InterPARES 3 Project

International Research on Permanent Authentic Records in Electronic Systems

TEAM Canada

Implementing Digital Records Preservation in Small and Medium-sized Archives

Luciana Duranti

InterPARES Project Director

Mexico City

30 September 2008

The Goal of InterPARES 1 & 2

To develop the body of theory and methods necessary to ensure that digital records produced in databases and office systems as well as 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 them and of society at large, regardless of technology obsolescence and media fragility.

The Digital Records Challenge

- They do not exist as physical entities, but are constituted of **linked digital components** (the "manifested" record differs from the "stored" record, if there is one—we might have only form data, content data, and composition data)
- Their **original** manifestation disappears when they are saved: **we cannot maintain or preserve digital records**, but only the ability to re-produce or re-create them, or to protect their capacity to instruct or enable the making of records
- The facility of reproduction and manipulation makes it **difficult to identify the final, official, reliable or accurate version**
- **Technological obsolescence** makes digital records inaccessible in a very short time span
- **Intellectual property and privacy rights** are hard to protect

The Digital Records Challenge (cont.)

The systems that contain records, contain bad records, primarily because of **lack of identifiable contexts and relationships** among themselves and with records outside the system

Most systems that should contain records do not, because the entities in them **lack fixed form and stable content**.

- In interactive systems, each user intervention or input from another system causes a change of content and/or form (Alsace-Moselle).
- In dynamic systems they depend for their content upon data extracted from a variety of other systems which may have variable instantiations (VanMap).

These digital entities, regularly produced by government agencies, only exist in the hands of the government. They are either **trusted implicitly**, when they should not be, **or not trusted**, when they are trustworthy (VPD)

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 related 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 data 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 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 a record, although is a record in becoming.
- 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 a relational database sufficiently simple to be maintained by the designated preserver, the archives (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

http://vanmappub.vancouver.ca - VanMap Public ActiveX Viewer version for Windows - Microsoft Internet Explorer provided by ICBC

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

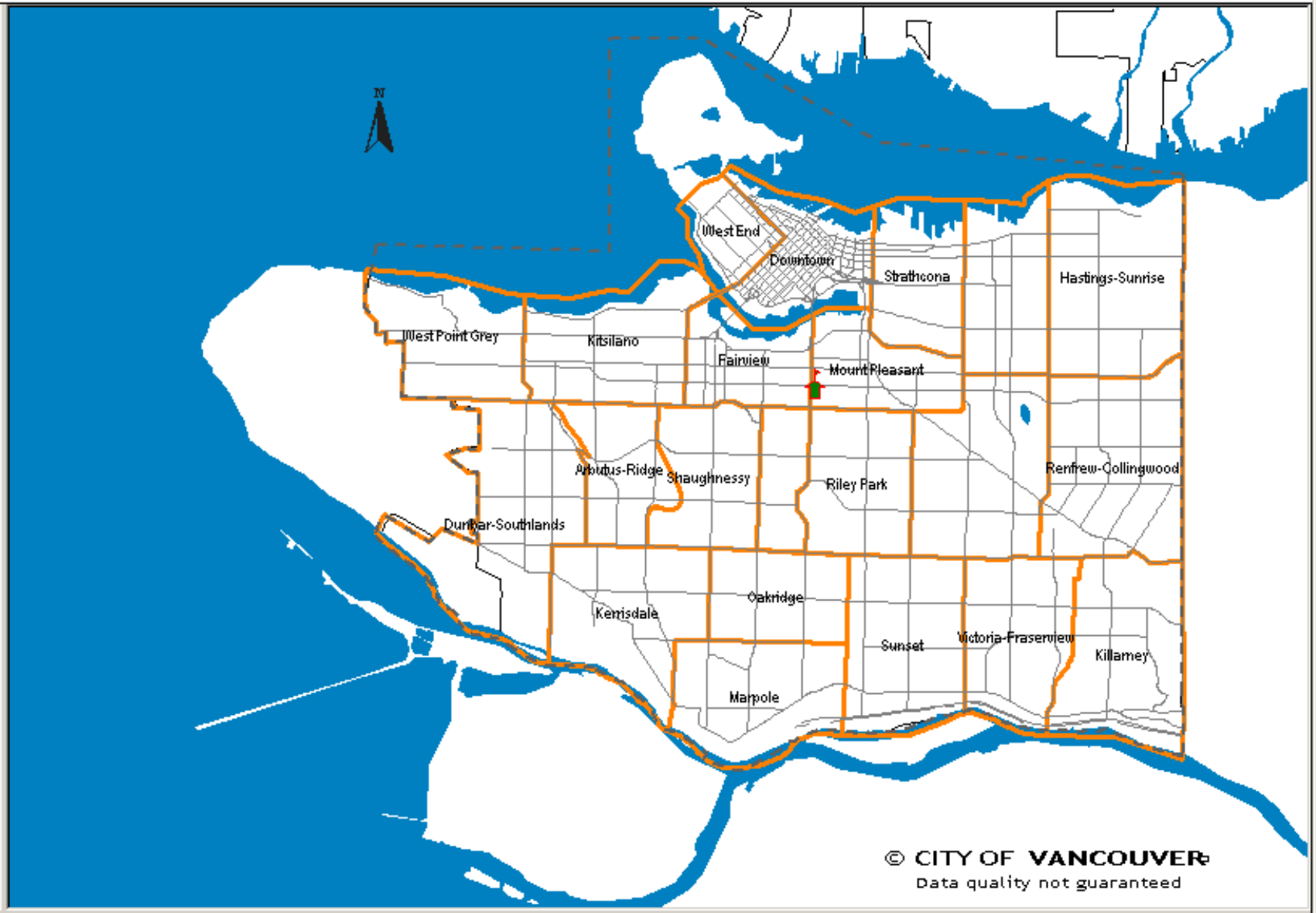
Help

About the Data

Applications

Select Application:

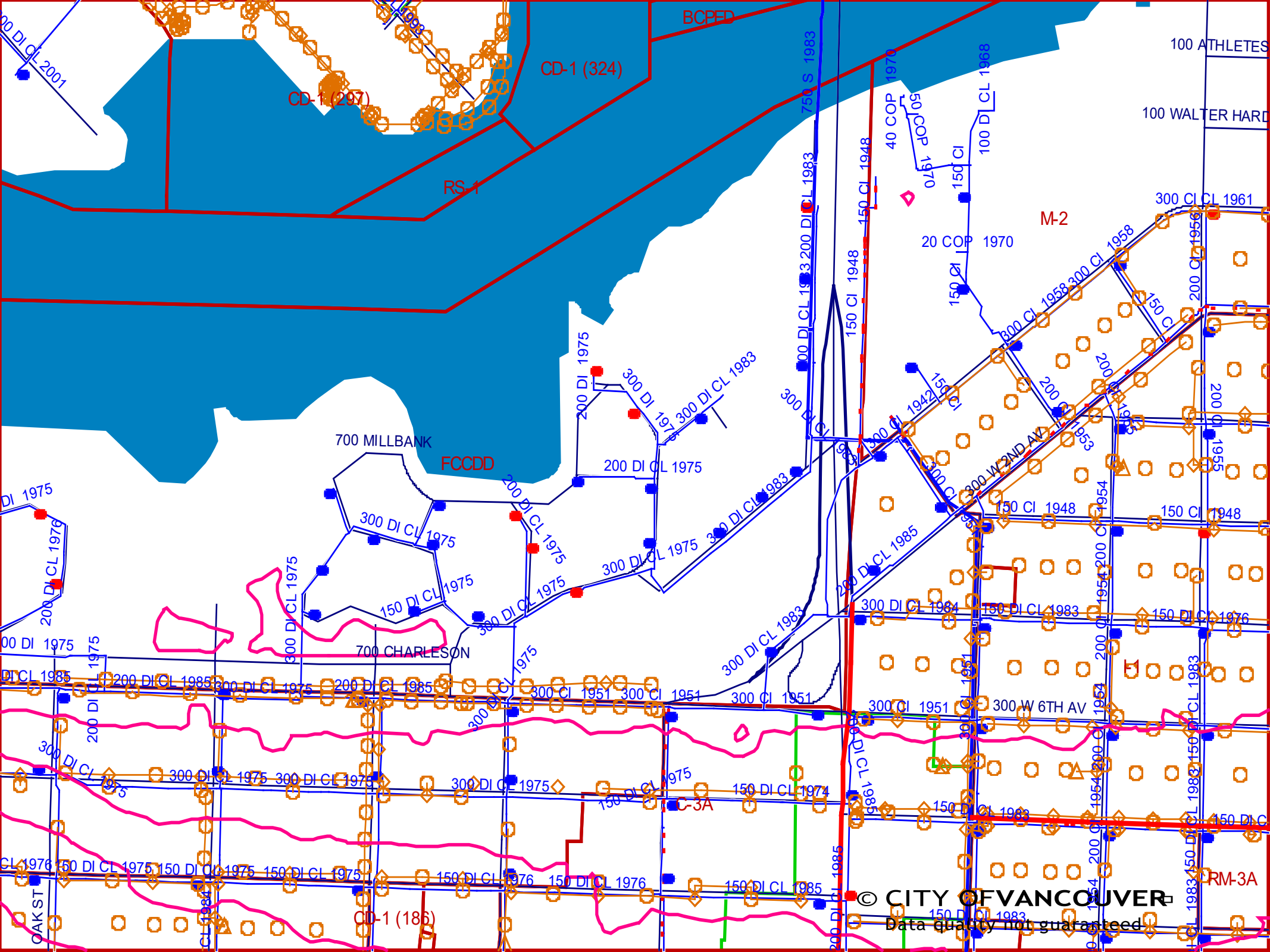
- [ ] Olympic Venues
- [x] City Boundary
- [ ] Public Art
- [x] Public Places
- [ ] The Road Ahead
- [ ] City Projects
- [x] Child Care Facilities
- [ ] Web Cameras
- [ ] Non-Market Housing
- [ ] City-owned Properties
- [x] Cemetery
- [ ] Urban Agriculture
- [ ] Contour Lines
- [ ] Truck Routes
- [ ]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

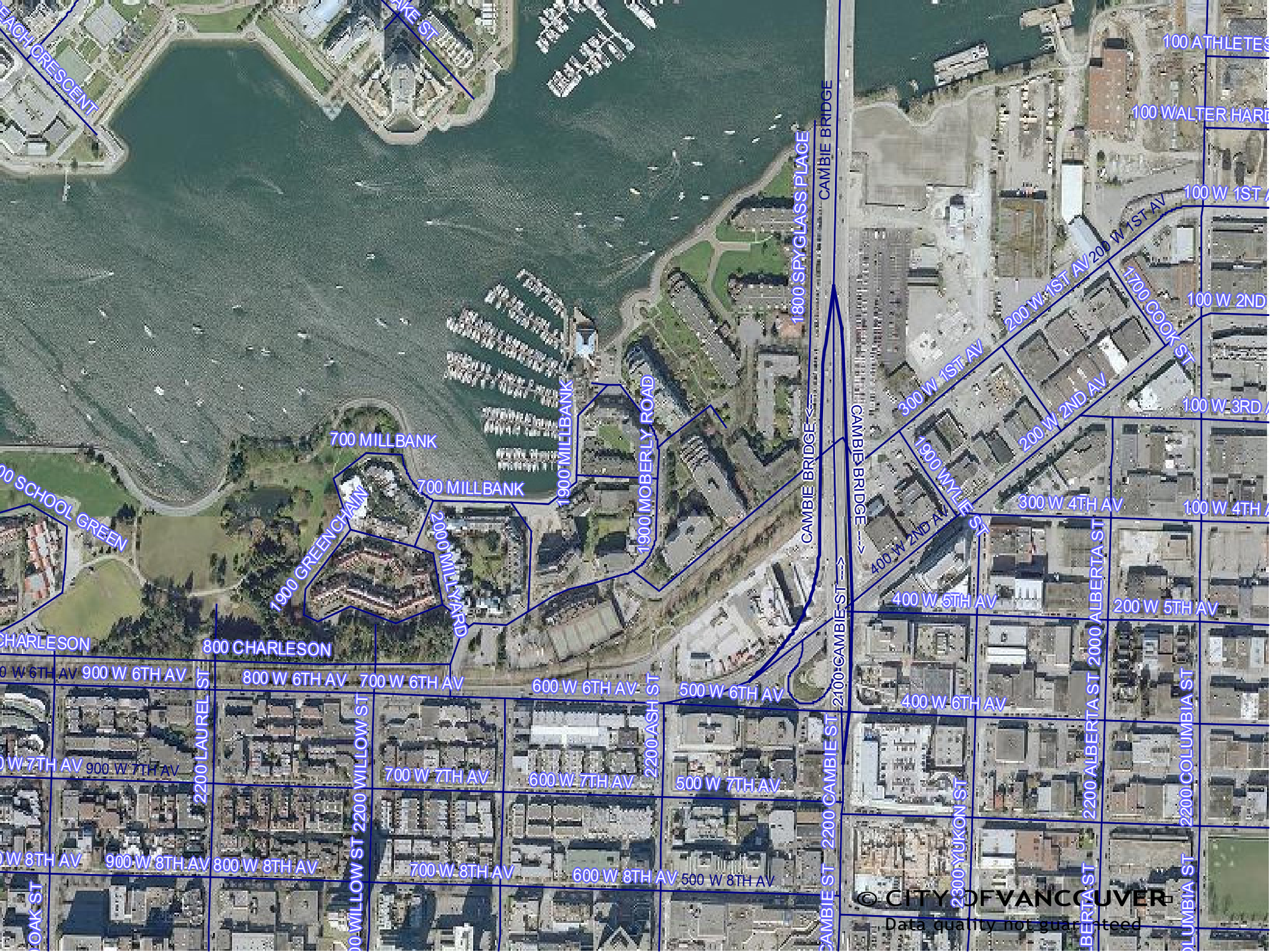


Water Bodies : Strait of Georgia | 0 feature(s) selected | 1 : 109,043 | 14.4 x 10.0 (mi)

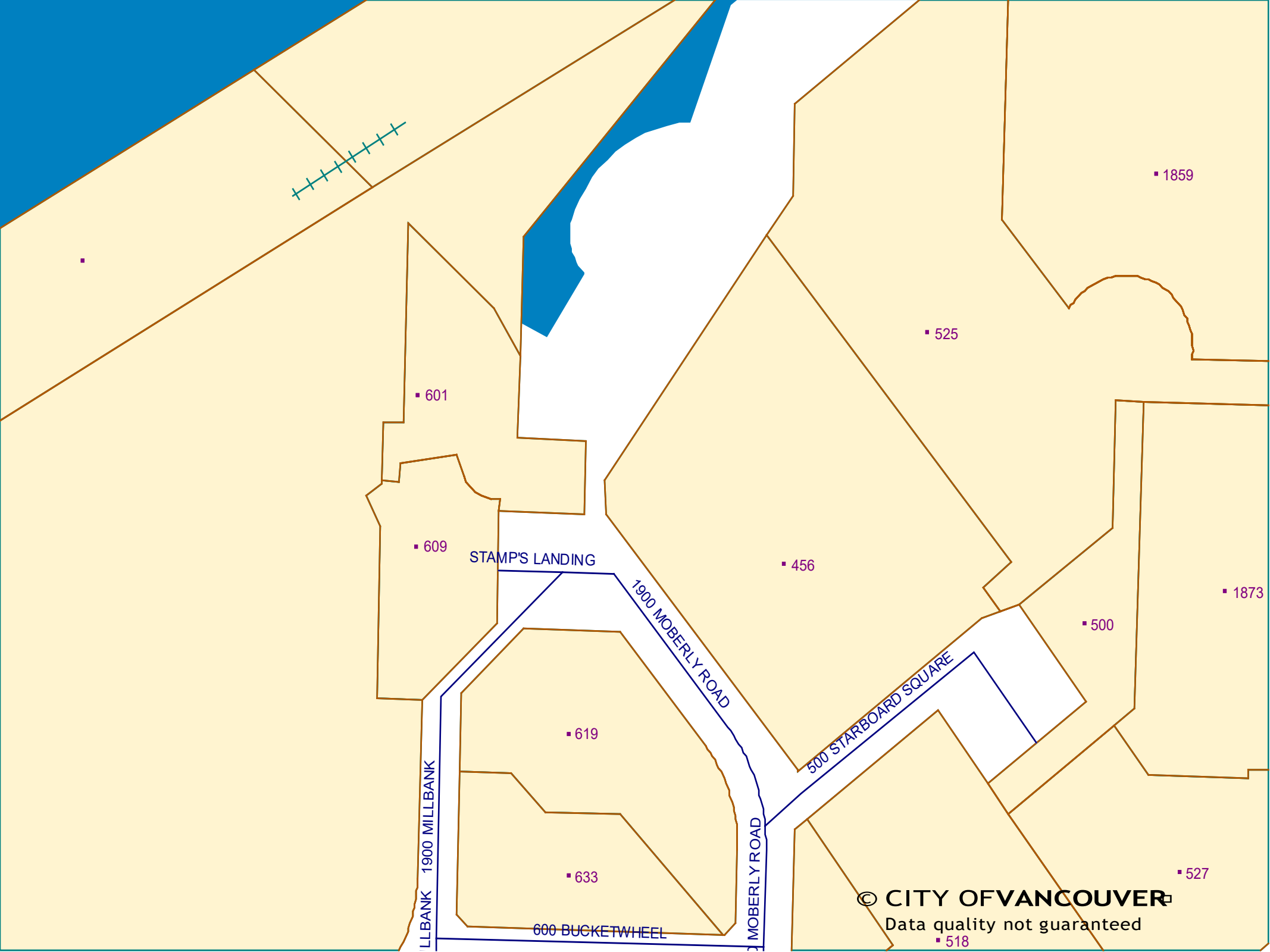
Internet

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


BCPED
CD-1 (324)
CD-1 (297)
RS-1
M-2
FCCDD
700 MILLBANK
700 CHARLESON
100 ATHLETES
100 WALTER HARD
300 W 6TH AV
300 W 2ND AV
C-3A
RM-3A
CD-1 (186)
OAK ST
© CITY OF VANCOUVER
Data quality not guaranteed


100 ATHLETES
100 WALTER HARD
100 W 1ST A
100 W 2ND
100 W 3RD A
100 W 4TH A
1800 SPYGLASS PLACE
CAMBIE BRIDGE
CAMBIE BRIDGE <---
CAMBIE BRIDGE --->
200 W 1ST AV
200 W 1ST AV
1700 COOK ST
300 W 1ST AV
200 W 2ND AV
1900 WYLER ST
300 W 4TH AV
400 W 2ND AV
400 W 5TH AV
200 W 5TH AV
2000 ALBERTA ST
2200 COLUMBIA ST
2300 YUKON ST
2100 CAMBIE ST
2200 CAMBIE ST
700 MILLBANK
700 MILLBANK
1900 MILLBANK
1900 MOBERLY ROAD
2200 MILLYARD
1900 GREENCHAW
00 SCHOOL GREEN
CHARLESON
800 CHARLESON
0 W 6TH AV
900 W 6TH AV
800 W 6TH AV
700 W 6TH AV
600 W 6TH AV
500 W 6TH AV
400 W 6TH AV
0 W 7TH AV
900 W 7TH AV
700 W 7TH AV
600 W 7TH AV
500 W 7TH AV
0 W 8TH AV
900 W 8TH AV
800 W 8TH AV
700 W 8TH AV
600 W 8TH AV
500 W 8TH AV
2200 LAUREL ST
2200 WILLOW ST
2200 ASH ST
OAK ST
E BEACH CRESCENT
E ST
© CITY OF VANCOUVER
Data quality not guaranteed


1859
525
601
609
456
1873
500
619
633
527
518
© CITY OF VANCOUVER
Data quality not guaranteed


1859
525
601
609
STAMP'S LANDING
456
1900 MOBERLY ROAD
1873
500
500 STARBOARD SQUARE
619
1900 MILLBANK
633
600 BUCKETWHEEL
MOBERLY ROAD
527
518
© CITY OFVANCOUVER
Data quality not guaranteed

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	200
	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	200

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

- Contains data that often do not exist anywhere else, especially in the correlated form showed on the GIS layers, and 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 take the views 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

http://vanmappub.vancouver.ca - VanMap Public ActiveX Viewer version for Windows - Microsoft Internet Explorer provided by ICBC

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

Help

About the Data

Applications

Select Application:

- Sea Level to 25m
- 26m - 51m
- 51m - 75m
- 76m - 99m
- 100m or higher

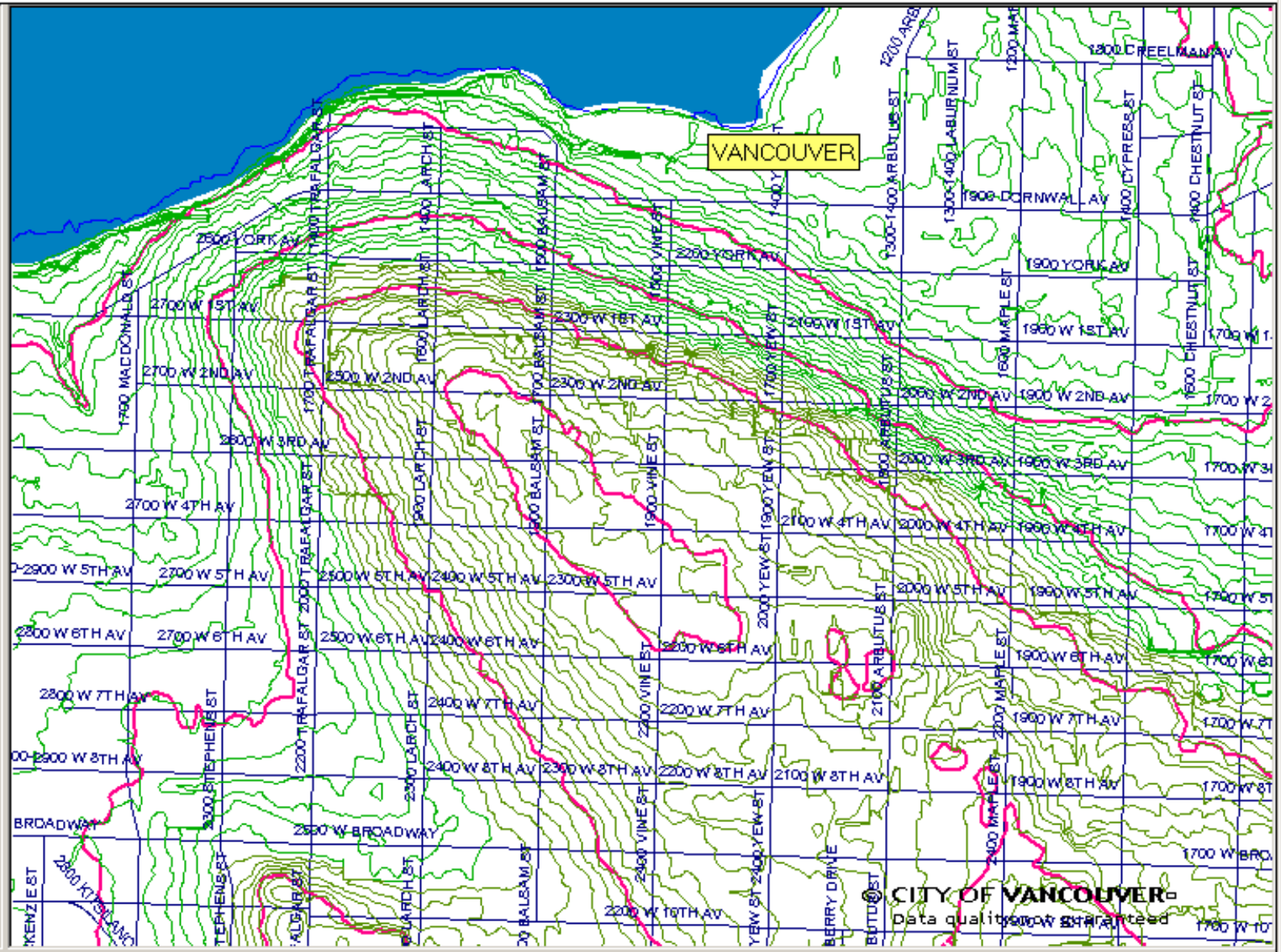
☑ 1-metre Contours (2002-NE)
- -10m to Sea Level
- Sea Level to 25m
- 26m - 51m
- 51m - 75m
- 76m - 99m
- 100m or higher

☑ 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

Start | Inbox - Microsoft Outlook | VanMap - City of Vanc... | http://vanmappub.... | VanMap | Microsoft PowerPoint - ... | 1:09 PM

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; therefore it requires 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.

InterPARES Principles

- Technology cannot determine the solution to the reliable and accurate creation of digital records or to their authentic preservation over the long term: **organizational needs** define the problem and **archival principles** must establish the correctness and adequacy of each technical solution
- Solutions to the digital records challenges are inherently **dynamic** and **specific** to the cultural, disciplinary, administrative and legal situations
- Preservation is a **continuous process that begins with records creation**
- We must be able to **presume records reliability**, till proof to the contrary is established
- We must be able to **infer authenticity** on the basis of the circumstances of records creation, maintenance and preservation

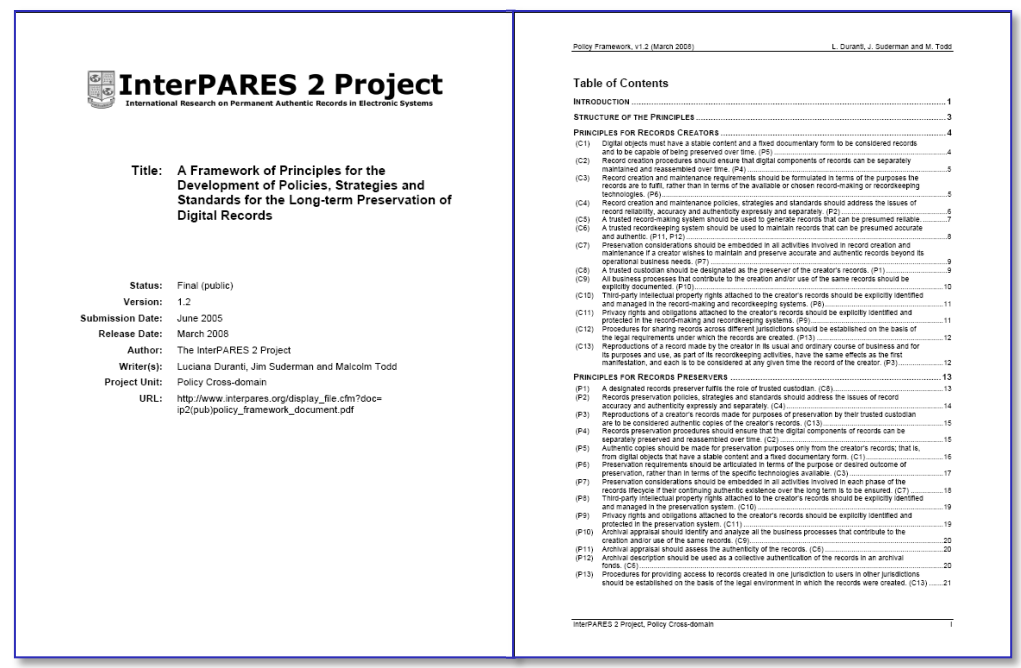
InterPARES Research Activities

- Identification of **what constitutes a record** in each type of system and in each context, and **what record has the force of an original**, based on archival and diplomatics theory and six thousand years of record making and recordkeeping
- Definition of **what a reliable, accurate and authentic record** is in the arts, science, law and administration on the basis of analysis of literature, surveys, interviews
- Development of the requirements for the design of **a trusted record making system, a trusted recordkeeping system** and **a trusted record preservation system**, on the basis of 42 case studies, and the modeling and diplomatic analysis of the results
- Development of **methods and procedures for the creation, maintenance, appraisal, selection and disposition, and long-term preservation of digital records**, based on archival theory, law, and issues related to organizational culture

Key IP 1 & 2 Final Products

- **Policy Framework**

  A framework of principles guiding the development of policies for records creating and preserving organizations



InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

Title: A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records

Status: Final (public)
Version: 1.2
Submission Date: June 2005
Release Date: March 2008
Author: The InterPARES 2 Project
Writer(s): Luciana Duranti, Jim Suderman and Malcolm Todd
Project Unit: Policy Cross-domain
URL: http://www.interpares.org/display_file.cfm?doc=ip2(pub)policy_framework_document.pdf

Policy Framework, v1.2 (March 2008) L. Duranti, J. Suderman and M. Todd

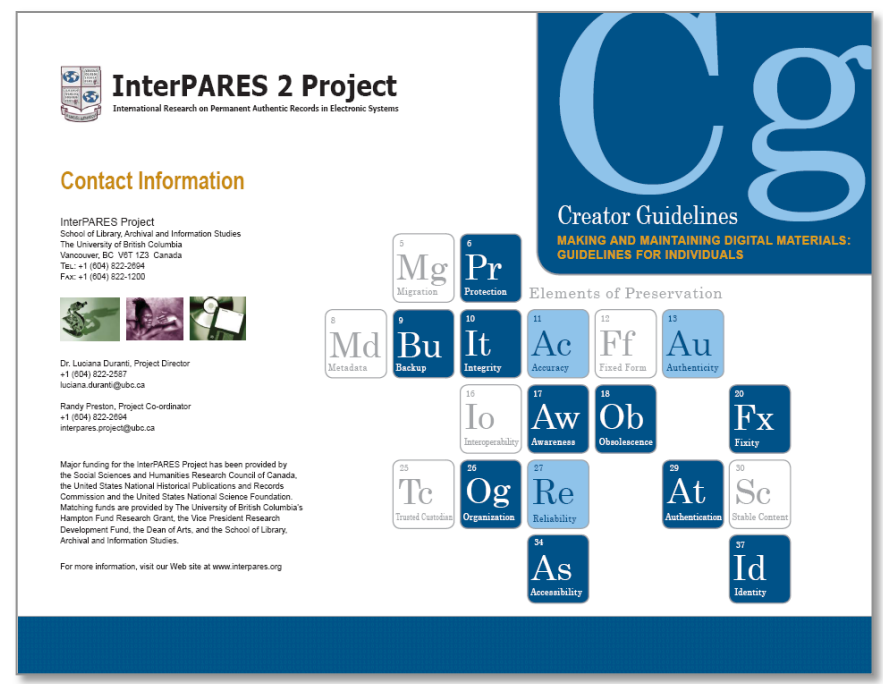
Table of Contents

InterPARES 2 Project, Policy Cross-domain i

IP 1 & 2 Final Products

- **Creator Guidelines**

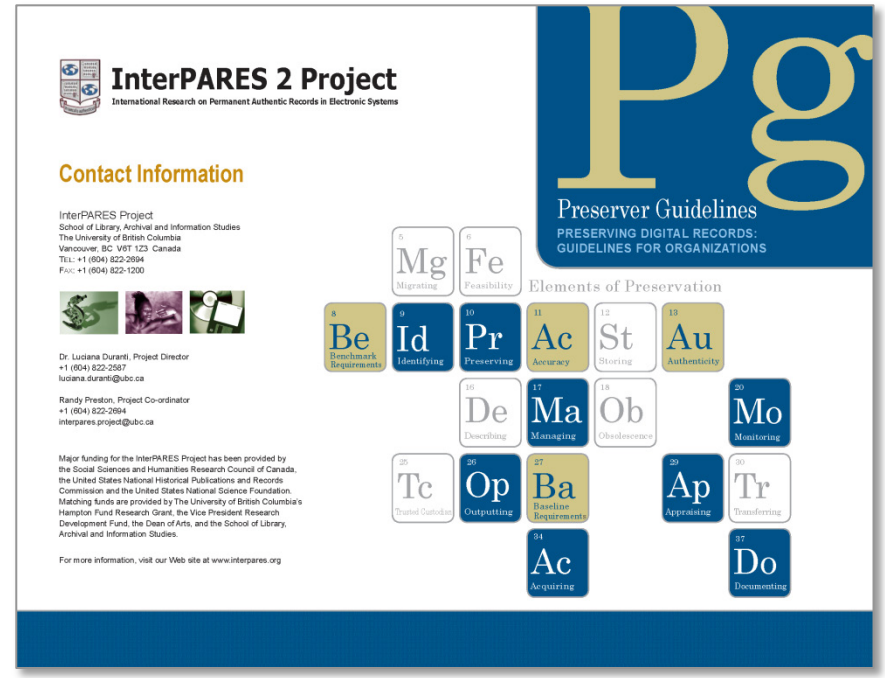
  Recommendations for making and maintaining digital materials for individuals and small communities of practice



IP 1 & 2 Final Products

- **Preserver Guidelines**

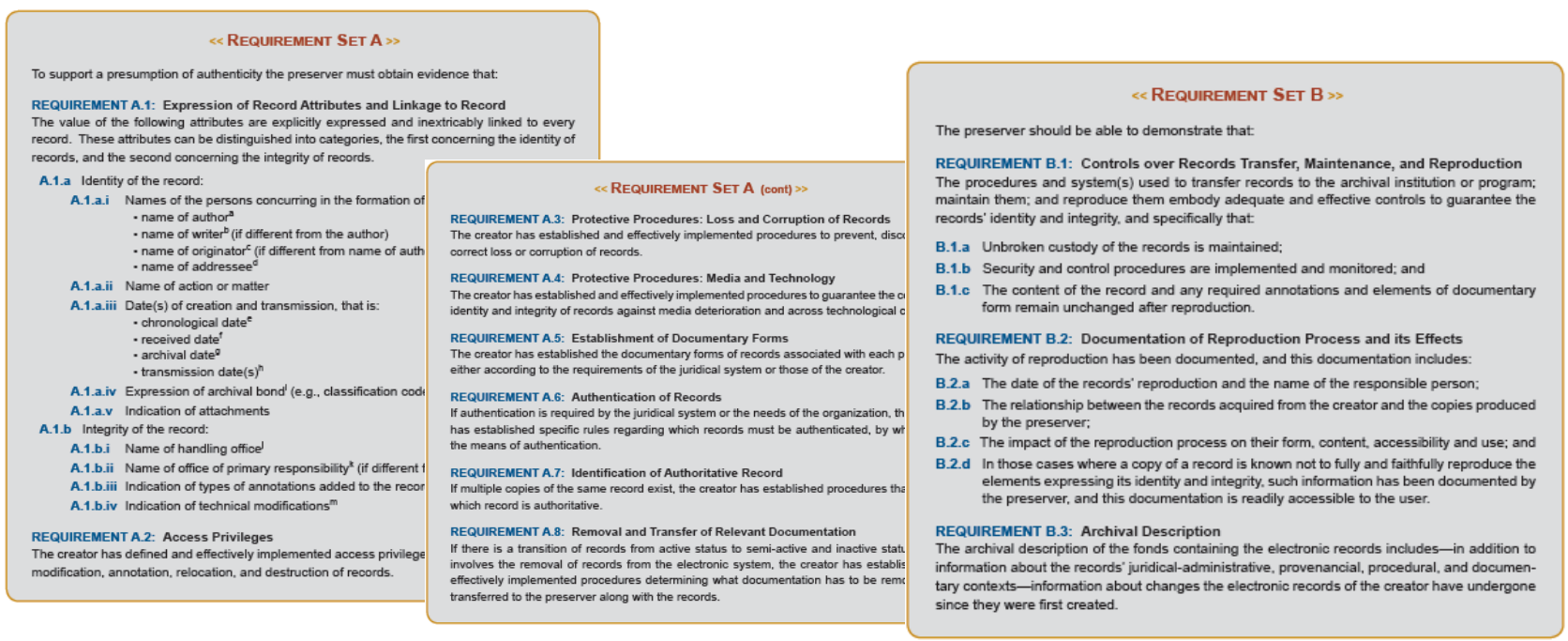
  Recommendations for digital preservation for archival institutions



IP 1 & 2 Final Products

- **Benchmark and Baseline Requirements**

  Authenticity requirements for assessing and maintaining the authenticity of digital records



<< REQUIREMENT SET A >>

To support a presumption of authenticity the preserver must obtain evidence that:

REQUIREMENT A.1: Expression of Record Attributes and Linkage to Record
The value of the following attributes are explicitly expressed and inextricably linked to every record. These attributes can be distinguished into categories, the first concerning the identity of records, and the second concerning the integrity of records.

A.1.a Identity of the record:
- A.1.a.i Names of the persons concurring in the formation of
  - name of author[a]
  - name of writer[b] (if different from the author)
  - name of originator[c] (if different from name of auth
  - name of addressee[d]
- A.1.a.ii Name of action or matter
- A.1.a.iii Date(s) of creation and transmission, that is:
  - chronological date[e]
  - received date[f]
  - archival date[g]
  - transmission date(s)[h]
- A.1.a.iv Expression of archival bond[i] (e.g., classification code
- A.1.a.v Indication of attachments

A.1.b Integrity of the record:
- A.1.b.i Name of handling office[j]
- A.1.b.ii Name of office of primary responsibility[k] (if different
- A.1.b.iii Indication of types of annotations added to the recor
- A.1.b.iv Indication of technical modifications[m]

REQUIREMENT A.2: Access Privileges
The creator has defined and effectively implemented access privilege
modification, annotation, relocation, and destruction of records.

<< REQUIREMENT SET A (cont) >>

REQUIREMENT A.3: Protective Procedures: Loss and Corruption of Records
The creator has established and effectively implemented procedures to prevent, disc
correct loss or corruption of records.

REQUIREMENT A.4: Protective Procedures: Media and Technology
The creator has established and effectively implemented procedures to guarantee the c
identity and integrity of records against media deterioration and across technological c

REQUIREMENT A.5: Establishment of Documentary Forms
The creator has established the documentary forms of records associated with each p
either according to the requirements of the juridical system or those of the creator.

REQUIREMENT A.6: Authentication of Records
If authentication is required by the juridical system or the needs of the organization, th
has established specific rules regarding which records must be authenticated, by wh
the means of authentication.

REQUIREMENT A.7: Identification of Authoritative Record
If multiple copies of the same record exist, the creator has established procedures tha
which record is authoritative.

REQUIREMENT A.8: Removal and Transfer of Relevant Documentation
If there is a transition of records from active status to semi-active and inactive statu
involves the removal of records from the electronic system, the creator has establis
effectively implemented procedures determining what documentation has to be rem
transferred to the preserver along with the records.

<< REQUIREMENT SET B >>

The preserver should be able to demonstrate that:

REQUIREMENT B.1: Controls over Records Transfer, Maintenance, and Reproduction
The procedures and system(s) used to transfer records to the archival institution or program; maintain them; and reproduce them embody adequate and effective controls to guarantee the records' identity and integrity, and specifically that:

- B.1.a Unbroken custody of the records is maintained;
- B.1.b Security and control procedures are implemented and monitored; and
- B.1.c The content of the record and any required annotations and elements of documentary form remain unchanged after reproduction.

REQUIREMENT B.2: Documentation of Reproduction Process and its Effects
The activity of reproduction has been documented, and this documentation includes:

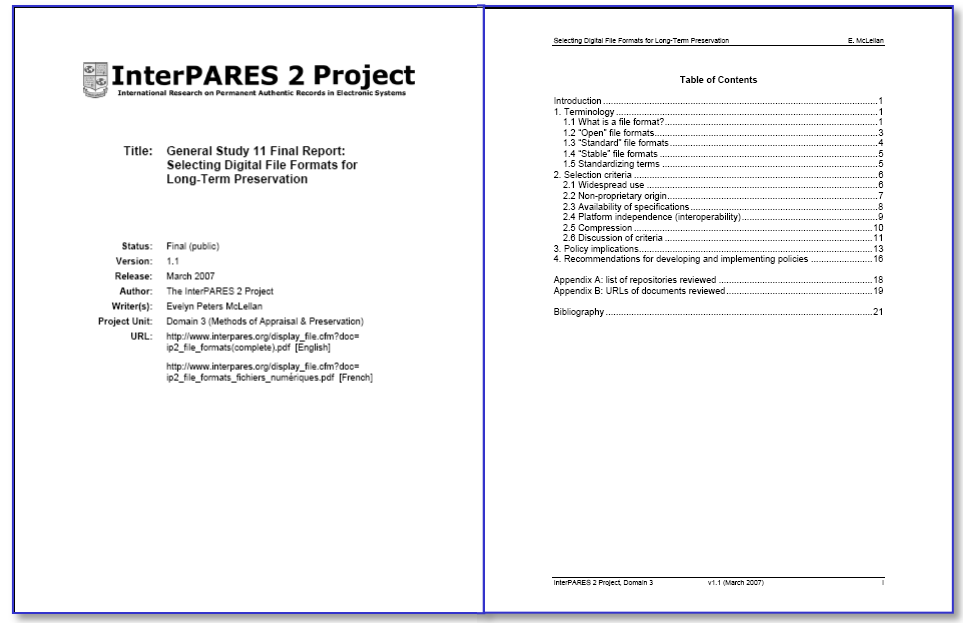
- B.2.a The date of the records' reproduction and the name of the responsible person;
- B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver;
- B.2.c The impact of the reproduction process on their form, content, accessibility and use; and
- B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user.

REQUIREMENT B.3: Archival Description
The archival description of the fonds containing the electronic records includes—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the electronic records of the creator have undergone since they were first created.

IP 1 & 2 Final Products

- **File Format Selection Guidelines**

  Principles and criteria for adoption of file formats, wrappers and encoding schemes



InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

Title: General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation

Status: Final (public)
Version: 1.1
Release: March 2007
Author: The InterPARES 2 Project
Writer(s): Evelyn Peters McLellan
Project Unit: Domain 3 (Methods of Appraisal & Preservation)
URL: http://www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf [English]
http://www.interpares.org/display_file.cfm?doc=ip2_file_formats_fichiers_numériques.pdf [French]

Selecting Digital File Formats for Long-Term Preservation E. McLellan

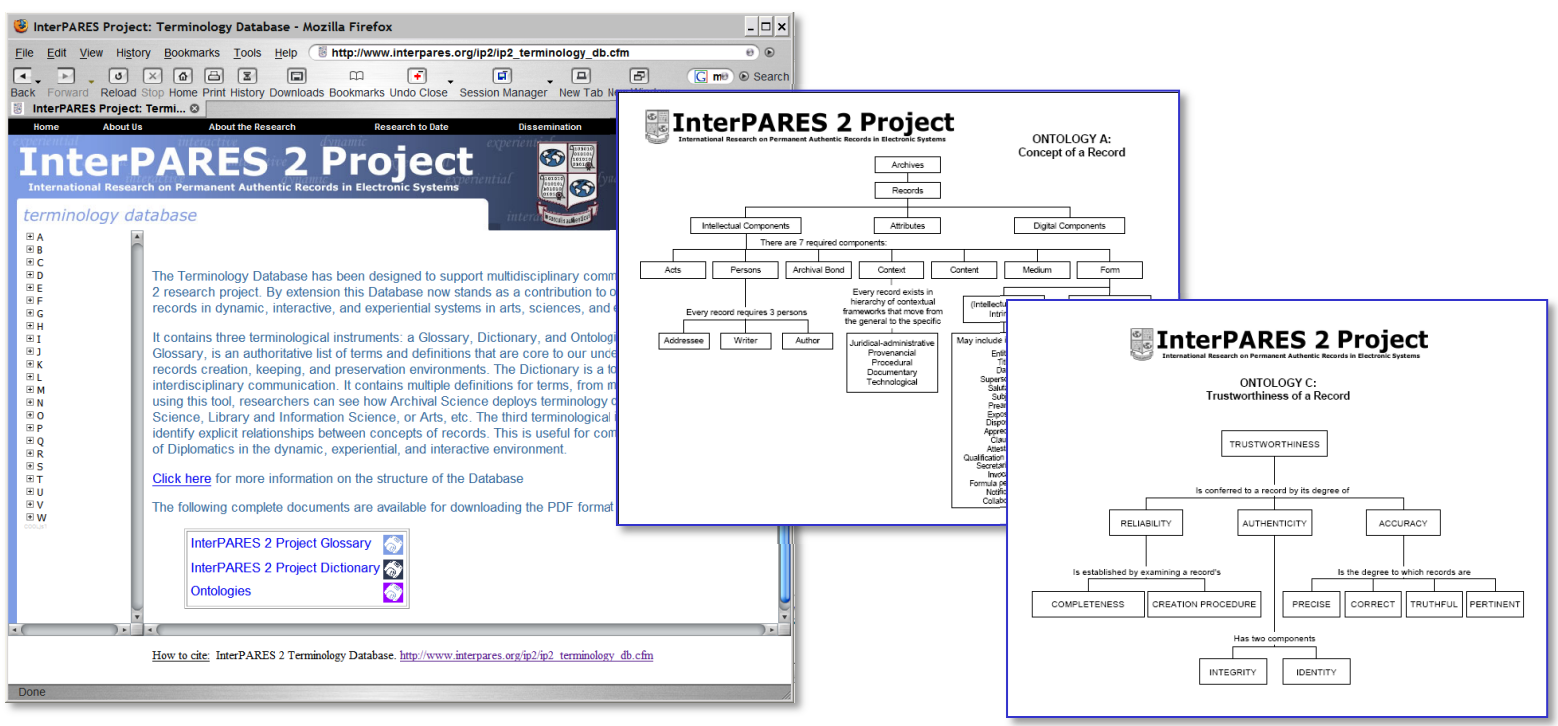
Table of Contents

InterPARES 2 Project, Domain 3 v1.1 (March 2007) i

IP 1 & 2 Final Products

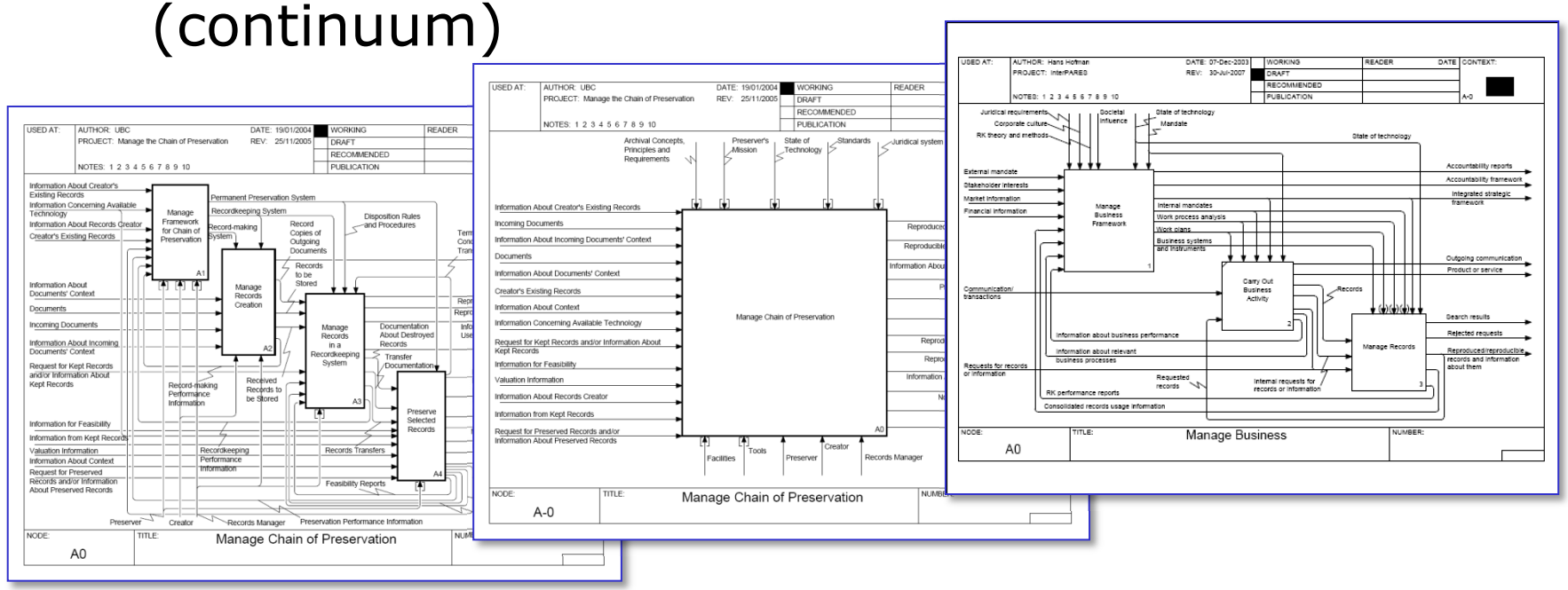
- **Terminology Database**

  Including a glossary, a dictionary and ontologies



IP 1 & 2 Final Products

- **Two Records Management Models**
  1. Chain of Preservation (COP) Model (lifecycle)
  2. Business-driven Recordkeeping (BDR) Model (continuum)



InterPARES Recommendation: Archivist as Trusted Custodian

The trusted custodian is a person who

- acts as a **neutral third party**, i.e., demonstrates that he/she has no stake in the content of the records and no reason to alter records under his/her custody, and that he/she will not allow anybody to alter the records either accidentally or on purpose,
- is equipped with the **knowledge and skills** necessary to fulfil its responsibilities, which should be acquired through formal education, and
- establishes a **trusted preservation system** that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved;
- But, mostly...

The Archivist's New Functions

1. Positions him/herself at the **beginning of the record life-cycle**, taking the role of "designated" trusted custodian
2. Assesses the **authenticity of the records** and **monitors it** throughout their existence
3. Identifies the records to be preserved at the moment of their creation and **monitors their transformation through time**

The Archivist's New Functions (cont.)

4. Determines the **feasibility of preservation** on the basis of the archives technological capacity
5. Determines a **preservation strategy** independently of technological trends (tries to influence the industry through the adoption of standards, but not viceversa) and maintaining the focus on interoperability
6. Controls the **accuracy of the records** after each conversion or migration
7. Develops **procedures** that address issues of **intellectual rights and privacy**

The Archivist's New Functions (cont.)

8. Recognizes to **archival description 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 1) their administrative, custodial and technological history, 2) the illustration of their scope and content, and 3) 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 whole in which the individuality of each member is subject to the bond of a common provenance and destination

The Archivist's New Functions (cont.)

9. Is constantly **involved in research and development projects** similar to those carried out by the industry, addressing questions like the following:
   - What entity constitutes the record in each dynamic or interactive system
   - What instance of such entity can be regarded as the record (manifested or stored entity)
   - How to keep such entities accurate and authentic through time
   - How to enable users to verify such authenticity over time

Why a Third Phase?

- A study of the effectiveness of workshop and seminar experiences for increasing archivists' skills in digital preservation and their ability to implement these skills in their repositories has shown that **very few participants were able to implement the skills** once they returned to their work environments
  - Wendy M. Duff, M., Amy Marshall, Carrie Limkilde and Marlene van Ballegooie (2006) "Digital Preservation Education: Educating or Networking?" *The American Archivist* 69(1): 188-212. In the context of ERPANET.
- Feedback on the outcomes of the two phases of InterPARES from archivists working in institutions smaller than national archives has consistently shown **concern about** their **downward-scalability and** their **relevance** to small and medium sized organizations

Goal of InterPARES 3

To **enable small and medium-sized** public and private **archival organizations and programs** (units within records creating organizations)—which are responsible for the digital records resulting from government, business, research, art and entertainment, social and/or community activities—**to preserve** over the long term **authentic records** that satisfy the requirements of their stakeholders and society's needs for an adequate record of its past.

3 Primary Components

1. **Research component**
   (short-term and long-term projects, including general studies and case studies related to policy, records or systems)
2. **Education and training component**
   (in the context of research projects, apprenticeships, activities credited as part of coursework, etc.)
3. **Knowledge-mobilization component**
   (workshops, seminars, colloquia, policy manuals and other publications, public lectures, etc., that meet the needs of both academic and community partners)

InterPARES 3 Composition

International Alliance

14 regional, national & multinational TEAMs:

TEAM (Theoretical Elaboration into Archival Management) Canada (including US); Africa; Brazil; Catalonia; China; Italy; Korea; Malaysia; Mexico; Netherlands & Belgium; Norway; Singapore; Turkey and UK & Ireland

Director: Luciana Duranti
Headquarters: UBC - SLAIS (facilities provided by UBC)
Summits: Once a year, each time hosted by a different country
Symposia: Once a year, each time hosted by a different country

TEAM Composition

- **Director**, principal investigator
- **Test-bed partners** (primary stakeholders)
- **Co-applicants** (individual academic researchers and professional practitioners)
- **Collaborators** (individual experts)
- **Resource partners** (e.g., ACA, NARA, LAC, CCI)
- **Graduate research assistants**

Expected Products

1. **Policies, strategies and procedures** for small archival organizations or programs, and **guidelines** for the records creators whose records fall under their responsibility.
2. **Action plans** for the specific case studies carried out in the course of the Project.
3. **Criteria to determine "most-at-risk" materials** e.g., checklist of age (date created, date last accessed), physical carrier, operating system, software used, equipment required and its availability, etc.

Expected Products (cont.)

4. **Guidelines for addressing digital preservation requirements** that apply to specific types of records, but not to other materials.

5. **Evaluation models for assessing the degree of success,** if any, of the chosen preservation action.

6. **Cost-benefit models** for various types of archival organizations or programs and for various kinds of records and/or systems.

7. **Ethical models** that identify and make explicit the consequences for individuals and society of types of preservation measures or lack thereof.

Expected Products (cont.)

8. **Training and education modules** for preservers, professional associations and university programs; and **awareness and education modules for non-archivists**, such as IT professionals, vendors and service providers; human resources and financial managers; doctors, communities of practice, members of the general public, etc.; and **a strategy for delivering them**.
9. **Position papers** directed to key regulating, auditing and policy-making bodies, advocating the vital need of embedding planned digital preservation in the requirements they issue for the activities they regulate, audit or control.

Methodology (general)

Action Research

- Practical, collaborative, pragmatic research directed toward producing solutions that are **directly useful** to a group of people
- Research **subjects are co-participants** and stakeholders in the process
- **Jointly** define research objectives and goals, co-construct research questions, pool knowledge and **develop solutions** and performance tests that implement specific strategies

Methodology (general - cont.)

Action Research

Two distinct methods of research:

1. **Prototype development research**
2. **Ethnographic research**

Methodology (general - cont.)

Prototype development research

- **User-centered, collaborative** prototyping approach that explores the interplay between theory and practice
- **Proof-by-demonstration**
- Comprises three major **iterative** stages:
  1. concept building (which we have done)
  2. system building
  3. system evaluation

Methodology (general - cont.)

Ethnographic research

- Creators of records, their users and archivists form a community of practice— **the archival environment**—for which social interaction creates meaning and defines values
- Researchers place themselves within an archival environment to **gain the cultural perspective** of those responsible for records
- **Observation** of the environment with detailed description, extensive **interviewing and analysis** of the documents

Case Study Methodological Steps

- **Identifying the Problem** -- Initially, each test-bed partner will identify a **body of digital material** for which a preservation plan has to be developed, be it already in the custody of the partner or not. Alternatively, the partner identifies a **policy need**, or a **system to be designed and implemented**. Most test-beds have already done this.
- **Context Data Collection** – Using archival methods, data will be collected about the context and limitations of each test-bed. The instrument intended to support consistency in data collection across all case studies is the "Template for Case Study Contextual Analysis". These data can be collected from analysis of web sites and of documentation identified or provided by the test-bed, and from interviews.

Case Study Methodology (cont.)

- **Case Data Collection—**Subsequently, using documentation, interviews, diplomatics, modelling, and text analyses, data will be collected either about the specific body of material, its documentary forms, technological constraints, functional or cultural meaning, etc., or about a system requirements or policy needs and constraints.

  For diplomatic analysis there is a template. For modelling, we use IDEFØ (Integrated Definition Function Modeling) modeling software. IDEFØ is a U.S. Information Processing Standard, as described in Publication 183 of the National Institute of Standards and Technology. A function model is a structured representation of the functions, activities and processes within the modeled system or subject area. For an introduction to IDEFØ modeling, see "Integrated Definition Function Modeling (IDEFØ): A Primer." For more detailed information, http://www.idef.com.

Case Study Methodology (cont.)

- **Answering the case studies questions—**On the basis of the case data collected, the members of the case study team answer the relevant set of "questions for researchers."

  Three sets of questions have been prepared, one for policy case studies, one for records case studies, and one for recordkeeping systems case studies.

Case Study Methodology (cont.)

- **First Iterations: Testing Different Solutions in Different Contexts** -- All TEAM members (co-investigators, collaborators, test-bed and resource partners, students), at the TEAM Plenary Workshop, reflect on the data collected and the information generated by the team of researchers for each case study and collectively articulate several possible solutions from which individual plans of action will emerge and be tested.

  These plans of action will include strategy, protocols, functional requirements, procedures, and expected outcome, as needed. The plans are then tested, and test results will include performance assessment of the plans against benchmarks and baselines established in extant research.

Case Study Methodology (cont.)

- **Comparison of First Iterations** -- The results of the tests will be shared among all TEAM researchers, and discussed during the following Plenary Workshop. An assessment of these results will then allow us to reflect on each action, and refine our respective plans of action, also in light of what has been done by the other TEAMs and following directives provided by the International TEAM at its annual Summit in October.
- **Second Iteration: Refining Solutions for Particular Contexts** -- After this assessment, the process will begin another cycle. This second iteration will account for anomalies in the test results, and benefit from the insight gained from a comparison across contexts. In so doing, it will refine our plans and performance measures.

Case Study Methodology (cont.)

- **Comparison and Reconciliation of Results** – At its Summit, the International TEAM reviews the work done comparing the findings across cultures and juridical systems and reconciling them
- **Reflection, Analysis, and Synthesis** -- Throughout the research, the co-investigators and collaborators will reflect on issues and processes and make explicit their assumptions and biases, thereby giving rise to theoretical considerations. The findings, recommendations and products of the case study will be summarized in a Case Study Report for which a Template has been prepared

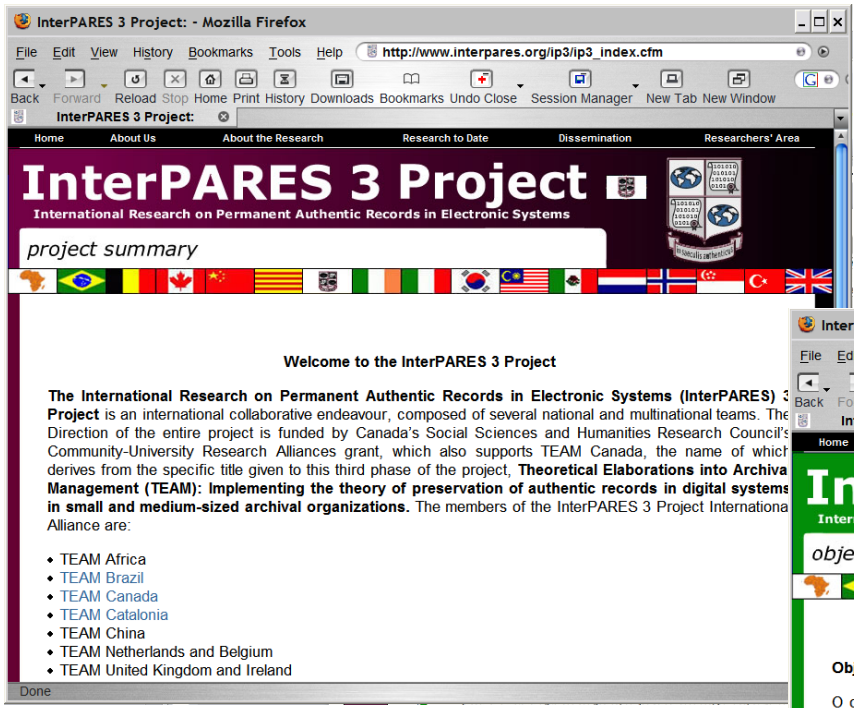
Examples of Case Studies

- **City of Victoria**: Digitized and Born Digital Building Permits; Scheduling Legacy Files
- **Collaborative University Project (UBC, UVic, SFU)**: E-mail Management, Preservation & Access
- **Belkin Art Gallery**: Policy and Procedures for Acquisition and Preservation of Digital Art
- **Canada Tourism**: Analysis of EDRMS Configuration and Related Policies and Procedures
- **North Vancouver Museum & Archives**: Policy and Procedures for Preservation of Digital Records from a Distributed, Multi-jurisdiction Recordkeeping System
- **UBC Alma Mater Society**: Web site Preservation & Access
- **UBC School of Journalism**: Preservation and Access System for High Definition Digital Video Archive
- **Vancouver Police Department**: Policy and Procedures for Management and Preservation of Digital Images Used as Evidence

Examples of General Studies

- Terminology Database
- Research Projects Database
- Bibliography Database
- Survey of e-mail clients recordkeeping capabilities
- Survey of web sites recordkeeping capabilities

InterPARES 3 Web Site



www.interpares.org