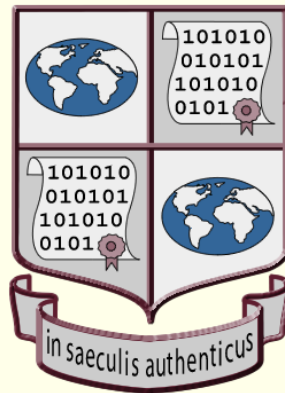


InterPARES Project

International Research on Permanent Authentic Records in Electronic Systems



The Appraisal of Digital Records

The Return of Diplomats as a Forensic Discipline

Luciana Duranti
The University of British Columbia



InterPARES Project

Luciana Duranti
Project Director

The Central Question

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

To answer we need first 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
 - The questions are: when, what, by whom, for whom, where, how often.
- **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, in different contexts, 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?**



Understanding the Record

“The reality of the record base must be an indispensable component of all acts of appraisal. **Without an understanding of documents and records, of their forms and of their functions, and of how they were created and used**, a plan can be so easily upset by the attractiveness of concentrating on information divorced from the realities of its documentary expression...., it is the record which is our special area of knowledge; it will be a sad day and a dangerous step when faith in planning replaces the study and knowledge of records.”

(Craig, The Acts of the Appraisers: The Context, the Plan and the Record 1992)



The Concept of Record

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



Digital Record Characteristics

- **Act:** an action in which the record participates or which the record supports (naturalness and impartiality)
- **Persons Concurring to Its Creation:** author, writer, originator, addressee, and creator
- **Archival Bond:** explicit linkages to other records inside or outside the system (uniqueness)
- **Identifiable Contexts:** juridical-administrative, provenancial, procedural, documentary, technological (interrelatedness)
- **Medium:** necessary part of the technological context, not of the record
- **Fixed Form and Stable Content**



Fixed Form

- An entity has 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)
- An entity has fixed form also 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)



Stable Content

- An entity has stable content if **the data and the message it conveys are unchanged and unchangeable**, meaning that data cannot be overwritten, altered, deleted or added to
- **Bounded Variability**: when **changes** to the documentary presentation of a determined stable content **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 (cont.)

- **Formal Elements:** constituent parts of the record documentary form as shown on its face, e.g. address, salutation, preamble, complimentary close
- **Metadata:** the attributes of the records that demonstrate its identity and integrity
- **Digital Components:** stored digital entities that either contain one or more records or are contained in the record and require a specific preservation measure



Stored and Manifested Records

- **Stored record:** it is constituted of the digital component(s) used in re-producing it, which comprise 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). Stored records may not have a corresponding manifested record(s).
- **Manifested record:** the visualization of the record in a form suitable for presentation to a person or a system. Sometimes, it does not have a corresponding stored record, but it 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

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 content and form of presentation may be either fixed or variable



Interactive Entities

- **Non-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** if all other necessary characteristics are present
- **Dynamic:** the rules governing the presentation of content and form may vary—**they are either information systems** if the other necessary characteristics are not present, **or potential records**, if they are present



Records Functions

(the way a record relates to an action)

- **Dispositive**: the record is the act (e.g., contracts)
- **Probative**: the record follows the act but is required to prove it (e.g., registries)
- **Supporting**: generated to be used in the course of activity (ies) as a source of information, often by multiple users (e.g., GIS)
- **Narrative**: generated on a purely discretionary basis only as a means of communication (e.g., most e-mails, memos, web sites)



Records Functions

- **Instructive:** provide **guidance** on the way in which external data or documents are to be presented (e.g., scores, scripts, regulations, manuals of procedure, instructions for filling out forms)
- **Enabling:** **enable** the performance of artworks (software patches), the execution of business transactions (interacting business applications), the conduct of experiments (a workflow generated and used to carry out an experiment of which it is instrument, byproduct and residue), the analysis of observational data (interpreting software), etc. Most of them are stored only records.



Interactive Information Systems

- 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)



Interactive Potential Records

- Entities where the variation is due to **data that change frequently**, because 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 (e.g. Land Registry)
- Entities where the variation is due to **data received from external sources** and not stored within the system (e.g. GIS)

They are presently not records but should be made into records if they fulfill one of the six records functions.



Some Key Questions

- If it is not possible to have a record in fluid form and with undetermined boundaries (i.e. an interactive dynamic record), should an entity with fixed form and stable content be generated for the purpose of making a record to be kept in a trusted recordkeeping system and perhaps preserved over the long term?
- If yes, **who** should make it?
- On the basis of **which criteria**?
- **When** in the entity's lifecycle?



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 (record)**, 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 (not records)** (i.e., 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.



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.



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



Address Search

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

Number

Street

or select

E 1st Av

Go

Legend Options 

Toolbox

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



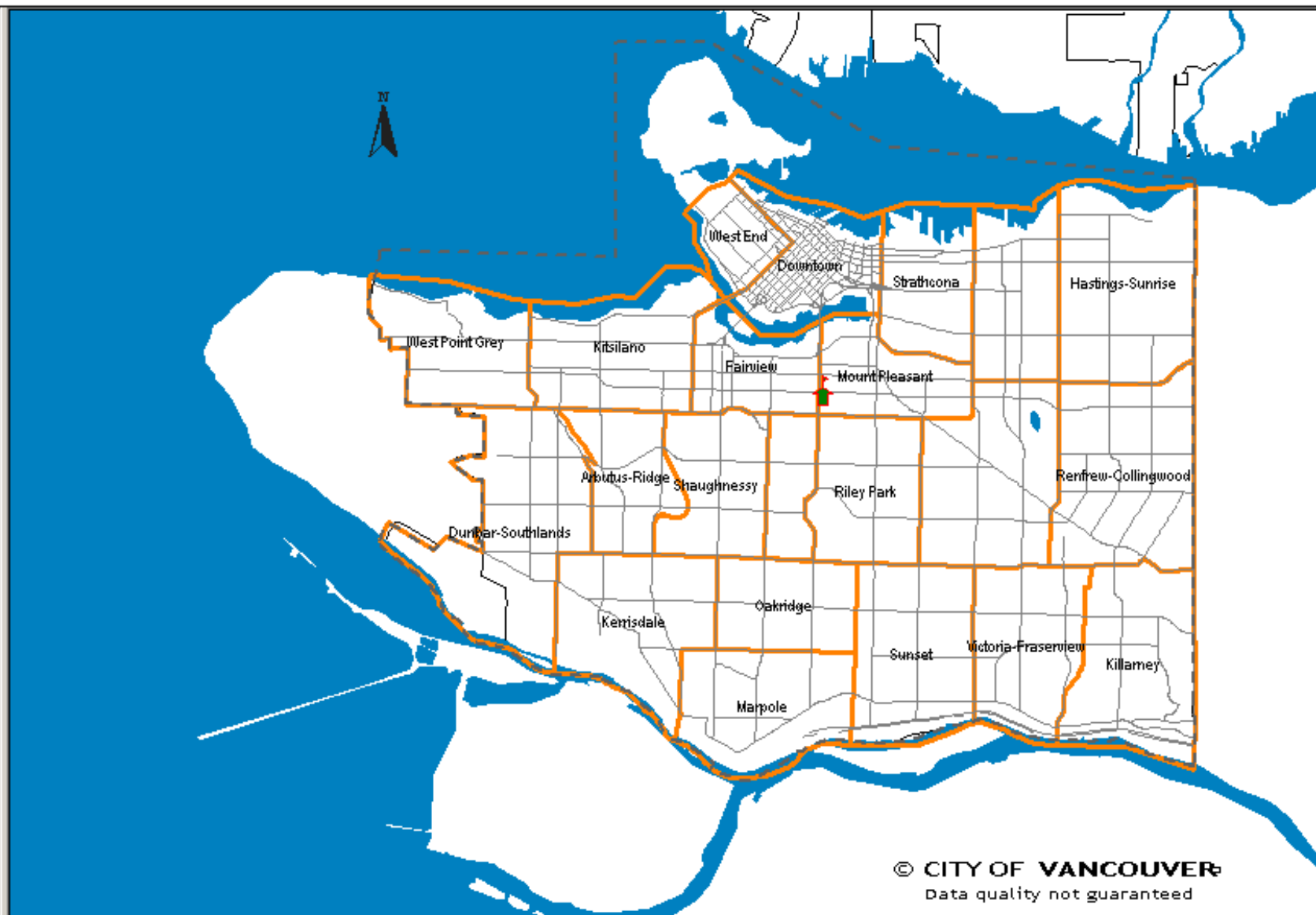
Help

[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

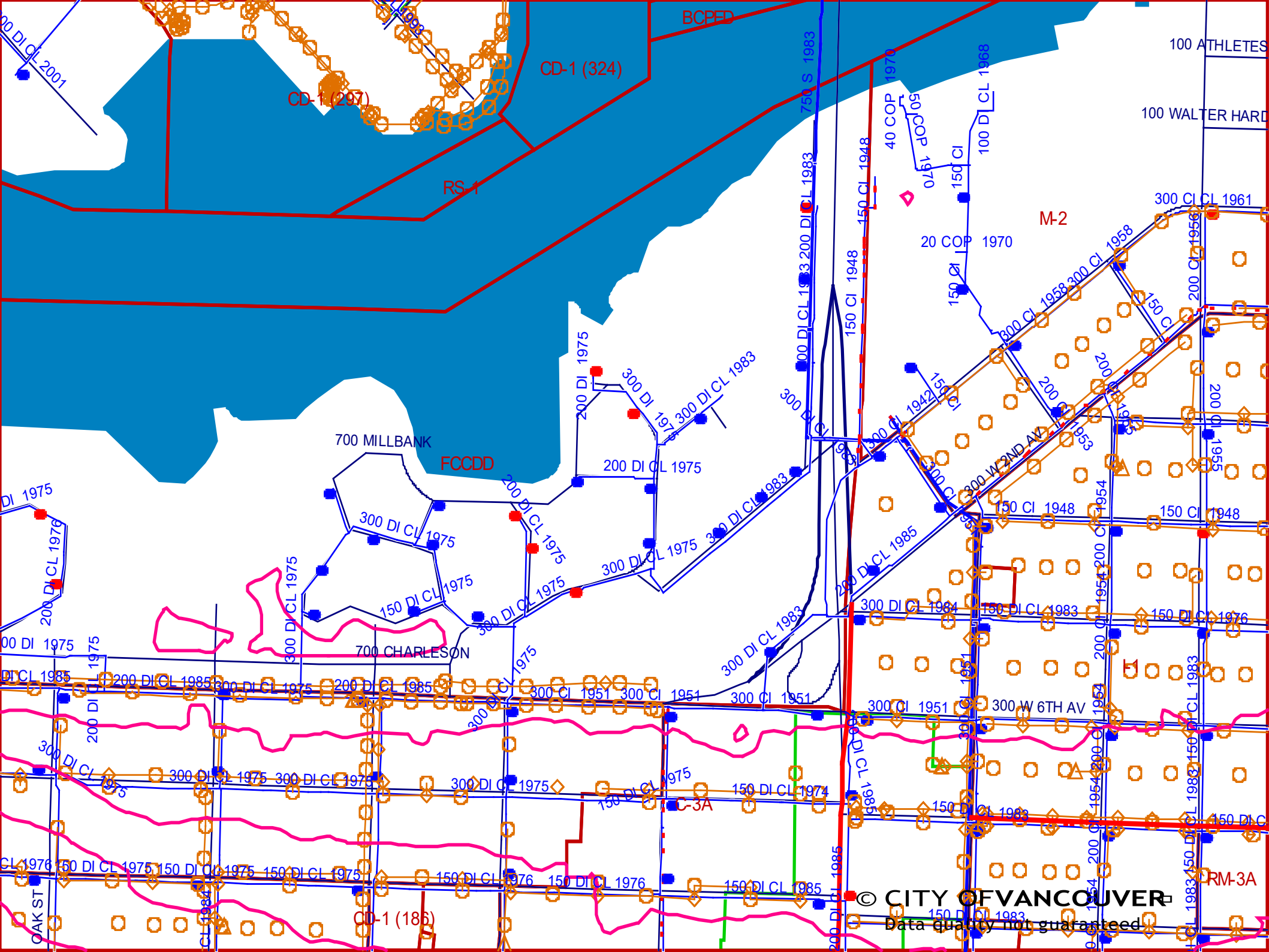
Water Bodies : Strait of Georgia

0 feature(s) selected

1 : 109,043

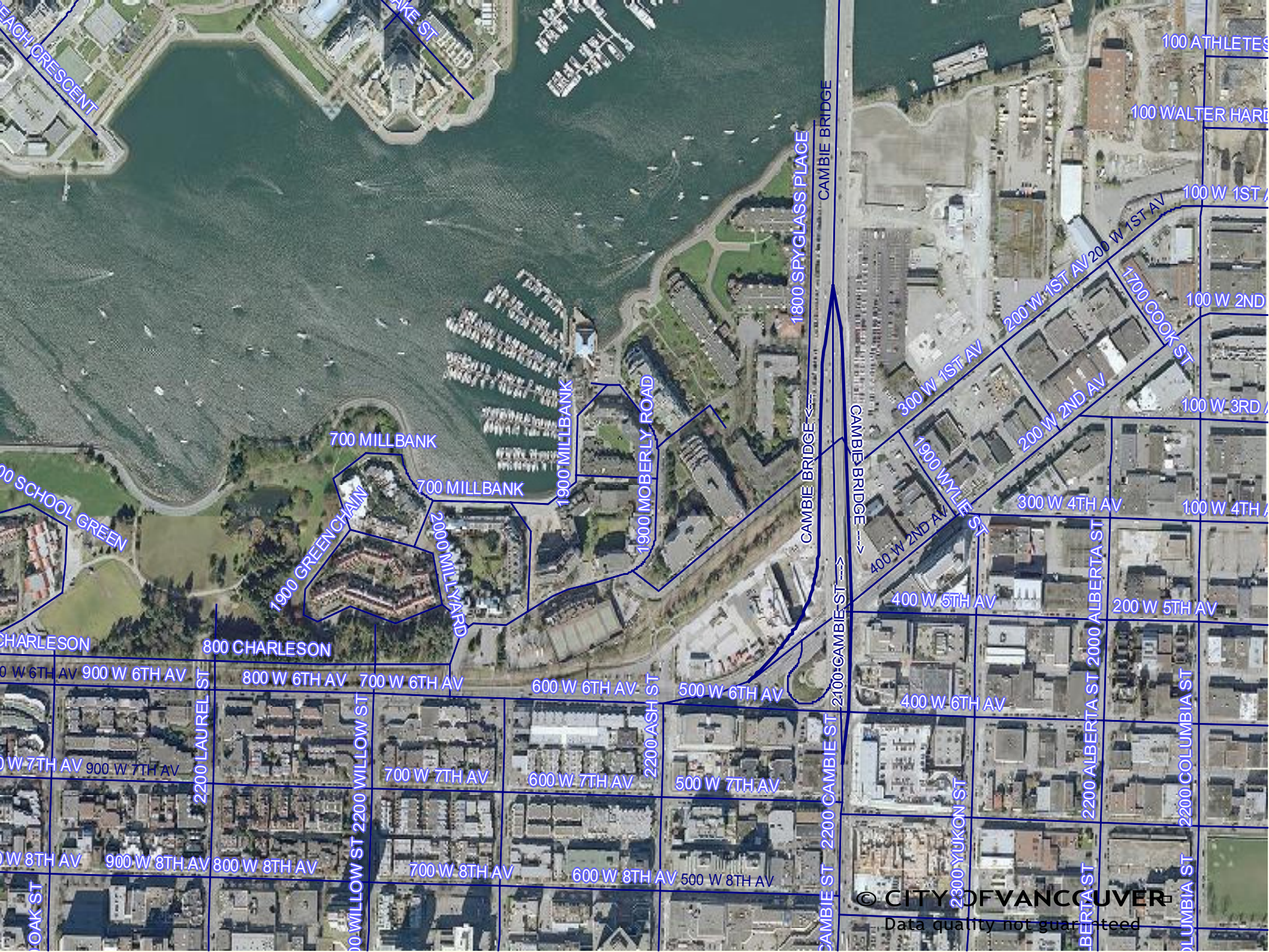
14.4 x 10.0 (mi)

 Internet



© CITY OF VANCOUVER

Data quality not guaranteed





601

609

619

633

456

525

1859

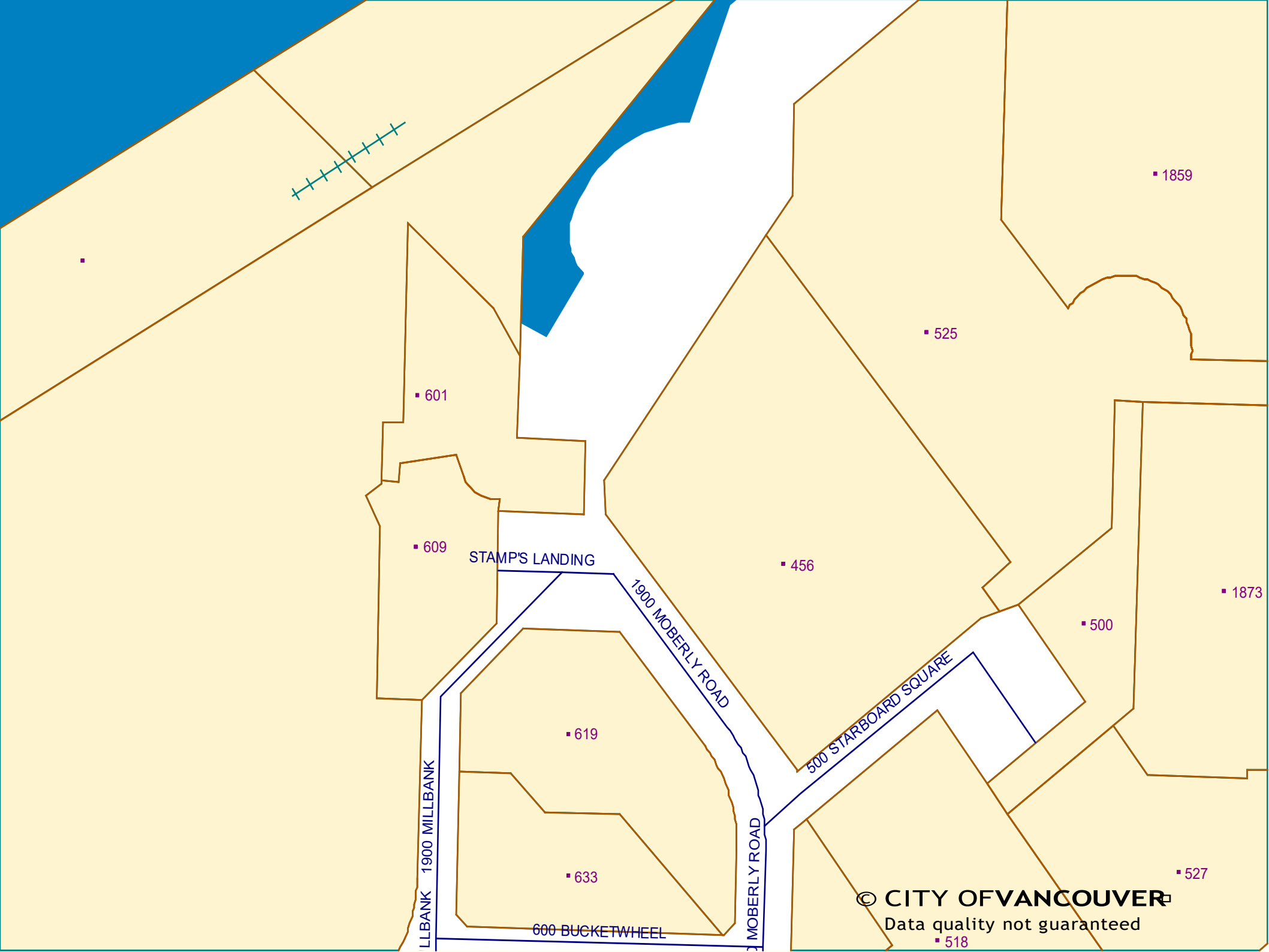
500

1873

527

518

© CITY OF VANCOUVER
Data quality not guaranteed



▪ 601

▪ 609

▪ 619

▪ 633

▪ 456

▪ 525

▪ 1859

▪ 500

▪ 1873

▪ 527

▪ 518

STAMPS LANDING

600 BUCKETWHEEL

1900 MILLBANK

1900 MOBERLY ROAD

MOBERLY ROAD

500 STARBOARD SQUARE

© CITY OF VANCOUVER

Data quality not guaranteed

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 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



Address Search

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

Number

Street

or select

Go

Legend Options

Toolbox

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



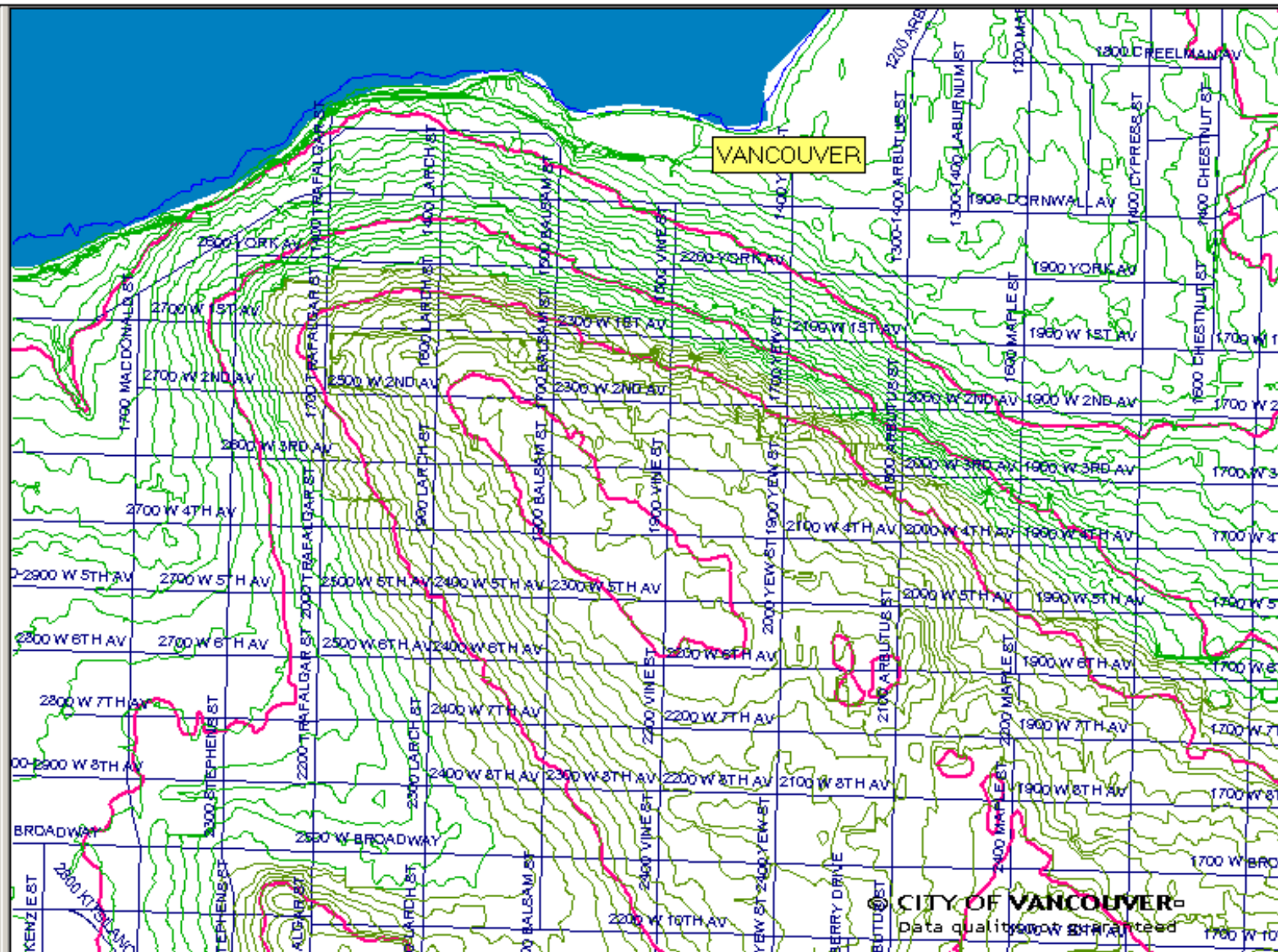
Help

[About the Data](#)

Applications

Select Application:

- ☐ Sea Level to 25m
- ☐ 26m - 51m
- ☐ 51m - 75m
- ☐ 76m - 99m
- ☐ 100m or higher
- ☒ 1-metre Contours (2002-NE)
 - ☐ -10m to Sea Level
 - ☐ Sea Level to 25m
 - ☐ 26m - 51m
 - ☐ 51m - 75m
 - ☐ 76m - 99m
 - ☐ 100m or higher
- ☒ 1-metre Contours (2002-SW)
 - ☐ -10m to Sea Level
 - ☐ Sea Level to 25m
 - ☐ 26m - 51m
 - ☐ 51m - 75m
 - ☐ 76m - 99m
 - ☐ 100m or higher
- ☒ 1-metre Contours (2002-SE)
 - ☐ -10m to Sea Level
 - ☐ Sea Level to 25m
 - ☐ 26m - 51m
 - ☐ 51m - 75m
 - ☐ 76m - 99m
 - ☐ 100m or higher
- ☐ 2-Metre Contour Lines
- ☒ 10-Metre Contour Lines
- ☒ Street Lighting



BC Regions : VANCOUVER

0 feature(s) selected

1 : 9,439

6,145 x 4,572 (ft)

Done

Internet

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 than 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 Antarctica** 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.



... if it were not possible to stabilize content and fix form...

- We could trade stability of content and fixity of form with the **ability to track changes**.
Record=last manifestation + log of changes + metadata
- We could think of the record as existing in two modes: **record in becoming** when the object is accessed to add information to it; **record** when the object is accessed for use.
Record=each manifestation accessed for use + metadata



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/diplomatic and technological, but also legal, ethical, financial
- The archivist must be all to all records, including the **potential records that s/he should contribute to design.**
- So... when, what, by whom, for whom, where, how often?



Conclusions (cont.)

When? ASAP after creation

What? Functions, records, systems

By whom? A team of records professionals, legal and financial experts and IT professionals

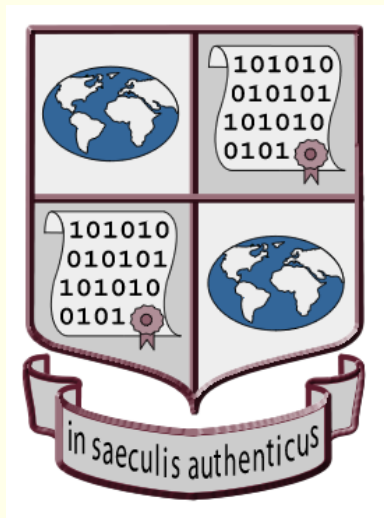
For whom? Creator, researchers, society at large

Where? In the environment in which they reside at any given time

How often? Every time the creator's systems are upgraded or changed. Once the records are acquired, the selection process ends and the appraisal is definitive.



InterPARES Web Site



www.interpares.org



InterPARES Project

Luciana Duranti
Project Director