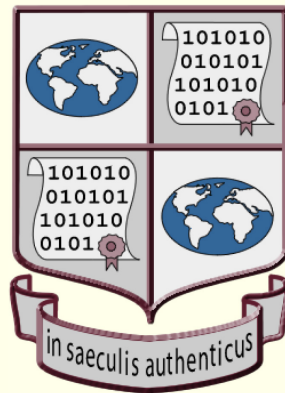# InterPARES Project

International Research on Permanent Authentic Records in Electronic Systems

The Appraisal of Digital Records

Assessing More Than Value

Luciana Duranti
The University of British Columbia

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, for the use of those who created it and of society at large, regardless of digital technology obsolescence and media fragility.

The Central Question

Has appraisal as we understood it in the paper world changed when confronted with the digital environment?

To answer we need to distinguish between The concept of appraisal and that of selection

Selection and Appraisal

- **Selection** is a process that starts at creation (some say before creation) and continues till the records are acquired by the archives for permanent retention. It is the prelude to preservation, increasingly part of preservation itself

- **Appraisal** is a decision, resulting from an attribution of value to records, a value defined by the preserver. Appraisal is meant to justify the choice of continuing retention for a possible permanent preservation

Selection of Digital Records

1. As early as possible in the life of the records, assess their continuing value to the creator and/or the designated preserver (appraise)

2. Assess the authenticity of the records considered for continuing preservation

3. Determine the feasibility of their preservation by the designated preserver (creator and/or delegated archives); and,

4. Constantly monitor all the records of the creator and, if warranted by the changes that they have undergone through time, revise the appraisal decision

Assessing the Authenticity of the Records

- Never done before (Jenkinson: all records are authentic) because the records we kept over time were physically the same made or received and kept by the creator

- It is based on the assessment of the identity and integrity of the records either by inference if there are procedures in place that control the making and keeping of the records and their use, or by verification, comparing multiple copies, through recognition by the authors and other similar procedures

- If done long before acquisition by the preserver it must be repeated before transfer to the archival program or institution

Determining the Feasibility of Preservation

- It means 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

- Digital components are to the stored record what elements of form are to the manifested record. All digital records have digital components, but not all have elements of form.

Monitoring the Records

- **All** the records (not only those appraised for continuing preservation) must be checked on a regular basis

- This involves looking for changes both 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

- The theory of appraisal has not changed: same values (including authenticity, which with paper was presumed to be a quality of all records)

- Fulfilling its requirements requires a selection process of a methodological complexity never needed in the paper environment:
  a) team work of all stakeholders and competent professionals: creator, designated preserver, IT and law specialists
  b) scalability of the process: action plans must be specific to the creator and the preserver
  c) focus on the creator and its functions: appraising the same records or digital objects from different viewpoints at the same time and many times before the disposal action
  d) selection of the most appropriate form of the record (music, science)
  e) identification of the records in systems containing different kinds of information, and, if the records do not exist, but should exist, ...creation of the records?

Case Study #1:
the Alsace-Moselle Land Registry

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

The Procedure of Inscription

- An electronic request for inscription is generated by the notary using custom software, which connects to the land registry in order to retrieve the information relative to the parties or parcels

- Once the request is received at the land registry office, it is dated. This date determines the inception of the rights on the property.

- For each request, an electronic file is created containing all of the associated documents (contract, cadastre, etc.), as scanned imaged files where they do not exist as digital data sets to which the request can be linked

The Procedure of Inscription (cont.)

4. A draft order of inscription is prepared. Inscriptions are also drafted directly in the database, but are not visible to outside users of the database until a judge has signed them; the draft order 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 of the land registry office provides him/her 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 draft order. At that moment, in a single step, the order is generated and signed, producing an inscription, and the relevant fields of the database are updated.

Digital Records in the Registry

- **The order**, 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); with links between the tables (using primary/foreign key mechanisms) that establish relationships between relevant data in the tables. 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 and 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 orders and add inscriptions to the registry.

What Is New?

- The system 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.

- They also compare the orders with the inscriptions every time their authenticity is questioned

Appraisal Issues

- While the acquisition of the orders by the District Archives, as stand-alone documents, poses no particular problems, that of the inscriptions does
- The digital inscriptions are not records, the land registry as a whole is.
- As a record, the land registry cannot be understood outside of its dynamic and interactive capabilities.
- The inscriptions cannot be authenticated outside the PKI infrastructure
- Migration to overcome obsolescence risks loss of interoperability

Possible Solution

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 users.

Inscriptions could then be exported to a file according to the XML schema and imported into relational database sufficiently simple to be maintained by the designated preserver (e.g., Microsoft Access).

Case Study #2: the VanMap

The cross-corporate GIS created by the City of Vancouver and used by staff in

- Engineering
- Planning
- Permits and Licenses
- By-law Enforcement

- Social Planning
- Police
- Fire and Rescue
- Parks and Recreation

VanMap
public edition

Contact Us

Legend Options

Address Search
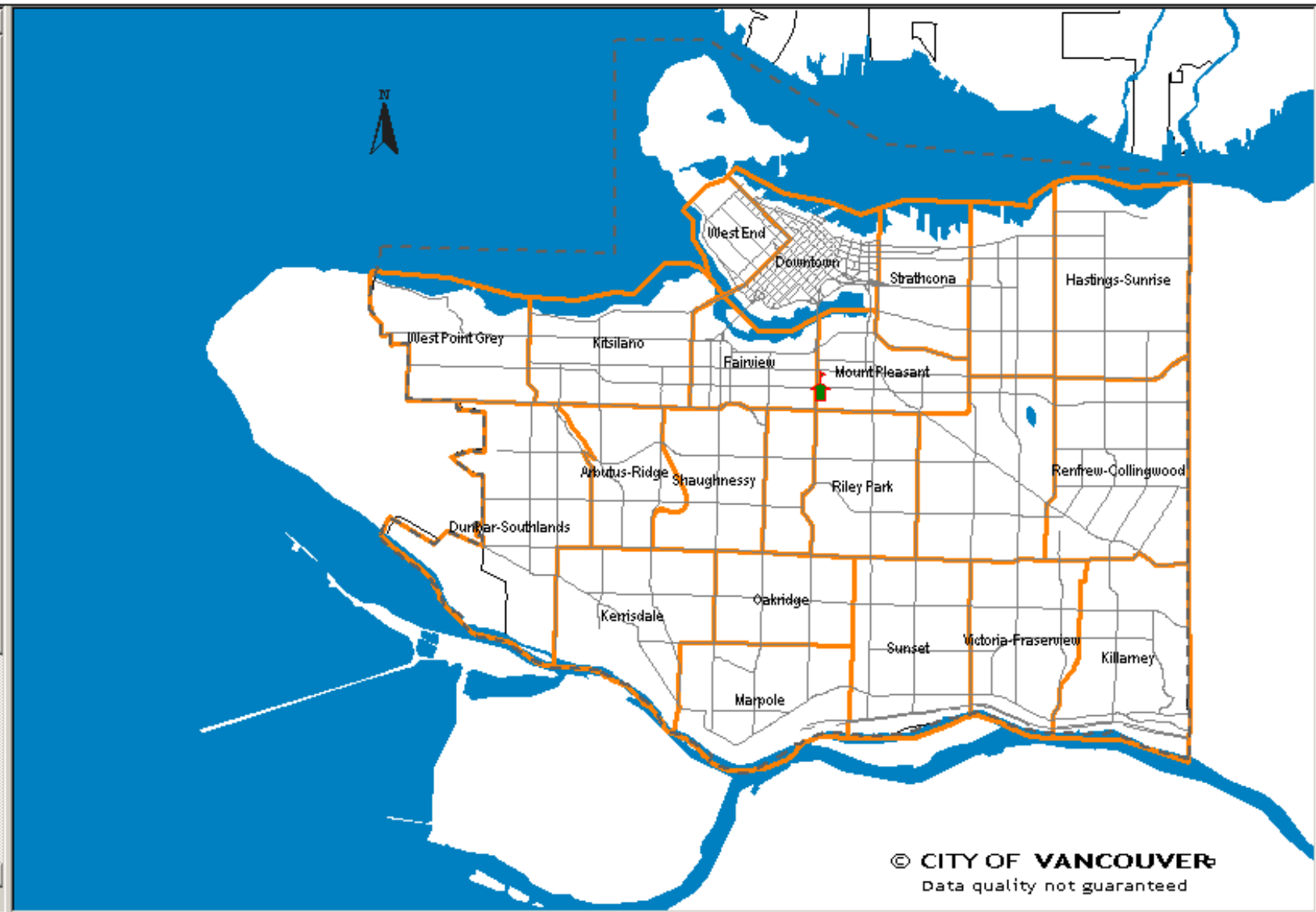
Options | Reset | Help

Number   Street

or select   E 1st Av   Go

Toolbox

Right-Click Menu | Help

STOP

Help
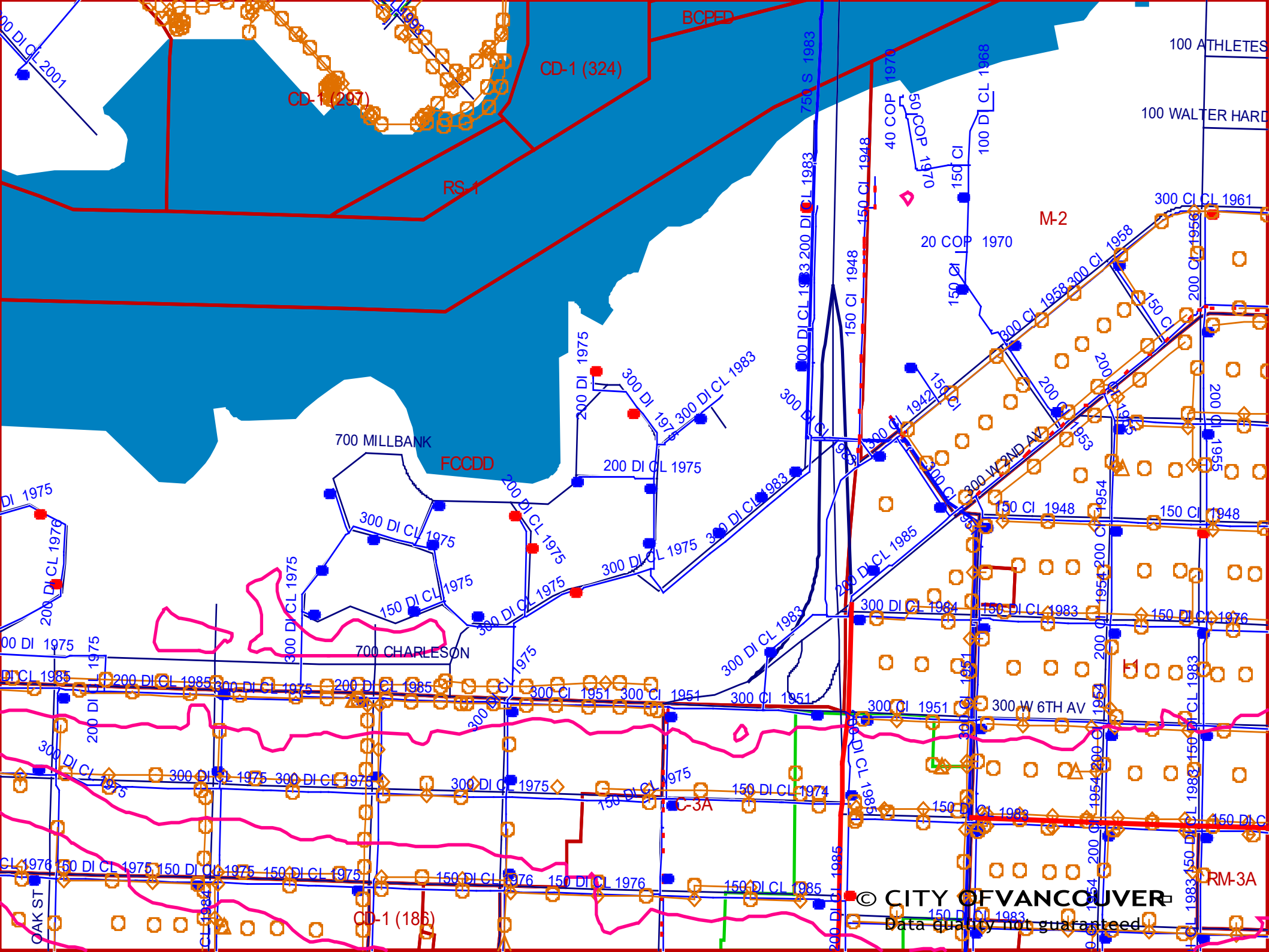
About the Data

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

Map labels: West End, Downtown, Strathcona, Hastings-Sunrise, West Point Grey, Kitsilano, Fairview, Mount Pleasant, Arbutus-Ridge, Shaughnessy, Riley Park, Renfrew-Collingwood, Dunbar-Southlands, Kerrisdale, Oakridge, Sunset, Victoria-Fraserview, Killarney, Marpole

© 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)

Internet

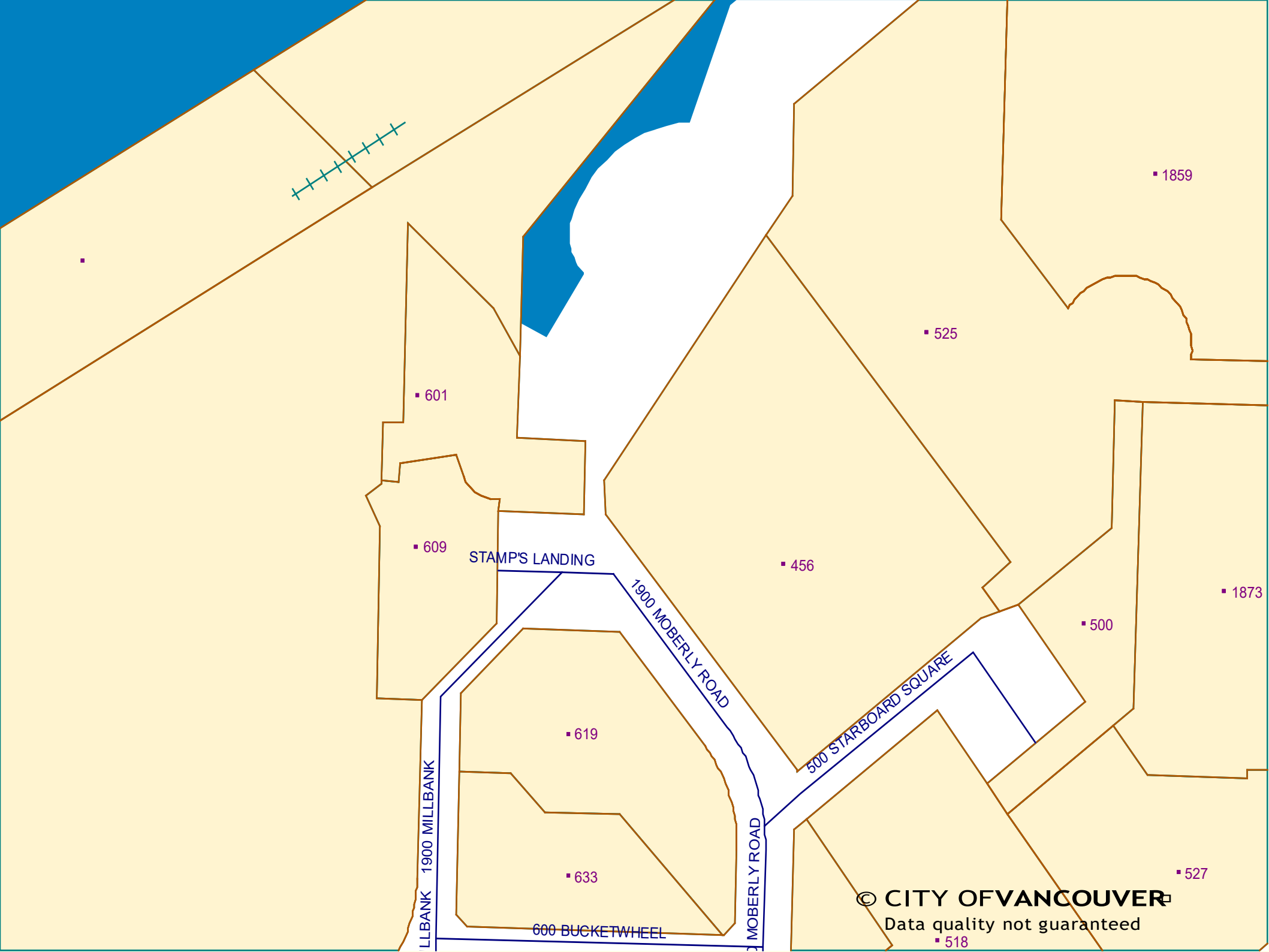Start | Inbox - Micro... | 2 Windows ... | Microsoft Excel | Symposium 2... | Presentation1 | ACA presenta... | 2 Internet ... | 12:42 PM

BCPED

CD-1 (324)

100 ATHLETES

CD-1 (297)

100 WALTER HARD

300 DI CL 2001

RS-1

M-2

300 CI CL 1961

300 CI CL 1958

20 COP 1970

40 COP 1970

50 COP 1970

150 CI

100 DI CL 1968

150 CI 1948

750 S 1983

200 DI CL 1983

300 DI CL 1983

150 CI 1948

200 DI 1975

300 DI 1975

300 DI CL 1983

300 DI CL 1975

200 DI CL 1975

300 CI CL 1958

300 CI CL 1958

150 CI

200 CI

1942 CI

300 W 2ND AV

700 MILLBANK

FCCDD

200 DI CL 1975

300 DI CL 1975

300 DI CL 1975

300 DI CL 1983

300 DI CL 1975

200 DI CL 1975

300 DI CL 1983

150 CI 1948

150 CI 1948

1954

150 CI 1948

DI 1975

300 DI CL 1975

150 DI CL 1975

300 DI CL 1975

300 DI CL 1975

300 DI CL 1985

300 DI CL 1964

150 DI CL 1983

150 DI CL 1976

200 DI CL 1976

150 DI CL 1975

700 CHARLESON

DI CL 1985

200 DI CL 1985

300 DI CL 1975

200 DI CL 1985

300 CI 1951

300 CI 1951

300 CI 1951

300 CI 1951

300 W 6TH AV

200 DI CL 1985

300 DI CL 1975

300 DI CL 1975

300 DI CL 1975

150 DI CL 1975

150 DI CL 1974

200 DI CL 1985

C-3A

OAK ST

CL 1976

150 DI CL 1975

150 DI CL 1975

150 DI CL 1975

150 DI CL 1976

150 DI CL 1976

150 DI CL 1985

200 DI CL 1985

RM-3A

CD-1 (186)

150 DI CL 1983

601

609

619

633

456

525

500

1859

1873

527

518

▪ 1859

▪ 525

▪ 601

▪ 609

STAMP'S LANDING

1900 MOBERLY ROAD

▪ 456

▪ 1873

500 STARBOARD SQUARE

▪ 500

▪ 619

1900 MILLBANK

MOBERLY ROAD

▪ 633

▪ 527

600 BUCKETWHEEL

© CITY OF VANCOUVER
Data quality not guaranteed

▪ 518

VanMap
public edition

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

Tax Attributes Report

urce: The city's Property Tax System. (see About Data for the latest update.)

ecord	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 ($)	Current Year Improvement Value($)	Asse
	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 linked to it, existing in a variety of offices, of a variety of local authorities, and whose data flow continuously in the Oracle database
- CAD drawings, satellite imagery, photographs, html pages
- Autodesk MapGuide
- Autodesk ActiveX Viewer
- Application servers
- Web server

VanMap is
a Dynamic Information System

- Data that often do not exist anywhere else, especially in the correlated form showed on the GIS layers, 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

- VanMap does not contain records

Is VanMap as a Whole 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

- It is an indivisible unit affixed to a medium

- It has all the diplomatic characteristics of a record in terms of persons and contexts

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 or as a memorial.

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 Taking Map Views?

- The preserver cannot do so or would become the creator of digital objects never used by the creator in the course of business, a creator of his/her own records

- It is not feasible to require City staff to save the map views in connection with the decisions based on them

- We have to preserve not what the staff member *saw* at a given point in time but what s/he *would have been able to see*

- A detailed documentation of the business process would support this preservation activity

How to Build a GIS 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 the system

Using 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 Grid and VanMap

- 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

- Queries based on the chronological date of the data are used to reproduce VanMap layers

- Queried data are loaded into a different GIS product

What Is 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 by the creator and the reason for having records rather that fluid information.

- The research GIS of the Archaeological Society of Arizona requires preservation of its ability to make the users detect underground materials from the layers showing vegetation and stratifications of the soil and therefore preservation of the records that suggest that excavations should be carried out.

- The Canadian Atlas of Antartica requires preservation of the content of external users interactions with specific layers and therefore preservation of the users' records that have changed the system output.
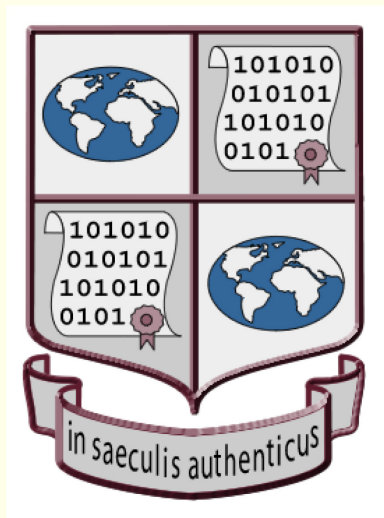
Conclusions

- **Appraisal**, differently from selection, is entirely conditioned by the **specific 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** on the basis of arguments that are not only be archival, but also legal, ethical, and financial

- The archivist as "jack of all trades" must be all to all records, including the **potential records**. It is a very tall order, but we can meet the challenge if we do not turn away from it, and might have much fun in the process.

InterPARES Web Site

www.interpares.org