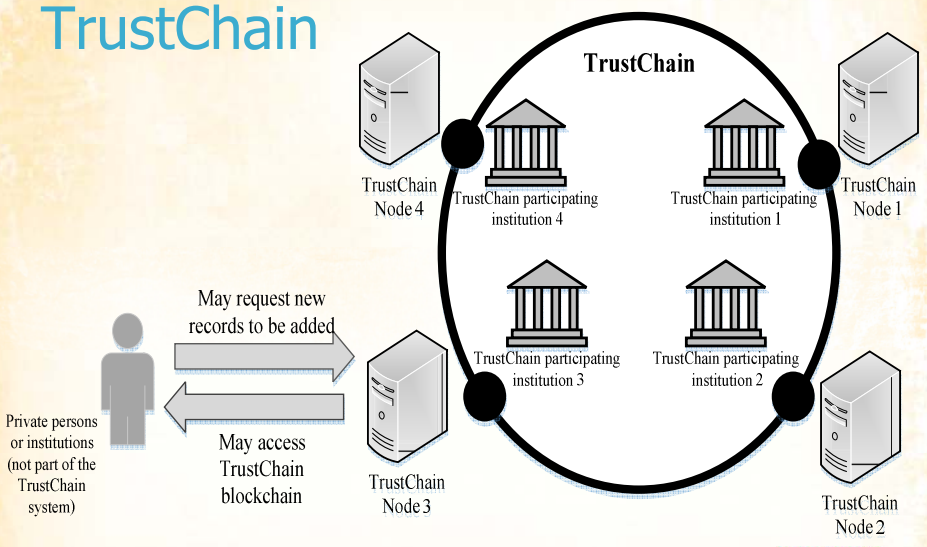
- Blockchain-based system called "**TrustChain**"
- Applies the concepts of
  - hash algorythms
  - blockchain
  - distributed consensus
- Presumptions:
  - private cloud blockchain
  - only approved nodes can write
  - everyone can read

TrustChain

The proposed **TrustChain** system

- relies on the involvement of a **group of trusted institutions**
- **the recordkeeping system in the creating office** and **the preservation system in the archives** would work **in concert** along the lifecycle of the records
- would provide confirmation of **integrity**, time of **creation/existence**, **sequence of records**, **non-repudiation**, **validity of e-signatures** whose certificate has not expired

TrustChain



InterPARES TrustAI

System Authentication

Yet, even the TrustChain would be a temporary and expensive expedient as a new chain would have to be generated at every step in the record lifecycle

Let's forget technological authentication and go back to trusted systems. **Electronic Documents and Records Management Systems (EDRMS)** share some fundamental requirements that support authenticity and authentication:

- The system software should be able to present old materials as they originally appeared (**backward compatible**), and allow the sharing of materials easily with other systems (**interoperable**)
- The **software** should have undergone **theoretical or empirical testing and peer review**; its **error rate should be known**; and it should have gained **general acceptance within the scientific community** (Daubert standard)
- The **formats** used should be **non- proprietary**, **platform independent and uncompressed**, with **freely available specifications (open format)** and software whose **source code is made (freely) available and can be modified (open source).**

InterPARES TrustAI

System Authentication (cont.)

- The **results produced by using the system should be repeatable, objective and verifiable.**
- The **specifications of the software must be maintained and availa**ble.
- If the software is customized, the **changes must be documented** (including comments in the software code).
- The **construction of the whole system must be documented**.

The integrity of any system (not only EDRMS) should be inferred from sufficient **security measures** to prevent unauthorized or untracked access to the computers, networks, devices or storage; and **stable physical devices** that will ensure the values they were provided with should be maintained until changed with authorization.

These devices include user names and permissions, passwords, firewalls and logs.

System Authentication (cont.)

Logs are an important part of the authentication of the system and the records. They are sets of files automatically created to track the actions taken, services run, or files accessed or modified, and the time, identity of the person undertaking the action and their location. They can be separated into:

- **Web logs** (Client IP Address, Request Date/ Time, Page Requested, HTTP Code, Bytes Sent, Browser Type, etc.).
- **Access logs** (User account ID, User IP address, File Descriptor, Actions taken upon record, Unbind record, Closed connection).
- **Transaction logs** (History of actions taken on a system to ensure Atomicity, Consistency, Isolation, Durability (ACID); Sequence number; Link to previous log; Transaction ID; Type; Updates, commits, aborts, completes).
- **Auditing Logs**. They demonstrate the integrity of the system; provide checks and balances, determine effective security policies, catch errors that occur, provide instantaneous notification of events, monitor many systems and devices through 'dashboards', allow to determine accountability of people, provide the snapshot for post- event reconstruction ('black- box'), and, if retained for a long enough time, have the capability to answer the Who- What- Where- When questions.

Authentication in an Archival System

- When the records are in the custody of an archival institution, **archival description** (that is, inventories) acquires a **primary authentication function**.
- The authentication function of archival description is a collective attestation of the authenticity of the documents or records in an archival fonds (the Canadian expression 'archival fonds' is equivalent to 'archive' in British usage, and 'archives' in Australian and American usage) as well as of all their interrelationships; in other words, **authenticity in their documentary context**.
- **Archival description provides a historical view of the records and of their transformations while maintaining the bond of their common provenance and destination**.
- Archival description of permanent digital records relies on metadata as evidence about a record's identity and integrity.
- The authenticity of the records in an organization's archives or an archival institution can be presumed if such organization or the archival institution has a trusted digital repository.

InterPARES Trust AI

Trusted Digital Repository

- Based on the **OAIS** model, which was not developed by archival specialists, neither was it intended for archival institutions. Rather, it was conceived as a preservation system internal to an organization, such as NASA.
- It offers a **conceptual framework for digital preservation** that describes, in a technologically neutral manner, the activities and the information that are necessary for trustworthy preservation.
- Effectively, it has defined the universe of discourse for digital preservation in a variety of contexts around the world. It details the authorized custody services of a **Trusted Third Party Repository (TTPR)** in order to ensure provable authenticity of the clients' digital records and serve as a source of reliable evidence. ISO 16363 2012
- It describes the services and processes to be provided by a TTPR for the clients' digital records during the retention period to ensure trust. It also details the criteria of 'trustworthiness' and the particular requirements of TTPR services, hardware and software systems, and management.

InterPARES Trust AI

Chain of Preservation

- The InterPARES project recognized that digital preservation requires a **Chain of Preservation** (COP) that ensures that digital records survive uncorrupted from creation through their migration from one system to another.
- The phrase 'Chain of Preservation' was chosen to indicate that all the activities to manage records throughout their existence are linked, as in a chain, and are interdependent. If a link in the chain fails, the chain cannot do its job. Any break in how digital information has been preserved could make it impossible to assert that what remains is what it should be.
- The COP is realized by implementing controls that ensure that the requirements for preservation are satisfied throughout the life of the records. **The COP is reflected, after the fact, in data that demonstrate that these requirements have been satisfied, the identity and integrity metadata discussed earlier**.

Preserved Records Authentication

- By scrutinizing the digital records preservation practice in the context of the authenticity metadata, it is possible to say that **digital records authentication can be broken down into at least two tiers**.
- The first, and most important, tier for the presumption of authenticity is to **audit the integrity of the preservation system in which records are kept.**
- To date there are a number of ways integrity metadata are included in this tier, such as the use of checksums, plus the conduct of visual inspection and the comparison with duplicated material in a parallel system.
- The second tier for the presumption of authenticity moves from audit of integrity to the **verification of identity**. Identity metadata are provided during the process of making, transmitting, receiving and storing the digital records. They rely on how the creator's system works to encapsulate the entirety of the state of the digital evidence as used by the creator in the ordinary course of business.
- Thus, **to authenticate material that has been preserved in a system other than the one in which it was generated and/ or received, it is necessary to authenticate all digital systems.**

Migration and Format Changes

Some activities are routinely carried out by archives to ensure the continuing integrity of the records, as well as the ability to verify it. Appropriately qualified witnesses should be able to attest that the organization has implemented the following:

- A **controlled process of migration** to the archives' technological environment (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;
- **Implementation of privileges concerning the access, use and reproduction of the records within the archives**; and

Migration and Format Change (cont.)

Procedures to

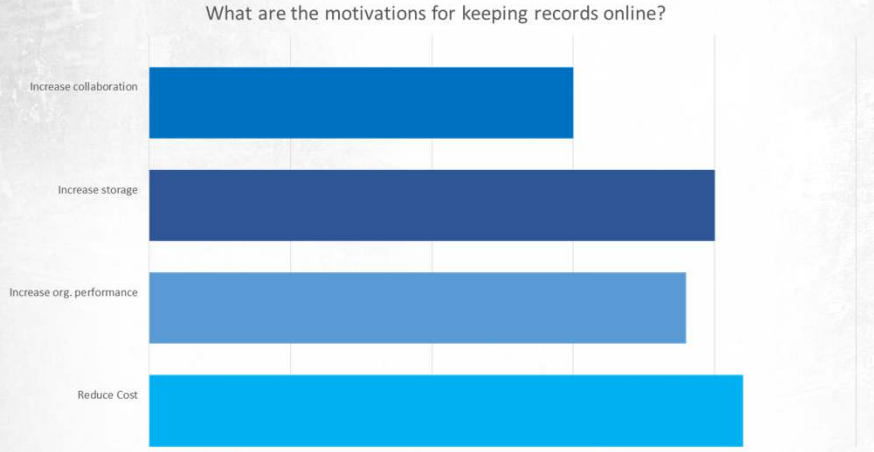
- **prevent, discover and correct loss or corruption of records;**
- guarantee the continuing identity and integrity of the records against media deterioration and across technological changes though **continuing conversion and migration**;
- assign **responsibility for and means of authentication of individual records**, when required; and
- ensure **redundancy, internally and remotely** (LOCKSS).

Consideration should be given to

- verifying that the organization has selected **preservation formats** using accepted criteria, such as widespread adoption, non-proprietary origin, published specifications, interoperability (platform independence), and lack of compression or lossless compression.

Would it not be much easier to entrust the records to a Cloud Provider?

Records in the Cloud



InterPARES Trust AI

Issues

- Data ownership
- Availability, access, and reliability
- Retention and disposition
- Storage and maintenance
- Security
- Location and transfer
- End of service
- Preservation
- Trustworthiness

InterPARES Trust AI

Data Ownership

- When a user entrusts its records to a provider and uses the latter's platform and application to generate additional data, the **provider will create data** related to actions about data processing, management, etc.
- While the content created and/or stored in the cloud by the user is owned by such user, **the metadata created by the provider are not**, and, as the user needs them to demonstrate the integrity of the records, contractual agreements should determine whether and how the user has the **right to access and use the provider's metadata**.

Availability, Access, Reliability

- **Availability** is a fact, while **access** is a right, but the latter cannot be satisfied without the former
- In a cloud environment, **availability of the stored records** implies also the **availability of the infrastructure** (i.e. the amount of time that a system is expected to be in service is 100%), which facilitates the retrieval and readability of the data, because technical difficulties might slow a FOIA process and the owner of the data, being liable for providing access to them, may be sanctioned
- **Reliability** is the characteristic of behaving consistently with expectations: one must consider not only availability of the records through redundancy but also **consistency** and **accuracy** of access.

Records Retention and Disposition

- **Compliance** is difficult to verify.
  - **transfer** from a system to another for retention might involve loss of authenticity
  - **destruction** might involve
    - a breach of confidentiality or privacy,
    - persistence of some of the copies and related metadata, and
    - persistence of the metadata generated by the provider about the user's data.

Records Storage and Maintenance

- Storage and maintenance impact the quality of the records and they ability to serve as legal evidence, especially in legal jurisdictions where the **authenticity** of the record is an inference made from the **integrity** of the system where the data reside (Canadian Government Standard Board 74:32 2017).
- Contractual agreements do not generally specify how records are maintained **across changing technologies and data formats**, and they generally say that users are responsible for backing up their data. All maintenance procedures, including proper storage, care, custody, and data control, are referred to by providers as "backup procedures."

Records Security

- It is protection of the system/records from **unauthorised access, use, alteration or destruction.** In a world where integrity of a system is an inference from which one infers integrity of the record, from which one infers its authenticity and then trustworthiness, **security is the new authenticity**.
- Individuals enforce security with something they know (e.g. password), they own (e.g. tokens), or they are (e.g., biometrics of eyes, fingerprints, private keys in a PKI environment)
- A cloud provider enforces it through encryption and should **produce audit trails and access logs** and capture, maintain and make available **metadata** associated with access, retrieval, use and management of the data, in addition to those linked to the data themselves.
- **The security issue links directly to the matter of data location and cross-border data flow.**

InterPARES Trust AI

Records Location and Transfer

- The cloud is the platform of choice for **mobile applications** and the data generated using them, as well as those created in **smart devices.** Records can be in data centres anywhere in the world
- The location of the records is a criterion in **determining the law that applies** in case of litigation
- National strategies used to require that records resides within the boundaries of the country where they were created (very expensive for data centres, if Europe or North America).
- The international strategy no longer requires that, thereby underscoring the importance of **multilateral agreements** among countries for collaboration in security (new safe harbour)

InterPARES Trust AI

End of Service – Contract Termination

- If the provider ceases to exist or terminates one or more of its services (for breach, inactivity, or convenience), the records will be **deleted** or **inaccessible**
- Free services do not have an established duration and may close accounts **unilaterally**, requiring users to delete software and applications, and preventing them from accessing the data left with the provider
- When the data are given back to the user it is not certain that they will be in a **usable** and **interoperable** format
- If the contract is terminated by the user, the restitution of the data may be **expensive** and the data may not be in accessible formats. Also, the user may not have **the right to access the metadata** generated by the system for its recordkeeping or legal purposes, and may have no guarantee that the provider will **destroy** every copy of the data held in the data centers

Records Preservation

- Preserving records in the cloud is a **black box process**
- Providers **may not know where the records are**, can and do **subcontract** some of their services to other providers, potentially maintaining servers or being registered as providers in different countries.
- One cannot expect that the same hardware and software will remain in service for as long as the records must be preserved, or that the technologies replacing them will be **compatible** with the previous ones.
- Standards give information about preservation formats but there is **no way of controlling compliance**
- There is **no way of ensuring and verifying authenticity**

Final Remarks

- **Authentication is about proving that something is what it purports to be.**
- When seeking to prove the authenticity of digital material, the traditional methods of authenticating paper and other forms of analogue evidence do not apply
- Although traditional means of authentication, such as **proof of chain of custody** and **trustworthiness of the preserver**, still have a role in assessing that an entity is what it claims to be, these criteria on their own are insufficient to demonstrate the authenticity of digital records.
- Even the term 'demonstrating' is at issue in the digital environment because, at most, authenticity may be 'inferred' from several factors rather than shown, due to the fact that digital material is perpetually being reproduced in the process of maintenance and use, and the entity under consideration is always new.
- The significant differences between authentication of analogue material and that of electronic evidence lie in two fundamental concepts: **system integrity and security**, and the **records significant properties**.

InterPARES TrustAI

Final Remarks: System Integrity and Security

- System integrity and security come into play in jurisdictions, such as Canada, where authentication is based on an inference made from the technological environment in which the records exist.
- Because of the vulnerability of digital material and the difficulty of establishing authenticity by examining the digital entity itself, its identity and integrity can be deduced from the system's requirements on access, use, management and such like.
- This implies that strict policies and procedures must exist for controlling all the records in the system, as well as any interaction with them within the system and from outside.
- However, this happens only in **institutions and organizations that have their own information and preservation systems subject to mandated standards of practice**.

InterPARES TrustAI

Final Remarks: Significant Properties

- When organizations and individuals entrust their records to **cloud providers**, it is not possible to verify the integrity of servers where the digital material is stored.
- Thus, it is only possible to examine the security measures agreed upon in the contract between provider and user and make an inference of authenticity from them.
- But, we can still **determine authenticity on the basis of the records significant properties**.
- The most significant properties of digital records are the **attributes necessary to establish their identity and integrity through time**.
- These metadata are produced when a digital entity is generated (they contribute to establishing its identity), during its use and management (they help to establish its integrity) and after the entity is selected for permanent preservation in an archives to ensure that its authenticity remains verifiable over time.
- Some of significant properties though are logs, which are not accessible if the records are entrusted to a cloud provider, who owns all the properties it adds.

InterPARES Trust AI

Conclusion

The **fundamental difference** between the authentication of analogue and electronic records is in the fact that, while analogue material can be authenticated on its face and only exceptionally is circumstantial or extrinsic evidence necessary, the **authentication of digital material**

- **is always an inference** based on extrinsic elements such as significant properties, and
- **relies on circumstantial evidence** such as
  - the integrity of the system hosting it,
  - the policies and procedures controlling it, and
  - the technology encrypting or securing the access to it.

However, it might be possible in the new future to use **Artificial Intelligence based on archival concepts to authenticate archival materials**. To find out stay tuned on the **InterPARES Trust AI** project!

InterPARES Trust AI

THANK YOU

luciana.duranti@ubc.ca

www.interpares.org

www.interparestrustai.org

www.ciscra.org