A Case study in Geospatial E-science: The Cybercartographic Atlas of Antarctica

D. R. Fraser Taylor, Tracey P. Lauriault and Peter L. Pulsifer Geomatics and Cartographic Research Centre, Carleton University, Ottawa http://gcrc.carleton.ca

Presentation to ANAI Seminario Internazionale InterPARES 2 e Seminari Nazionali sul Digitale, **Archive e Digitale: Quale Futoro?** Milan, Itlay, December 2006.



The geospatial sciences, including cartography, surveying, remote sensing, geodesy, photogrammetry, geography, oceanography and others, are making increasing use of Internet technology. Internet mapping for example is increasingly popular with the general public through Mapquest, Google Maps and more recently, Google Earth. This presentation will consider the challenges of archiving a complex example of geospatial e-science – the Cybercartographic Atlas of Antarctica Case Study C06 for InterPARES 2.



"The organization, presentation, analysis and communication of spatially referenced information on a wide variety of topics of interest and use to society in an interactive, dynamic, multimedia, multisensory format with the use of multimedia and multimodal interfaces" *(Taylor 1997)*

=> Reflection of the Changing Nature of Cartography



Cybercartography in the Information Era

- Helps organize a growing volume of data
- Analyzes and synthesizes complex data new ways (e.g. computer visualization)
- Combines data from remote databases in an interoperable fashion
- Communicates the results to the public in new and more understandable ways

=> Mapping is central in the information era



Maps and Mapping in the Information Era

The map as an artifact:

• New multimedia and multisensory products on-line

The map as an organizing concept:

- Conceptually and geographically link information sources in an integrated framework
- Spatial Data Infrastructures
- Multimedia links add new context to improve understanding

The map as a metaphor:

- The 'Cybercartographic Atlas' as a metaphor for spatially referenced qualitative and quantitative information of all kinds
- Multiple representations to challenge fixed ideas
- Maps involving art and imagination not just 'real' locations



Major elements:

- Multisensory / multimodal vision, sound, touch and eventually smell
- Multimedia formats and the new telecommunications technologies (eg. Web 2.0, mobile devices
- User-centric, interactive Understanding and engaging the user in new ways
- Integrated Part of an information / analytical package including both qualitative and quantitative information
- Compiled by interdisciplinary teams which include cybercartographers but which also involve disciplines not normally associated with cartography
- Collaborative New research partnerships (academia, government, industry, NGO)
- Applied to a wide range of topics of interest to society and responds to societal demand including topics not usually "mapped"

=> Holistic approach: combination of these elements and informed by the interaction between theory and practice



What is Cybercartography?





Ongoing Research: The Use of Sound





Ongoing Research: The Use of Sound





Ongoing Research: The Use of Sound





Ongoing Research: Games & Engagement



Game developed by J-P Fiset (GCRC) & the 'game cluster' - 2005

Improving the engagement: Developing a mod game

•Modification and repurpose of an existing game

- Teach about:
 - Antarctica environment
 global warming
 scientific process
 different perspectives
- Engaging manner



Challenges in Archiving Geospatial Data

Technological Obsolescence

- Fast product development cyber
- Lack of backward compatibility (ArcView 8.1 and 3.0)

Interoperability

Between software, hardware, peripherals and file types

Proprietary Ownership

- Impedes preservation
- Information is in the private sector

Data Refreshing

- Moving from one medium to a new medium
- Risk of losing functionality



Challenges in Archiving Geospatial Data (cont.)

Data Migration

- From one hardware software configuration to another
- From one generation computer to another
- Moves software and operating systems

Data Emulation

- One software tool to make software act as if it were something else
- Long term viability is unknown
- Storage Capacity
- Expanded Metadata Descriptions for New Geospatial Products

(Lauriault & Taylor, 2004)

Framework Based on Bleakly, Denise R., (2002). Long-term Spatial Data Preservation and Archiving: What are the Issues?



Challenges in Archival Geospatial Data (cont.)

- Data archiving versus information preservation
 - Data are more powerful if kept in their context
 - Refreshing and copying to new media removes the context and the objects may be preserved but not archived (*i.e.*, context is lost)
- Clearinghouses are libraries not archives
 - Provides distributed access to data but does not store them nor archive them
 - National Satellite Land Remote Sensing Data Archive is a clearinghouse and an archive
 - GeoConnections Access Portal is a clearinghouse
- Paper preservation inadequate as it does not capture the important interrelationships.

(Lauriault & Taylor, 2004)



Challenges for archiving geospatial data are compounded by the rapid emergence of multimedia and multimodal online products such as cybercartographic atlases and the availability of user-centric data on personal devices.



- "An unprecedented firestorm is incinerating Canada's digital research wealth." (SSHRC Report 2002)
- "Indeed information technology is responsible for much of the loss as storage technology has given a false sense of security against loss and obsolescence." (NCASRD 2005)
- "In archiving terms the last quarter of the 20th century has some similarities to the dark ages. Only fragments or written descriptions of the digital maps produced exist. The originals have disappeared or can no longer be accessed." (Taylor 2005)



The Atlas is a complex digital entity which produces a rich variety of interactive, experiential and dynamic documents. Further complexity is introduced by the fact that the Atlas draws on remote databases in several countries to create these products in real time. But like more than half of the 22 case studies for InterPARES 2 the digital objects in the Atlas were not considered records from the perspective of diplomatic and archival theory.



"Interactions between humans and computer systems, <u>experiences</u> enabled or mediated by experiential systems, and <u>processes</u> which are carried out with at least some degree of spontaneity by dynamic systems are not the residue of action. They are not means of remembering either what was done or what is to be done. In short, they are not records."

Duranti & Thibodeau, 2006



In theory this should not affect the presentation of the atlas. InterPARES' description of records "...must be independent and transcend any and all approaches to preserving them. (Duranti & Thibodeau, 2006)

"The status of a record is not higher than that of data, document or publication. We must know the nature of the material only in order to decide what to do with it, how to manage it so that its nature remains intact and it can continue to fulfill the purposes for which it was generated." (Duranti, 2006, personal communication).



But in practice entities such as the Atlas are less likely to be archived given that they are not considered records from the perspective of preservers.

Acquisition policies, other things being equal, are less likely to choose digital entities which are not records given that a major acquisition decision for archivists is not what to include but what to leave out.



Create "records" which meet the archival definition

- 1. "Freezing" the Atlas at fixed points in time using grid computing approaches
- 2. Create documents describing the processes by which the Atlas was created and operated including extensive metadata

Some of the elements of the existing Atlas, such as the enabling software "Nunaliit" and some of the Atlas modules already have the potential to meet the archival definition of records



InterPARES 2 has already done this by including concepts such as "bounded variability" and accepting that digital components might themselves constitute a record (Duranti & Thibodeau, 2006). Thus radically expanding the concept of electronic record used in InterPARES 1.



This could go further because defining the term record depends on whose perspective is being used as well as which level of abstraction is chosen (Duranti & Thibodeau, 2006). There are at least four perspectives:

- •That of the author/writer
- •That of the creator
- •That of the user
- •That of the preserver

There is no reason why only one perspective should dominate.



At one level of abstraction the Cybercartographic Atlas of Antarctica does not contain "records" from the perspective of the preserver but at another it does. "As a whole the Cybercartographic Atlas can be described as an interactive and experimental enabling record of the Geomatics and Cartographic Research Centre at Carleton University made in the course of a research project funded by the Social Sciences and Humanities Research Council of Canada, with the administrative records of which it shares an archival bond." (Duranti & Thibodeau, 2006)



The way ahead lies in an innovative combination of both approaches. Geospatial Escience must give much more attention to creating records which can be preserved and the preservers must listen to other perspectives on the definition of the term "record".

This is indeed already happening.

Those involved in the creation of the Cybercartographic Atlas are building in elements of preservation and archival strategies during the creation process and the archivists involved have accepted the importance of concepts such as data quality and the metadata used to describe and measure it. This includes and expands on the archival concept of provenance.



Data quality, including accuracy, is central to evaluating all scientific records, whereas archivists are concerned with the authenticity of a record, not its quality. In deciding what e-science records to archive data quality could be an important selection criteria for archivists which has not been the case in the past.



Much has been learned from the interdisciplinary interaction in the InterPARES project and useful guidelines have been created but many challenges remain.

"Record" comes from the Latin "recordari" to remember (Duranti & Thibodeau, 2006) but that which is not remembered may be forgotten!

The debate over what to "record" is of great importance if we do not resolve this problem satisfactorily. Then the increasing volume of interactive, personalized information of WEB 2.0 which do not meet the archival definition of record will be lost. This is already happening and as the interactive, experimental and dynamic information environment becomes the dominant source of information on many aspects of life in the 21st century, we are in danger of losing our cultural heritage.



• The biggest problem is the lack of a coordinated digital information strategy and above all, the resources and institutional structures required to effectively archive the growing volume of digital data

We can solve the theoretical problems and devise methods to preserve the data but the resources and political will required to address they key resource and institutional issues are currently lacking.

• Numerous studies and recommendations for action have been made in Canada, the latest of which was a national summit in December 2006 organized by Library and Archives Canada but to date no significant act has been taken.

