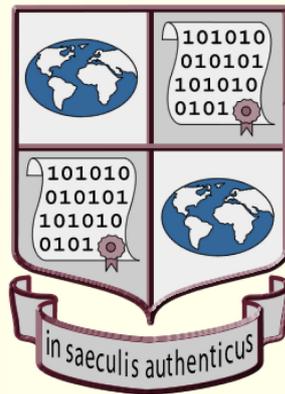


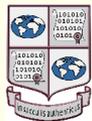
InterPARES Project

International Research on Permanent Authentic Records in Electronic Systems



The Appraisal of Digital Records

The View of InterPARES



InterPARES Project

Luciana Duranti

Project Director

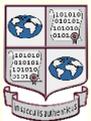
InterPARES 2 Goal

To ensure that the portion of society's recorded memory digitally produced in dynamic, experiential and interactive systems in the course of artistic, scientific and e-government activities can be created in accurate and reliable form and maintained and preserved in authentic form, both in the long and the short term, for the use of those who created it and of society at large, regardless of digital technology obsolescence and media fragility.



The selection function

- a process that starts at creation and continues till the records are acquired by the archives for permanent retention
- the prelude to preservation, increasingly part of preservation itself
- different from **appraisal**—which is an attribution of value to records, a value external to the records and defined by the acquiring institution or program—that is supposed to justify the choice of continuing retention for a possible permanent preservation



The selection of digital records

1. the preserver must assess the **authenticity** of the records considered of continuing value;
2. the preserver must determine the **feasibility** of the preservation of the authenticity of the records;
3. the disposal decision must be made very **early** in the life of the records; and,
4. the preserver must constantly **monitor** the records of the creator and, if warranted by the changes that they have undergone through time, revise the disposal decision



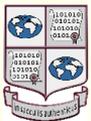
Assessing the authenticity of the records

- Never done before (Jenkinson: all records are authentic)
- It is based on the assessment of the identity and integrity of the records
- It can be done by inference if there are procedures in place that control the making and keeping of the records and their use
- It can be done by verification by comparison of multiple copies, recognition by the authors and other similar procedures
- If done long before acquisition it must be repeated before transfer to the archives



Determining the feasibility of preservation

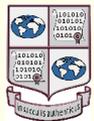
- Deciding whether the digital components embodying the essential elements that confer identity and ensure the integrity of the records can be preserved, given the preserver's current and anticipated capabilities
- A digital component is a digital object that contains all or part of the digital record, and/or data or metadata necessary to order, structure, or manifest its content, and that requires specific methods for preservation



Procedure for determining feasibility of preservation

Comprises three steps:

- The appraiser determines both the record elements containing informational content and those elements and attributes that need to be preserved according to requirements for authenticity, as formulated by the designated preserver (e.g. TNA).
- The appraiser identifies where these crucial record elements are manifested in the digital components of the record.
- The appraiser reconciles these preservation requirements with the preservation capabilities of the designated preserver



Monitoring the records

- All the records (not only those appraised for permanent preservation) must be checked on a regular basis
- This involves looking for changes in their technological context and in their use
- In some cases it may be necessary to repeat the appraisal because of changes that can affect the feasibility of preservation or because the records result from other functions or present different characteristics
- In most cases, monitoring produces minor revisions to the documentation on the selection and to the terms and conditions of transfer
- Example of the UBC Students Registration System



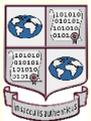
Radical changes

- Methodological complexity
- Scalability
- Quality of the records
- Focus on record creators and their function
- Appraisal as an iterative process (not re-appraisal)
- The material we appraise



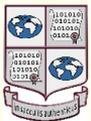
What is a Record?

- **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 Records Characteristics

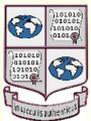
- **Medium:** physical carrier, part of the technological context, but the record must be affixed to it.
- **Stable Content and Fixed Form**
- **Archival Bond:** explicit linkages to other records inside or outside the system
- **Five Necessary Persons:** author, writer, originator, addressee, and creator
- **Act:** an action in which the records participates or which the record supports
- **Five Necessary Contexts:** juridical-administrative, provenancial, procedural, documentary, technological



Fixed Form and Stable Content

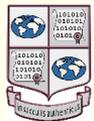
Fixed Form:

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



Fixed Form and Stable Content (cont.)

- **Stable Content:** 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



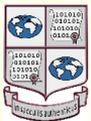
Manifested and Stored Record

- **Stored record:** the digital component(s) used in reproducing one or more than one record, which include the data to be processed in order to manifest the record (content data and form data) and the rules for processing the data, including those enabling variations (composition data)
- **Manifested record:** the visualization or materialization of the record in a form suitable for presentation to a person or system. Sometimes, it does not have a corresponding stored record, but is re-created from fixed content data when a user's action associates them with specific form data and composition data (e.g. a record produced from a relational database)



Types of Digital Records

- **Legal: dispositive**, e.g., contracts; **probative**, e.g., a land registry)
- **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, web sites)
- **Instructive**: delineate 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 (a workflow generated and used to carry out an experiment of which it is instrument, byproduct and residue), analysis of observational data (interpreting software), etc.



Digital Entities Behaviour

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—they are records if the other conditions are satisfied.

Interactive: They present variable content, form, or both, and the rules governing the content and form of presentation may be either fixed or variable—they may be records



Interactive Entities Behaviour

- **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**



VanMap
public edition

Contact Us

Legend Options

Address Search [Options](#) | [Reset](#) | [Help](#)

Number Street

or select

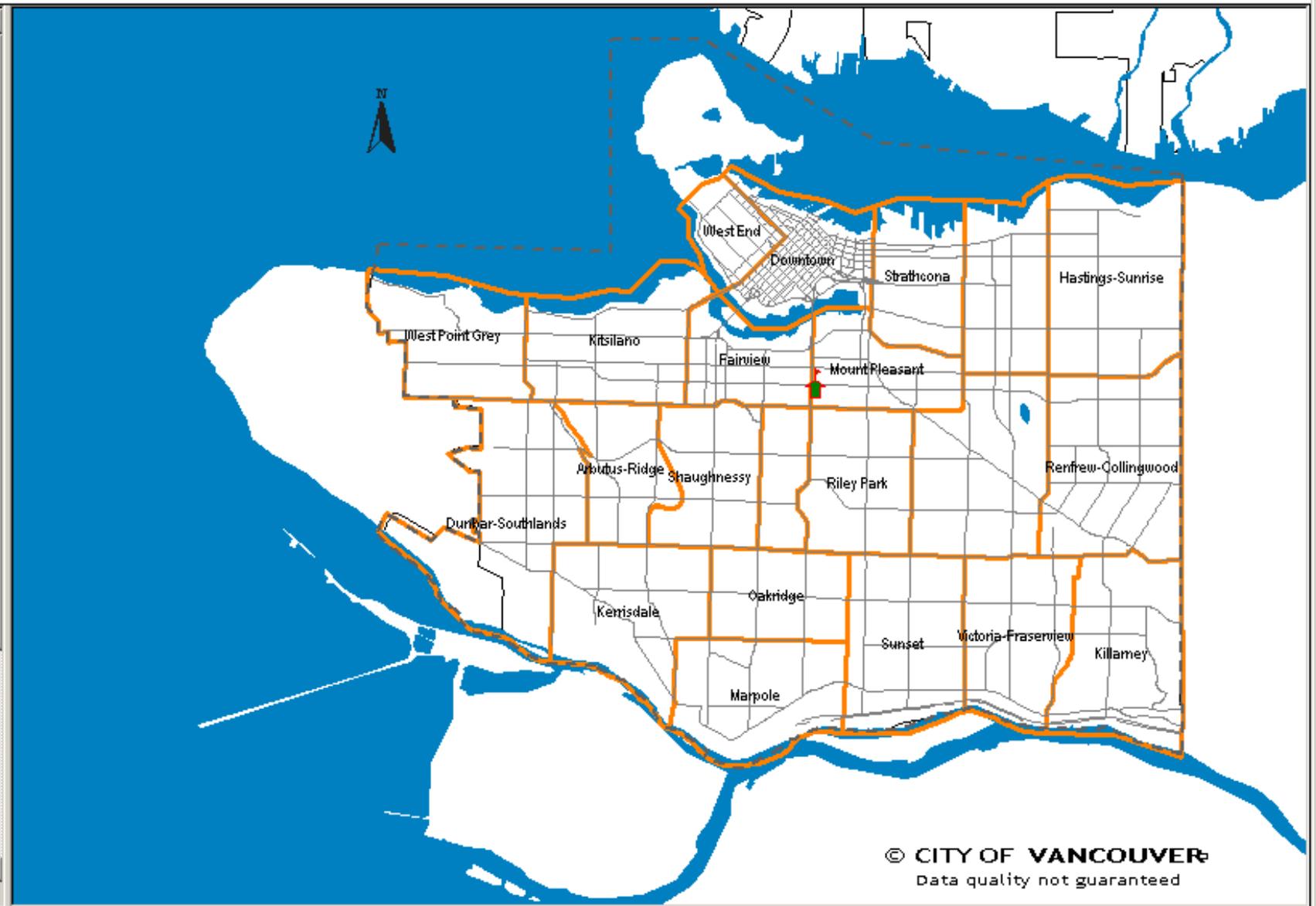
Toolbox [Right-Click Menu](#) | [Help](#)

Help [About the Data](#)

Applications

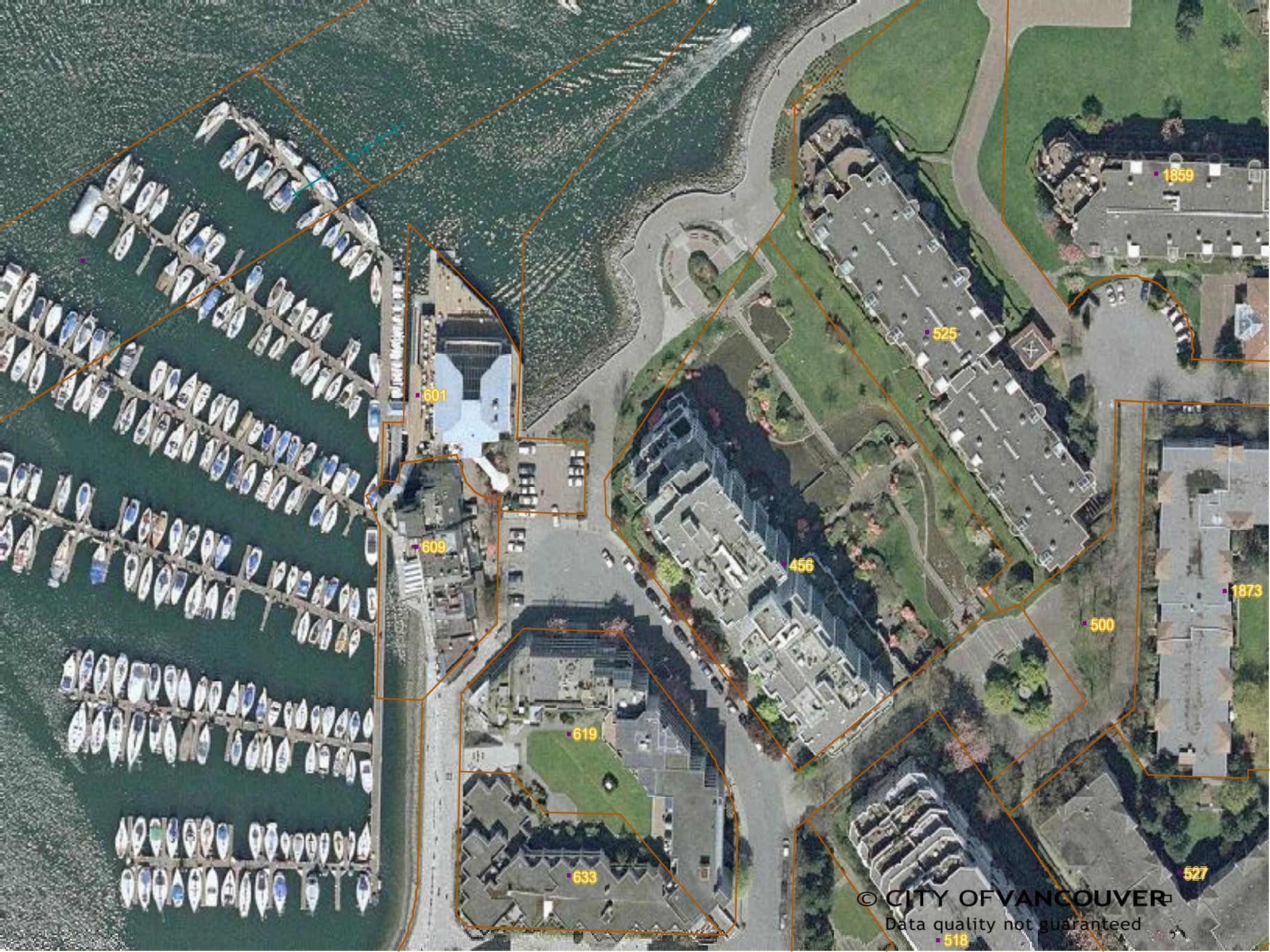
Select Application:

- Olympic Venues
- City Boundary
-
- Public Art
- Public Places
- The Road Ahead
- City Projects
- Child Care Facilities
- Web Cameras
- Non-Market Housing
- City-owned Properties
- Cemetery
- Urban Agriculture
- Contour Lines
- Truck Routes
- Canada Line Alignment
- Traffic Related
- Sewer
- Water
- City Streets Network
- View Cones
- Subdivision Categories
- DCL Areas
- Zoning Districts Types
- Zoning Districts
- Business Improvement Areas
- False Creek Navigable Chan
- Youth
- Property Information
- Facet Grids
- Administrative & Service Area
- Shore Lines (2002)
- Foreshore Contour Line
- Water Bodies



© CITY OF VANCOUVER
Data quality not guaranteed

Water Bodies : Strait of Georgia 0 feature(s) selected 1 : 109,043 14.4 x 10.0 (mi)



601

609

619

633

456

525

500

527

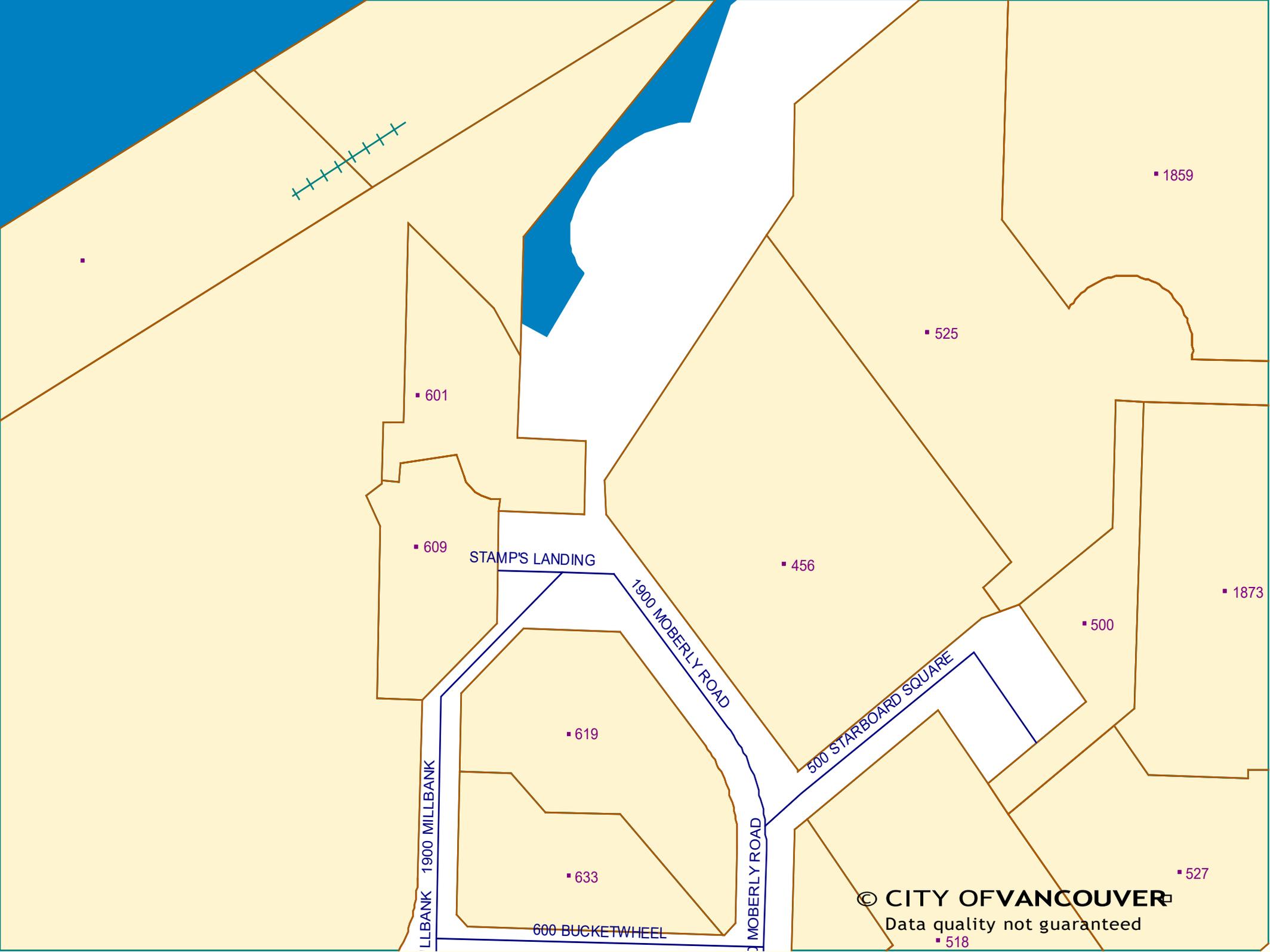
1859

1873

© CITY OF VANCOUVER

Data quality not guaranteed

518



▪ 1859

▪ 525

▪ 601

▪ 609

STAMPS LANDING

▪ 456

1900 MOBERLY ROAD

▪ 1873

▪ 500

500 STARBOARD SQUARE

▪ 619

1900 MILLBANK

▪ 633

MOBERLY ROAD

▪ 527

600 BUCKETWHEEL

© CITY OF VANCOUVER

Data quality not guaranteed

▪ 518



Print date: 13-Feb-07,01:52 PM

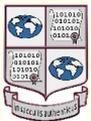
Tax Attributes Report

Source: The city's Property Tax System. (see [About Data](#) for the latest update.)

Record	PID or Strata Plan Number	Assessment Roll Number (Folio)	From Address Number	To Address Number	Street Name	Property Postcode	Lot	Block	District Lot	Plan	Legal Description	Previous Year Land Value(\$)	Previous Year Improvement Value(\$)	Current Year Land Value (\$)	Current Year Improvement Value(\$)	Assessment Year
	VAS1831	164632770001		601	STAMPS LANDING	V5Z 3Z1	1		FC	VAS1831	LOT 1 PLAN VAS1831 DISTRICT LOT F C NEW WESTMINSTER AN UNDIVIDED 795 /808 SHARE IN THE COMMON PROPERTY T HEREIN EX PLAN 18514, 18515, & 1851 6.	570000	1124000	584000	1135000	2007
	VAS1831	164632770002		601	STAMPS LANDING	V5Z 3Z1	2		FC	VAS1831	LOT 2 PLAN VAS1831 DISTRICT LOT F C NEW WESTMINSTER AN UNDIVIDED 13/ 808 SHARE IN THE COMMON PROPERTY TH EREIN EX PLAN 18514, 18515, & 18516 .	1200	39600	7800	40000	2007

VanMap Technical Components

- Oracle Spatial database
- Other databases
- CAD drawings, satellite imagery, photographs, html pages
- Autodesk MapGuide
- Autodesk ActiveX Viewer
- Application servers
- Web server



A Dynamic System

- Some data are overwritten without being saved



A Dynamic System

- Some data are overwritten without being saved
- The data are viewed as maps but these views are not saved



A Dynamic System

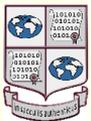
- Data are overwritten without being saved
- The data are viewed as maps but these views are not saved
- New layers are being added all the time



Is VanMap a Record?

Yes!

- It is made and received in the course of a practical activity
- It is an instrument and a by-product of that activity
- Five persons are involved in its creation and the 5 relevant contexts are identifiable
- It is an indivisible unit affixed to a medium



InterPARES Project

Luciana Duranti

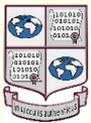
Project Director

Is VanMap a Record?

No!

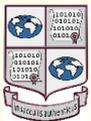
- It lacks fixed form and stable content
- It is not set aside for action or reference

Thus, it cannot be used to render an account of the decisions made.



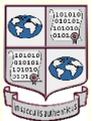
Can VanMap Become a Record?

- Yes, if we introduce fixed form and stable content
- We need to configure the system so that as each layer is updated the data are saved rather than overwritten
- Then we need to develop a means of reproducing VanMap as it was on any given date



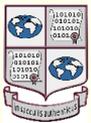
What About the Map Views?

- It is not feasible to require City staff to save the map views
- We would preserve not what the staff member *saw* at a given point in time but what s/he would have *been able to see*
- Improved business process documentation would fill in the gap



Building a Preservation Environment...

- Step 1: save the empty layers
- Step 2: add metadata to the layers
- Step 3: store the data in a secure environment
- Step 4: create infrastructure independence
- Step 5: migrate to new/neutral technology platforms
- Step 6: reproduce VanMap



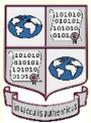
...using data grid technology

- Software developed by San Diego Supercomputer Center to manage large volumes of data
- Implemented as the Storage Resource Broker (SRB) which manages several large data repositories



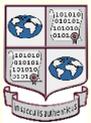
Data Grid Technology

- Manages data and their associated metadata



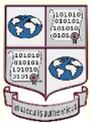
Data Grid Technology

- Manages data and their associated metadata
- Separates the data from dependence on original creating infrastructure



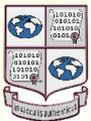
Data Grid Technology

- Manages data and their associated metadata
- Separates the data from dependence on original creating infrastructure
- Maintains audit trails of all operations performed on the data



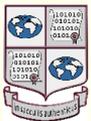
Data Grid Technology

- Manages data and their associated metadata
- Separates the data from dependence on original creating infrastructure
- Maintains audit trails of all operations performed on the data
- Manages access and retrieval



Data Grid Technology

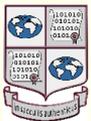
- Manages data and their associated metadata
- Separates the data from dependence on original creating infrastructure
- Maintains audit trails of all operations performed on the data
- Manages access and retrieval
- Supports migration of data to new platforms



Data Grids and VanMap

The scenario:

- Data grid is inserted between the data storage systems and the access applications
- Each saved layer within the GIS is independently registered in the data grid
- Date-based queries are used to reproduce VanMap layers



Testing the Data Grid

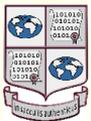
The test:

1. Selected data transferred from Vancouver to San Diego Supercomputer Center
2. Data stored in technological environment similar to original environment
3. Data registered in an SRB data grid
4. Data queried for specific dates
5. Queried data loaded into a different GIS product



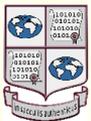
What Gets Preserved?

- The data themselves
- The ability to see the data available on a given day and time
- The ability to render the data as interactive maps
- Presentation elements such as colours and fonts do not necessarily have to be preserved given the costs of doing so



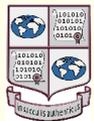
The same solution for all GIS?

- No!
- What is identified as the record to be generated and maintained over time depends on the use of the data and the purpose of having records rather than fluid information. Accountability is not always an issue as in the VanMap case.
- The research GIS of the Archaeological Society of Arizona requires preservation of its ability to make the users detect underground materials from the vegetation and stratification of the ground and therefore the records that suggest that excavations should be carried out.
- The Canada Atlas of Antarctica requires preservation of the content of external users' interactions with specific layers and therefore the users' records that have changed the output.



Case Study #2: the Alsace-Moselle Land Registry (1891-1924-present)

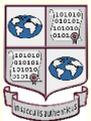
The registry is required by the French real estate law, as the means to fulfill the requirement that the juridical status of property (including the various forms of mortgages on the property) must be made publicly available to interested third parties by means of transcription within a land registry.



The Land Registry Function

A request by the notary for inscription in the registry must be verified by a land registry judge, an officer of the French State, who

- establishes the validity of the transaction,
- verifies the identity of all parties,
- verifies that all formal requirements relative to the inscription have been met, and
- verifies that no other rights on the property conflict with the new one.



The Land Registry Records

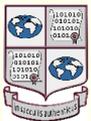
- An *ordonnance of inscription*, dictating the information to be transcribed within the register; because it is written and signed by the land registry judge, *the ordonnance* is considered to be *authentic documentary evidence within French evidence law*, that is, its trustworthiness is superior to all other written and oral evidence;
- The inscription within the register itself, transcribed by a specialized land registry clerk, and individually signed by the judge. The inscription is presumed to be valid until proof of the contrary, that is, interested parties may rely on its validity during the ordinary course of their business; it may however be contradicted by appropriate documentary evidence;
- The date of an application for inscription in the register fixes the legal effect of the rights, establishing a set order by which the third party rights on a property are recognized.



Computerization of the Registry

1994—creation of a specifically dedicated administrative body, the GILFAM (Groupement pour l'Informatisation du Livre Foncier d'Alsace-Moselle) to “facilitate and speed up the process of requests for new inscriptions, automate information exchange between the registry and the cadastre, optimize information storage and enable remote consultation of the register.”

2002-2006—data from 40,000 scanned registers were input in a database



Phase Two

- Deployment of a region-wide public-key infrastructure (PKI) which enabling judges to digitally sign each new entry into the database. The infrastructure integrates a number of state-of-the-art security technologies and methodologies, including digital signatures and biometric access.
- Research of an agreement between the GILFAM and the Archives de France on a method whereby land registry “records” may be transferred to the Archives for their historical value *while retaining the functionality of the registry database*.



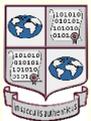
Change of Legislation

- Following the 1999 European Directive on digital signatures
- 2002, the Alsace-Moselle registry law was amended in order to recognize the legal value of the land register held on a data-processing media, in accordance with:
- Article 1316-1 of the French Civil Code: “Electronic written documents may be admitted as evidence in a manner similar to paper-based written documents, with the condition that the author of the document may be duly identified, and that it be manufactured and preserved under conditions which guarantee its integrity.”



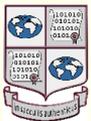
The Procedure of Inscription

1. Electronic requests for inscription (RIN) will be created by the notary using custom software, which will connect to the land registry in order to retrieve the information relative to the parties or parcels, as most of them will already be registered within the registry;
2. Once the request is received at the land registry office, it is immediately dated. The system does not allow entering a date anterior to the last date and time of entry; it is not possible to modify the date and time once it has been assigned to the request;
3. For each request, an electronic file is created containing all of the associated documents (contract, cadastre, etc.), as scanned imaged files;



The Procedure of Inscription (cont.)

4. A “projet d’ordonnance” is prepared, using the information contained in the RIN, if it was used. Inscriptions are created directly in the database, but are not visible to outside users of the database until a judge has signed them; the projet is transferred to the judge’s “in box” in the form of an XML document;
5. The judge is responsible for the required verifications; however, the custom software developed for the land registry provides him with a “before” and “after” view of the inscription, that is, of the changes to the registry which the inscription will effect in the database;
6. After identifying himself through biometric (fingerprint) scan and inserting a smartcard with his private signature key, the judge signs the “project d’ordonnance.” At that moment, in a single step, the ordonnance is created and signed, generating an inscription, and the relevant fields of the database are updated.



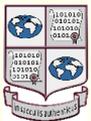
Digital Records in the Registry

- The ordonnance, which becomes the inscription, listing the information relative to the land parcel, the parties to the transaction, and the nature of the transaction. It is delineated in fields, using XML tags, and may thus be readily processed. It is authored by the judge, who dates and signs it.
- The tables of a relational database (that is, one table records the characteristics of land owners, another of land parcels, another of the *charges*, another of the mortgages); links between the tables (using primary/foreign key mechanisms) establish relationships between relevant data in the tables. It is now possible to query and list the information in the database based on any criteria. The two most important views offered by the computerized land registry are (a) the ownership history of a given land parcel and (b) the set of land parcels owned by a particular individual.



Technological Structure

- An Oracle database, containing the land registry data;
- Personal (Windows) computers, for registry clerks, running web-based applications for consulting the registry and managing the inscription process;
- Plugs-ins for commercial notarial software for integration with the land registry;
- Personal computers, for land registry judges, running web-based applications for consulting the registry and for finalizing inscriptions to the registry; in addition, those computers are equipped with biometric identification peripherals, and digital signature software;
- A PKI infrastructure, linking together all land registry offices and the central database, so that judges may sign ordonnances and add inscriptions to the registry.



Ordonnances

Each judge workstation is equipped with the necessary software to digitally sign ordonnances. To secure access to the judge's private key, the system uses a three-part authentication process:

1. The judges must have in their possession a smartcard holding their private signature key; the card is introduced in a card reader connected to the judge's workstation;
2. After entering the card, the judges must provide biometric identification using a fingerprint scanner;
3. Judges must also provide a password to finally gain access to their signing key.

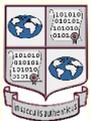


Inscriptions

An inscription is not a single bitstring, but is constituted from data stored in different fields of a database record.

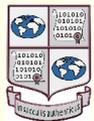
Two software mechanisms ensure that the content of the database remains in accord with the signed ordonnances at the procedural and legal origin of the inscription process:

1. A top-down procedure traverses sequentially all of the ordonnances stored on the system, verifying each digital signature and comparing the ordonnance to the inscription, determine that the inscription linked to the ordonnance has not been modified;
2. A bottom-up procedure performs a similar process, but proceeds from the inscriptions;



What Is New?

- Rather than expecting digital signatures to perform a one-time service of demonstrating the authenticity of the ordonnance in a court of law, it uses the digital signatures to provide continuous authentication services, that is, regularly performed declarations of the integrity and origin of the data;
- Digital signatures provide an extreme assessment of the integrity of data: if even a single bit of the signed data is modified, the signature fails. Using a comparison mechanism between the inscriptions and the signed ordonnances provides the necessary flexibility for system evolution when migration changes the bitstrings, but what about interoperability?

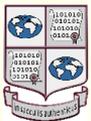


Preservation issues

The creator is mandated to preserve all of the digital entities under examination for regular transfer to the District Archives.

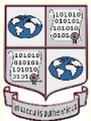
The digital entities are kept within a relational database system. Data relative to the inscriptions are kept within tables, linked together through relationships.

The ordonnances and their signatures are kept as stand-alone files, and linked to the relevant inscriptions.



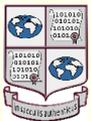
Preservation Issues (cont.)

- The creator has not considered the problem of maintaining the digital entities, except through the mechanisms afforded by system vendors when upgrading the database management system.
- The creator has not considered the problem of maintaining the evidential value of digital signatures through technological evolution.



Preservation Issues (cont.)

- While the transfer of the ordonnances to the District Archives, as stand-alone documents, poses no particular problems, the transfer of the inscriptions does
- As a record, the land registry cannot be understood outside of its dynamic and interactive capabilities.



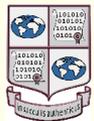
Possible Solutions

- The definition of an XML schema which may serve as a translation device between the complex data model used by the land registry, and a less complex model, to be defined, sufficient to satisfy the needs of future researchers. Inscriptions could then be exported to a file according to the XML schema and imported into relational database sufficiently simple to be maintained by an archival institution (e.g., Microsoft Access).
- The Direction des Archives de France could grant to GILFAM the permission to act as the custodian of the land registry. Under such an agreement, GILFAM would be mandated to transfer its inactive records to an archival database and, using the same software infrastructure, provide access to the records as determined by the Direction des Archives de France. Under such a model, the full dynamic and interactive capabilities of the preserved land registry would remain, that is, as long as the land registry itself remains in operation.

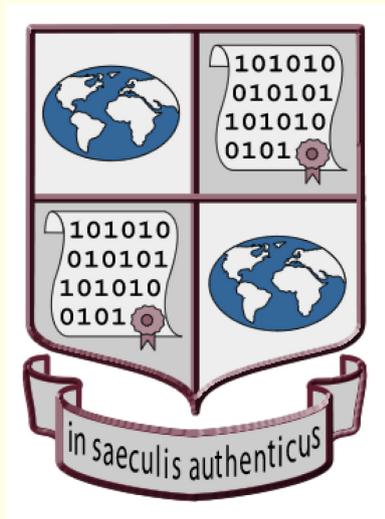


Conclusions

- Appraisal, differently from selection, is conditioned by context
- It requires a clear relationship between the creator and the designated preserver
- Selection and appraisal must be based on trust
- Appraisal must be clearly motivated
- The criteria must not only be archival, but also legal, ethical, and financial
- The archivist as “jack of all trades.”



InterPARES Web Site



www.interpares.org



InterPARES Project

Luciana Duranti
Project Director