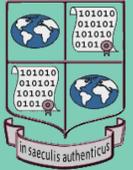


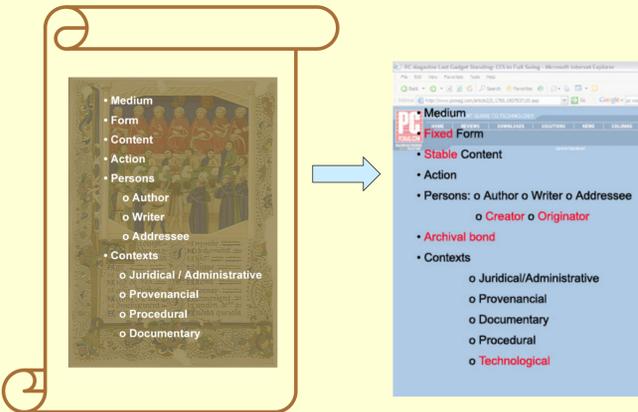


# Archival Diplomatics of Digital Records

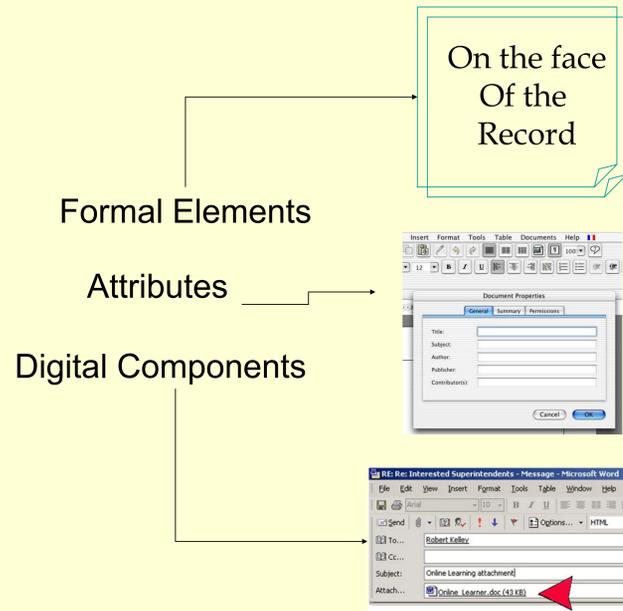
Dr. Luciana Duranti  
The University of British Columbia



## The Concept of Record



## Digital Record Characteristics



## Lifecycle of Digital Records

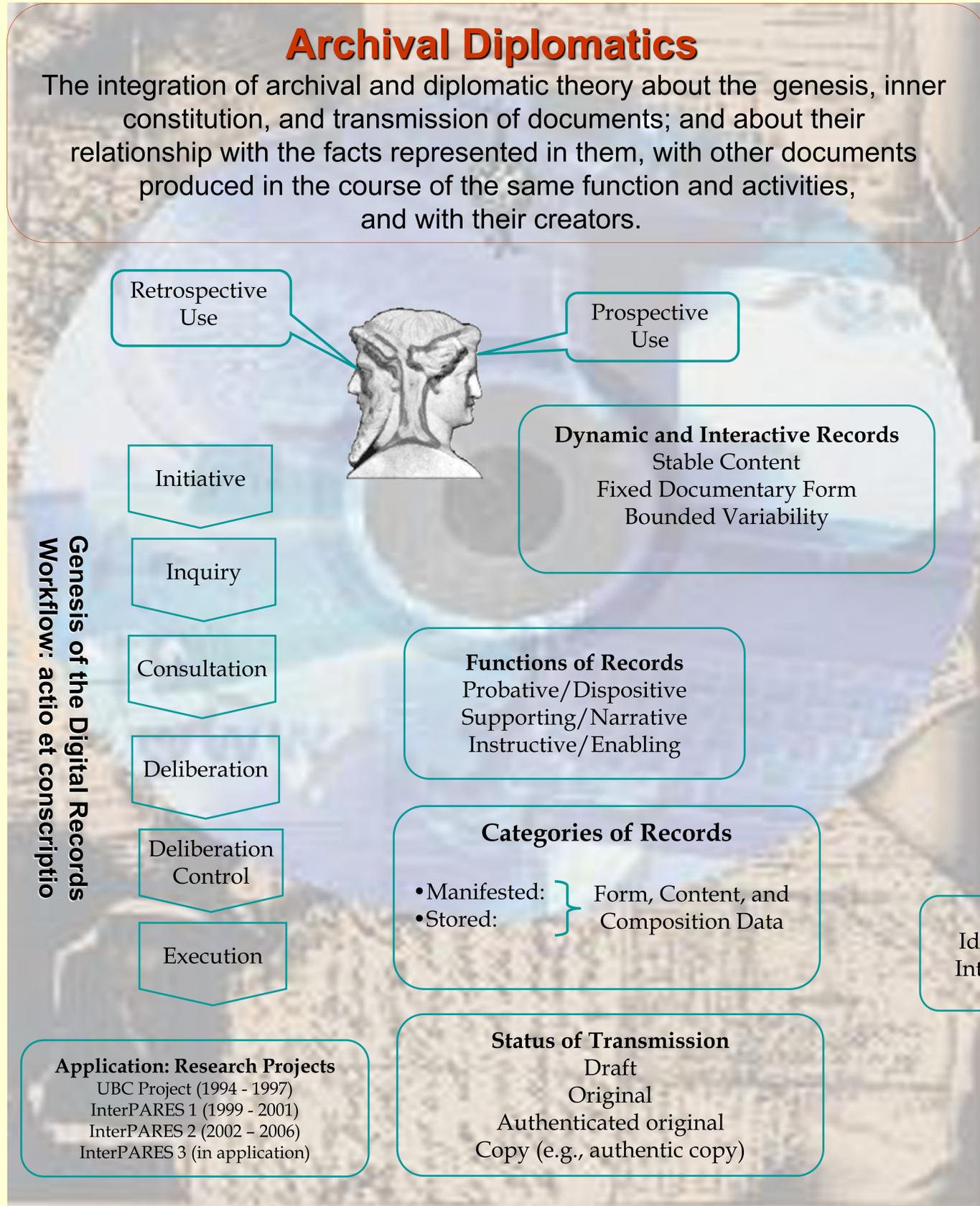
Phase 1: Records of the creator  
Phase 2: Authentic copies of the records of the creator



Luciana Duranti  
Email: [luciana@interchange.ubc.ca](mailto:luciana@interchange.ubc.ca)  
[www.interpares.org](http://www.interpares.org)

## Archival Diplomatics

The integration of archival and diplomatic theory about the genesis, inner constitution, and transmission of documents; and about their relationship with the facts represented in them, with other documents produced in the course of the same function and activities, and with their creators.



## The Concept of Trustworthiness

### Reliability

The trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about.

### Accuracy

The degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter.

### Authenticity

- identity
- integrity

The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption.



### Digital Signature



✓ As a Means of Authentication

**Metadata**  
Identity Metadata  
Integrity Metadata

Authenticity Requirements  
Benchmark Requirements  
Baseline Requirements

# The Object of Archival Diplomatics

- **Record:** any document created (i.e., made or received and set aside for action or reference) by a physical or juridical person in the course of activity as an instrument and by-product of it
- **Document:** recorded information (i.e., information affixed to a medium in an objectified and syntactic form)
- **Information:** a message intended for communication across time and space
- **Data:** the smallest meaningful piece of information

# Digital Record Characteristics

## Fixed Form:

- its binary content is stored so that the message it conveys can be rendered with the same presentation it had on the screen when first saved (different digital presentation)
- if the same content can be presented on the screen in several different ways in a limited series of possibilities, we have a different view of the same record having stable content and fix form (different documentary presentations e.g. statistical data as a pie chart, a bar chart, or a table )

# Digital Record Characteristics

- **Stable Content:** the fact that the data and the message in the record are unchanged and unchangeable, meaning that data cannot be overwritten, altered, deleted or added to
- **Bounded Variability:** when changes to the form are limited and controlled by fixed rules, so that the same query or interaction always generates the same result, and we have different views of different subsets of content, due to the intention of the author or to different operating systems or applications

# Digital Record Characteristics

- **Medium:** necessary part of the technological context, not of the record
- **Archival Bond:** explicit linkages to other records inside or outside the system
- **Five Necessary Persons:** author, writer, originator, addressee, and creator
- **Act:** an action in which the records participates or which the record supports
- **Five Necessary Contexts:** juridical-administrative, provenancial, procedural, documentary, technological

# Digital Record Characteristics

- **Formal Elements:** a constituent part of the record documentary form as shown on its face (intrinsic or extrinsic)
- **Metadata:** the attributes of the records that demonstrate its identity and integrity (authenticity)
- **Digital Components:** entities that either contain one or more records or are contained in the record and require a specific preservation measure

# Types of Digital Records

- **Static:** They 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
- **Interactive:** They present variable content, form, or both and the rules governing the context and form of presentation may be either fixed or variable

# Interactive Entities

- **Not-dynamic:** the rules governing the presentation of content and form do not vary, and the content presented each time is selected from a fixed store of data. Ex. Interactive web pages, online catalogs, records enabling performances—they are records
- **Dynamic:** the rules governing the presentation of content and form may vary—they are potential records

# Interactive Dynamic Entities

- Entities where the variation is due to data that change frequently (the design permits updating, replacement or alterations; allows data collection from users or about user interactions or actions or uses these data to determine subsequent presentations)
- Entities where the variation is due to data received from external sources and not stored within the system (VanMap)

# Interactive Dynamic Entities

- Entities produced in dynamic computing applications that select different sets of rules to produce documents, depending on user input, sources of content data, and characteristic of content (weather sites)
- Entities produced by evolutionary computing where the software generating them can change autonomously (scheduling and modeling of financial markets; edutainment sites)

# Stored and Manifested Records

- **Stored record:** a digital component used in reproducing more than one record and that includes the data to be processed in order to re-produce the manifested record and the rules for processing the data, including those enabling variations (think formularium=database form, database tables, logical model)
- **Manifested record:** the visualization or materialization of the record in a form suitable for presentation to a person or system (think deed)

Both have form, content and composition data

# Records Functions

- *Ad substantiam* and *ad probationem* (dispositive and probative=legal records)
- **Supporting:** generated to be used in the course of multiple activities as a source of information (e.g., GIS)
- **Narrative:** generated as an instrument of communication but not required by the juridical system (e.g., most e-mails, reports, web sites)

# Records Functions

- **Instructive:** delineates the form in which external data are to be presented (e.g., scores, scripts, regulations, manuals of procedure, instructions for filling out forms)
- **Enabling:** enable performance of artworks (software patches), execution of business transactions (interacting business applications), conduct of experiments (workflows generated and used to carry out the experiment of which it is , instrument, byproduct and residue), analysis of observational data (interpreting software)

# Status of Transmission

- The original is the first received record (the first complete effective record)
- The complete draft is the record saved to the system before being transmitted across time or space
- Retrieving the record creates copies in the form of original
- Inserts=vidimus
- Authenticated records
- Authentic copies

# Trustworthiness

- **Reliability:** the trustworthiness of a record as a statement of fact, based on the author's competence and the control on the genesis of the record
- **Authenticity:** the trustworthiness of a record as a record. An authentic record is one that has not been tampered with or otherwise corrupted, based on identity and integrity
- **Accuracy:** the correctness and precision of content, based on the competence of the author and the controls on the process by which data are recorded and transmitted through space (i.e., between persons, systems or applications) and time (i.e., when stored off line, or when the hardware or software is upgraded or replaced)

# Trustworthiness: Authenticity

**Identity** refers to the attributes of a record that uniquely characterize it and distinguish it from other records. Identity Metadata: the names of the 5 persons concurring in its creation; its date(s) of creation and transmission; the matter or action in which it participates; the expression of the archival bond; documentary form; digital presentation; the indication of any attachment(s); digital signature; name of the person responsible for the record

# Trustworthiness: Authenticity

**Integrity** is the wholeness and soundness of a record. A record has integrity if it is intact and uncorrupted, that is, if the message that it is meant to communicate in order to achieve its purpose is unaltered. This means that a record's physical integrity, such as the proper number of bit strings, may be compromised, provided that the articulation of the content and its required elements of form remain the same. The Integrity Metadata are: name(s) of handling persons over time; name of person responsible for keeping the record; indication of annotations; indication of technical changes (e.g. format, encoding, upgrading, changes to digital components, migration); indication of presence or removal of digital signature; planned removal from the system; transfer to a custodian, deletion; existence and location of duplicates outside the system

# Trustworthiness: Authentication

**Authentication:** a means of declaring authenticity at a point in time.

The **digital signature** is a seal, not a signature. It is functionally equivalent to medieval seals, which were not only a means of verifying the origin of the record and the fact that it was intact, but also made the record indisputable and incontestable, that is, had a non-repudiation function. The analogy is not perfect, because the medieval seal was associated exclusively with a person, while the digital signature is associated with a given person and a specific record, and because the former is an expression of authority, while the latter is only a mathematical expression.

# **Trustworthiness: Trusted Record-Making System**

A set of rules governing the making of records, and a set of tools and mechanisms used to implement these rules. In order to generate reliable and accurate records, every record-making system should include in its design integrated business and documentary procedures in a workflow structure linked to classification and file plan, record identity metadata schemes, records forms specifications, and record-making access privileges

# Workflow of Digital Records

Integration of *actio* and *conscriptio* (moment of action and moment of its documentation)

- Initiative
- Inquiry
- Consultation
- Deliberation
- Deliberation Control
- Execution

# **Trustworthiness: Trusted Recordkeeping System**

A set of rules governing the keeping of records, and a set of tools and mechanisms used to implement these rules. Every recordkeeping system should include in its design integrity metadata schemes, a classification scheme and filing plan, a linked retention schedule, a registration system, a retrieval system, recordkeeping access privileges, and procedures for maintaining authentic records.

# New Concept of Records Life-Cycle

Based on the recognition that re-productions of digital records

- if made by the creator in the course of and for the purposes of its business, are **records of the creator**, while
- if made by the preserver in the course and for the purposes of archival functions, are **authentic copies of the records of the creator**

# Digital Records Life-Cycle

It comprises two phases related to the status of transmission of the record:

Phase 1: Records of the creator: the re-productions behave and have to be treated as originals every time they are used and acted upon

Phase 2: Authentic copies of the records of the creator: they cannot be treated as originals because the creator has never used or acted upon them after re-production. If the records were reactivated we would again have the records of the creator

# As a Consequence...

- The creator can decide at any given time to stabilize its own fluid entities and to give them the most useful, accessible, interoperable form, or the form that best serves its present and projected needs, and have as a result an entity that we can call the record of the creator
- The preserver can only preserve what it receives from the creator by making an authentic copy of it, and has no right to stabilize it or alter its documentary form—only the format

# Alternatively

- The preserver could trade stability of content and fixity of form with the ability to track changes.  
Record=copy of last manifested entity+log of changes+ metadata
- The preserver could think of the record as existing in two modes: as a record in becoming when the entity is stored; record when the entity is accessed for use.  
Record=copy of each manifested entity + metadata

# It Depends on the Purpose of Preservation

1. To show the record as it was experienced by those who interacted with it when it was created—impossible
2. To re-create or re-execute the record as it was—emulation
3. Show parts of the record—migration of digital components
4. Keep the documentation of the record and of the interaction between it and the users when it was created
5. To enable a subsequent non-identical re-productions in which the essence of the record is conveyed, not necessarily its form and behaviour

# Appraisal of Digital Records

Composed of the following new activities:

- determine the entities-records to be acquired
- assessing the authenticity of the records
- determining the feasibility of preserving them
- monitoring the records until implementation
- keeping track of all system's and record's documentation

# Preservation

It involves the creation of authentic copies of the records of the creator. Their authenticity is guaranteed by:

- a controlled process of migration of the acquired records to the archives technological environment (always keeping the records 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
- the implementation and monitoring of privileges concerning the access, use and reproduction of the records within the archives

# Preservation (cont.)

- the establishment of procedures to prevent, discover, and correct loss or corruption of records, as well as
- procedures to guarantee the continuing identity and integrity of the records against media deterioration and across technological changes; and
- if authentication of individual records is required, by the existence of rules determining responsibility for and means of authentication.

# Archival Description

Archival description acquires a primary authentication function

- The authentication function of archival description is a collective attestation of the authenticity of the records of a *fonds* and of all their interrelationships as made explicit by their administrative, custodial and technological history, the illustration of their scope and content, and the hierarchical representation of the records aggregates
- The unique function of archival description is to provide an historical view of the records and of their becoming while presenting them as a universe in which the individuality of each member is subject to the bond of a common provenance and destination

# Sources

- L. Duranti and K. Thibodeau, “The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES,” *Archival Science* 6 (2006): 13-68
- L. Duranti, “Pour une diplomatique des documents électroniques,” *Bibliothèque de l’Ecole des chartes*, special issue on “Exportations de la diplomatique” (Paris: Ecole des chartes, 2004): 105-125
- InterPARES site: [www.interpares.org](http://www.interpares.org)