InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Case Study Proposal
CyberCartographic Atlas of Antarctica
Focus 2 - Science

Tracey P. Lauriault
Carleton University

June 2002

CyberCartography an Introduction

The Geomatics and Cartographic Research Centre (GCRC) has received an Innovation in the New Economy (INE) development grant from the Canadian Social Science and Humanities Research Council (SSHRC) to resource the work of preparing a full collaborative research grant proposal. The research aims to explore and increase the understanding of the new theoretical construct of CyberCartography, which sees the map as a new organizing mechanism for digital information in the information era. The term “CyberCartography” was first introduced in 1997 by D.R.F.Taylor at the International Cartographic Association (ICA) meeting in Stockholm.

CyberCartography transforms socio-economic, scientific and environmental data into interactive representations that allow the user to explore and understand spatial patterns and relationships in new ways. The Cybermap is a foundational set of geospatial coordinates with attached attributes that acts as an organizing mechanism for storage, analysis and dissemination of information specific to a variety of topics.

CyberCartography for the New Economy is cartography that is highly interactive and engaging to the user in new ways - “edutainment”, using multimedia formats and myriad telecommunication technologies compiled by a multidisciplinary team and involving new partnerships among research centres, national mapping agencies, the private sector, NGO’s and educational institutions. CyberCartography can be applied to a wide range of topics of interest to society that is multisensory using vision, hearing, touch and eventually smell and taste. It is an information/analytical package including text, graphs, photographs, videos, tables, statistics, models, voice commentary, sound, music, virtual reality and live web camera images.

As an applied discipline the utility of Cartography is determined by its application to the real world challenges of the New Economy. In theoretical terms this project will develop an increased understanding of a new form of multimedia and multisensory perception and communication from a multidisciplinary perspective. This project will develop models showing how CyberCartography can be used to help people navigate through increasing volumes of available data and information.

The Human Oriented Technology (HOT) research team members in the Department of Psychology, Carleton University are key collaborators in advancing knowledge on how different types of information should be structured and presented to enable users to navigate through complex databases and to simplify user interaction with technology. HOT will explore three key issues missed by multimedia developers, namely increasing individual usability, enhancing knowledge acquisition, and developing better end-products by identifying usability bottlenecks in current applications; user needs assessment, developing multimedia prototypes for each of the applications; and carrying out prototype usability testing. HCI and GCRC will advance multisensory research into virtual navigation and informational spaces by exploring multi-information display situations revealing how users navigate virtual and information space in real time and respond to myriad information presentation approaches. Additionally, narrative hyperfiction (NH) a new computer-based literary medium combining traditional modes of story telling with aspects of virtual reality, computer gaming and multimedia art will be researched. NH includes audio-visual data with written narrative allowing readers to select their own story-paths. Improving mapping tools and user interfaces for hyperfiction designers and readers. The cognitive analysis of NH will be integral to the project and will involve applied experiments in reader-response patterns using textual and structural analysis software.

The project will develop two prototype products to test and apply the theory of CyberCartgraphy; namely Canada's Trade with the World Atlas and the CyberCartographic Atlas of Antarctica. Mapping for the blind and tactile maps where possible will also be tested through the development of these projects. It is proposed that the CyberCartographic Atlas of Antarctica become an InterPARES 2 Case Study as part of Focus 2, Scientific activities. More specifically the work of will be under WG 2.1 Creation and Maintenance of Scientific Records Case Study chaired by D.R.F.Taylor.

Description of the CyberCartographic Atlas of Antarctica:

The CyberCartographic Atlas of Antarctica is a proposed Case Study for the InterPARES Domain 1, Task Force 2, Working Group 2.1 Records Creation & Maintenance of Scientific Records chaired by D.R.F. Taylor, Chancellors Professor of International Affairs, Geography and Environmental Studies and Director of the Geomatics and Cartographic Research Centre (GCRC) at Carleton University in Ottawa, Canada.

The Atlas of Antarctica is a SSHRC Innovation In the New Economy (INE) Collaborative Research Project in partnership with the Human Oriented Technology (HOT) Research Laboratory and in collaboration with an international team of Antarctic Scientists and multimedia visualization experts who will share their expertise, laboratories, human resources and data to the project. The Atlas of Antarctica will be a dynamic, interactive Internet and CD-ROM product to prototype the concept of CyberCartography. The historical evolution of the project highlights its collaborative nature beginning with the Canadian Committee for Antarctic Research (CCAR) which discussed and approved the Atlas project in 1999. The Atlas concept was later presented to the Scientific Committee on Antarctic Research's Working Group on Geodesy and Geographic Information (SCAR WG-GGI) meeting in Tokyo, in July 2000. The project was also presented to the Antarctic Treaty's Committee for Environmental Protection (CEP) in Amsterdam in September 2001. Finally, two CyberCartographic Antarctic Workshops were held in Argentina in November of 2001 and Ottawa, May 2002 were attended by leading members of Scientific Committee for Antarctic Research (SCAR), particularly scientists from the domains of biology, geodesy, geology, and geo-visualization expertise from Australia. It is expected that the project will be officially re-endorsed as a SCAR project in July of 2002.

Antarctica is often referred to as the 'Continent of Science' where exploration is for research and where treaties create and environment conducive to collaboration for international scientific study and not exploitation. The CyberCartographic Atlas of Antarctica will therefore incorporate scientific and environmental data into a Canadian led project important to both Canada's International political, economic and scientific interests and those research scientists. The emerging Antarctic Science and Global Linkages Strategy has positioned Antarctica as very important to Canada. Additionally, the Atlas will be of significant educational value, particularly for its user communities, which are decision-makers, scientists and the general public. The project will be widely disseminated over the Internet, become an authoritative source on Antarctica for use in Virtual Universities and be designed in such a way to enable the development of course modules. Additionally the work will be documented in both paper and online journals and become a catchment for existing and future research documents as well as for those because of the project. The CyberCartographic Atlas of Antarctica will reside on its own server and be linked to server nodes in a number of participating countries ensuring wide international distribution and will be a first for the scientific community.

The Atlas of Antarctica will be a unique dynamic product to represent a region where the interaction of physical, biological and human influenced characteristics and processes are of prime scientific and policy importance. There is a growing need for a common method of assembling and portraying information that comes from a wide range of studies and observations. Additionally, Antarctica is a region where the available data and information comes from observations, measurements, and research by scientists and monitoring or survey agencies from many countries, using a variety of techniques, languages, standards and verification procedures. There is a need for a practical, flexible and economic means to assemble and display this information and to identify important gaps such as the ability to archive and document the myriad data, processes, distributed systems, visualization models and digital objects. A comprehensive interactive Atlas, amenable to continental, regional or local scales and capable of superimposing the time dimension as well as overlaying both qualitative and quantitative data is a basic step in meeting this need. User interface testing in collaboration with HOT will ensure broad accessibility, while testing, developing, and implementing metadata and digital archiving processes throughout the course of the project will protect this important digital asset for future generations.

As a region, Antarctica is characterized by long distances, great disparity of scales of information and data with large areas where data are very sparse but with smaller areas of detailed information. A CyberCartographic Atlas can deal with both general information on a continental scale and larger scale vignettes on smaller areas on a variety of topics. Antarctica is a continent of very slow but important natural dynamic change in many characteristics, but one of sensitive environments that are at least in part subject to dramatic or catastrophic alteration as recently demonstration with the breaking up of large areas of the Antarctic ice shelves, in April and May of 2002.

A comprehensive, interactive, multidimensional CyberCartographic Atlas will be an important tool for meeting the needs of the Protocol on Environmental Protection to the Antarctic Treaty by assembling and displaying, in appropriate scales of space and time: information on environmental characteristics; past, current or proposed activities and their impacts; and as a means of monitoring and displaying changes, the results of response actions, and assessments of impacts for management or liability decisions. The Atlas will also provide the general public with a modern and up-to-date authoritative source of information about Antarctica in an interactive multimedia format.

The CyberCartographic Atlas of Antarctica enables decision-makers, scientists and the general public to be data users, explorers and contributors. Of special importance is the On-line Atlas of Antarctic Research maintained by the United States Geological Survey (USGS), the extensive text, images, photographs and map data resources of the USGS Antarctic Resource Centre, the resources of the British Antarctic Survey and Chinese Antarctic Center. The proposed Atlas is amenable to, and will build on, these and other existing databases such as the Antarctic Digital Database, the RADARSAT Mosaic of Antarctica, and the Global Change datasets. In the light of important initiatives a hub and node organizational approach will be adopted to enable a variety of approaches and autonomous contributions to the Atlas such as those previously noted with subject specific and/or organizational sub-nodes such as (e.g. Centro Nacional Patagonico (CENPAT) or the SCAR-Geodesy and Geographic Information (GGI)). Researchers who have important datasets on the region but do not have their own infrastructures will have opportunities to contribute their work. The overarching concept is to enable the space for people develop their own products independently yet be able to interoperate. The Atlas will be a distributed Open GIS Consortium (OGC) compliant Internet mapping tool (e.g. 'GML' - a means of standardizing handling of geographical data processor with XML/DTD/and other standards embedded in GML). To insist on standardized approaches for all nodes is currently unrealistic therefore it is proposed that “middleware” or “fusion ware” be used to allow nodes to transfer data to the coordinating hub. This will mean that existing data in a variety of different formats can be used without the effort and cost of conversion to a standard format. Sub-nodes, which do not have the technical capability to present data in a geospatial format, can link in with the nearest national node where the data can be prepared in an appropriate form. Organizations and individuals holding data on Antarctica who are not geospatial specialists could therefore contribute to the Atlas[1]. GCRC with its partners and collaborators will develop the product with the long term objective of transferring the project’s technology, knowledge and infrastructure to a lead organization who will maintain it as their own.

Antarctica for historical reasons, is a region rich in international scientific collaborations with a large amount of machine-readable and georeferenced data. The plethora of data on the region provides numerous thematic and data opportunities therefore decisions on content will be carried out in cooperation with partner agencies and the SCAR-GGI working groups. It is expected that data quality, availability, the degree of scientific research and topical subject matter will be key determinants influencing choice. A number of databases have already been identified including:

[1] The Jason Project VIII Frozen Worlds 2001-2002 (http://www.jasonproject.org/jason_project/jason_project.htm) or Students on Ice Learning Expeditions to Antarctica and the Arctic (http://www.studentsonice.com/splash/splash.htm).

Antarctic Digital Database[2], GRID-ARENDAL databases[3], RADARSAT[4], China Antarctic Data Centre, and the Italian Composite Names Gazetteer[5]. General themes of the Atlas are Geoscience, Life science, Physical science, Antarctic Environment, History (Science or General), Antarctic Treaty Systems, Climate, Fisheries, general education and knowledge for the public. These themes will be accessed via research and tested Internet user interfaces to meet the needs and user requirements of the identified target groups.

Earlier discussions have pointed to the domain of geodesy to inform framework data maps, and to explore recent research findings on plate models, point networks, evolution of astronomic observation systems, tracking technology for ships, and mapping protected areas, treaties boundaries, locating historic monuments, and etc. The SCAR-GGI WG has numerous important time series data sets, research papers, and monographs in myriad format. The biology of Antarctica is an important and popular subject. Atlas collaborators from Argentina hold important and significant biological research data on Seals and Penguins, particularly related to migration patterns, food sources, habitats, tracking of movements, in relation to subjects such as overfishing and climate change. The physical geography of Antarctica is also the subject of international research on topics such as sea ice models, ice change, ice shelves, icebergs, and breakups. Additionally, data are available on air and water temperature change, krill, sea caps, crustal plate movement models and meteorites. Finally, Antarctica is an important tourist destination and education topic for discussion in schools, therefore representing data and topics of interest while enabling the sharing of information to and from both these groups will also be integral to the Atlas (e.g. ship routes, history, and the ability to post image and text diaries).

The data are available in myriad formats such as satellite, radar, air photo images, scanned images, video, WebCams, dbases, text, and MP3s, which can be georeferenced. These data and topics are particularly conducive to being represented in new forms such as dynamic digital maps (e.g. distributed maps and dBases, large and small scales, bathymetric, thematic, name gazetteer, and etc.), charts, data visualization models (e.g. dynamic atmospheric and sea ecosystems), sound (e.g. stories, animal sounds, folkloric music and songs,), animation[6] (Ice Molecules, ice breakup models), stills, dynamic time series (e.g. ice falls, continental shifts) and real time (e.g. migration patterns of seals, ship and air craft tracking, geodetic points), and hypertext (e.g. georeferenced papers and narratives). Since the Antarctic Atlas users and data providers reside all over the world a multilingual interface strategy is also planned. CyberCartography can therefore be seen as an integrator of data and a bridge between disciplines and nations.

---

[2] Antarctic Digital Database, is the premier source of vector topographic data for Antarctica. First published on CD-ROM in 1993, the current version is available for download from the CGDI GeoGratis (http://www.nerc-bas.ac.uk/public/magic/add_home.html)

[3] GRID-ARENDAL databases is the UN's information Centre on the environment (http://www.grida.no/)

[4] RADARSAT Canadian Space Agency Earth-observation satellite data (http://www.space.gc.ca/csa_sectors/earth_environment/radarsat/default.asp)

[5] See the HISTORY AND STRUCTURE OF THE SCAR COMPOSITE GAZETTEER OF ANTARCTICA at http://www.pnra.it/LUOGHI_ANT/HTML_en/history.html.

[6] Moviemol is a program for visualization and animation of molecular structures for SGI and IBM RISC/6000 workstations and for PCs running DOS (or Linux, but only a beta version exists that unfortunately does not work properly). Moviemol typically takes sets of molecular coordinates and displays each of these sets or frames one after the other to create a "movie". (http://www.fos.su.se/physical/lars/moviemol.html) or Animation of Larsen B breakup, 31 January to 7 March 2002 (http://nsidc.org/iceshelves/larsenb2002/animation.html).

CyberCartographic Atlas of Antarctica InterPARES 2 Case Study Rationale

The Atlas of Antarctica will be the first international scientific CyberCartographic research collaboration considered to be of national and international import to decision-makers, scientist and the public. Antarctica is the 'Continent of Science' and the Atlas of Antarctica will be testing and prototyping the theoretical concept of CyberCartography, which may be the new way to navigate and explore in the information era. The dynamic, multidimensional, multisensory and multimedia Atlas will become an important scientific digital knowledge asset that will from its inception include archiving as an integral component of the project. The Atlas therefore provides the opportunity for a full phase implementation of a digital metadata and archiving processes. Partners and collaborators are from numerous scientific domains, providing Inter Pares II with a ready-made network an important and internationally recognized innovative digital geospatial product. Additionally there is a great opportunity for cross-domain collaboration between leaders in numerous scientific fields and the artistic, musical, and hypertext communities, which may provide critically, needed information for the multisensory archiving components of the Atlas.

International and national support for the project is significant. The Canadian Department of Foreign Affairs and International Trade (DFAIT) provided grant support from the Going Global Program, the Canadian Polar Commission (CPC) with cash and ongoing advice on content, as did the Canadian Committee for Antarctic Research (CCAR). International partnerships include the International Cartographic Association (ICA), Commission on Maps and the Internet and the Commission on Mountain Cartography with expertise from the United States, Austria, and Australia. The Centro Nacional Patagonico (CENPAT) in Argentina and a working group of the Scientific Committee for Antarctic Research (SCAR) are also part of the participants. Additionally, the Atlas of Antarctica has been favourably received by the Antarctic Treaty Consultative Meeting, the governing body for Antarctica while twelve nations have agreed to participate or have expressed an interest. The caliber of these collaboration efforts highlights the importance of this work and further reinforces the importance of the Atlas as a Knowledge asset for current endeavors and future generations.

Currently Antarctic data are readily available but are not well documented nor catalogued while the added component of spatiality suggests that metadata may be an afterthought. Additionally much of the data is in the hands of research scientists that have conducted research on a projects basis but not as members of large institutions, therefore the data may not be in machine readable formats (e.g. fish counts documented on paper captured on a ship by ship basis). In other cases data may be well documented and catalogued but are in a format that is platform specific, or captured at scales, levels of generalization, projections or levels of accuracy that are not compatible. Further domain specific data semantics and dbase structures complicate the interrelationship of spatial data for map rendering and analysis. Compatibility is attainable by way of adding server middleware or fusionware while mathematical, geometric transformations and aggregation may further enable map layering. Data transformation methods change the data, to enable map rendering which in turn requires changes in metadata, cataloguing and potentially new records. In other cases media conversion are required such as videos will need to be digitized while photographs, maps, and momentos may need to be scanned, transforming their formats but expanding their utility in a digital environment. In all cases data lineage (I.e. authenticity) will be key in preserving sources, authorship and copyright. Machine readability in a distributed Internet environment needs to be ensured for future generations, and this is of growing concern and unresolved problem for academia, industry and government alike.

The rendering of digital Internet maps and other visualizations from myriad data sets in a distributed environment are key to CyberCartography while the ability of visualization re-construction are important to the archiving of digital data (see Figure 1, Preliminary CyberCartography Project Schema). Time tagging of georeferenced information will therefore be of great import as will be the documentation of the processes of creating online digital maps, models and georeferenced visualizations. In a distributed environment databases are made accessible by way of standard server protocols and database-sharing tools captured formal agreements between organizations. A distributed environment also means that data are not housed by GCRC in a warehouse environment but are shared via adopted and defacto standards such as Open GIS Consortium (OGC) and ISO Technical Committee 211 standards. Data remain the full responsibility of the provider, to update, modify and maintain while data are described and catalogued in digital clearinghouses (e.g. GeoConnections Access Portal) in a number of metadata standards. Map and visualization renderings in most cases comprise access to numerous framework data sets (e.g. continent, water, topography) and thematic sets (e.g. seal migration, ice flows, water temperatures, geodetic stations and name gazetteers). This environment is an advantage in terms of information sharing, knowledge dissemination and creation while concurrently posing particular archiving challenges. There are a number of unresolved questions in terms of preserving digital geospatial data in a CyberCartographic environment, such as what is a record, i.e. the rendered map, the data, the metadata catalogue or is it the process? How to archive data in a non-warehoused international Internet collaborative environment? Can CyberCartography enable the creation of non-centralized autonomous consensual standards and processes to enable digital data archiving? How will archiving costs be recovered a distributed virtual Internet project?

Scientific knowledge on Antarctica is an international treasure that will be abstracted in the CyberCartographic Atlas. The creation of the Atlas will be a cumulative and interactive process pushing the frontiers of mapping, multimedia and Internet technologies to represent scientific data accessible to specific target audiences. Moreover the data will be discoverable via a human computer interaction designed user interface. The preservation of the data and the UI to inform new creators and for “re-analysis of existing data may lead to different conclusions. Thus, archived data allow for the formulation of new hypotheses and may unexpectedly change the relative importance of the data”[7]. Additionally, the project will take into consideration changing technological developments and the preservation of both hardware and software to ensure data accessibility through time and/or migration to new storage media and formats.

The project is a unique case study opportunity that is expected to be beneficial to those engaged in the archiving of digital information in the field of multimedia, the study of object re-creation and of particular import to the geomatics community that is grappling with these unresolved issues. Significant digital products like the Canada-U.S. Great Lakes Mosaic originally created in DOS based programs are not accessible in Windows NT while new distributed mapping projects emerge without any consideration for archiving such as Canada's Oceans Program Activity Tracking (OPAT) system. Members of the Association of Canadian Map Libraries and Archives

[7] Science Committee on Data for Science and Technology (CODATA) (http://www.codata.org/canada/datact98.html)

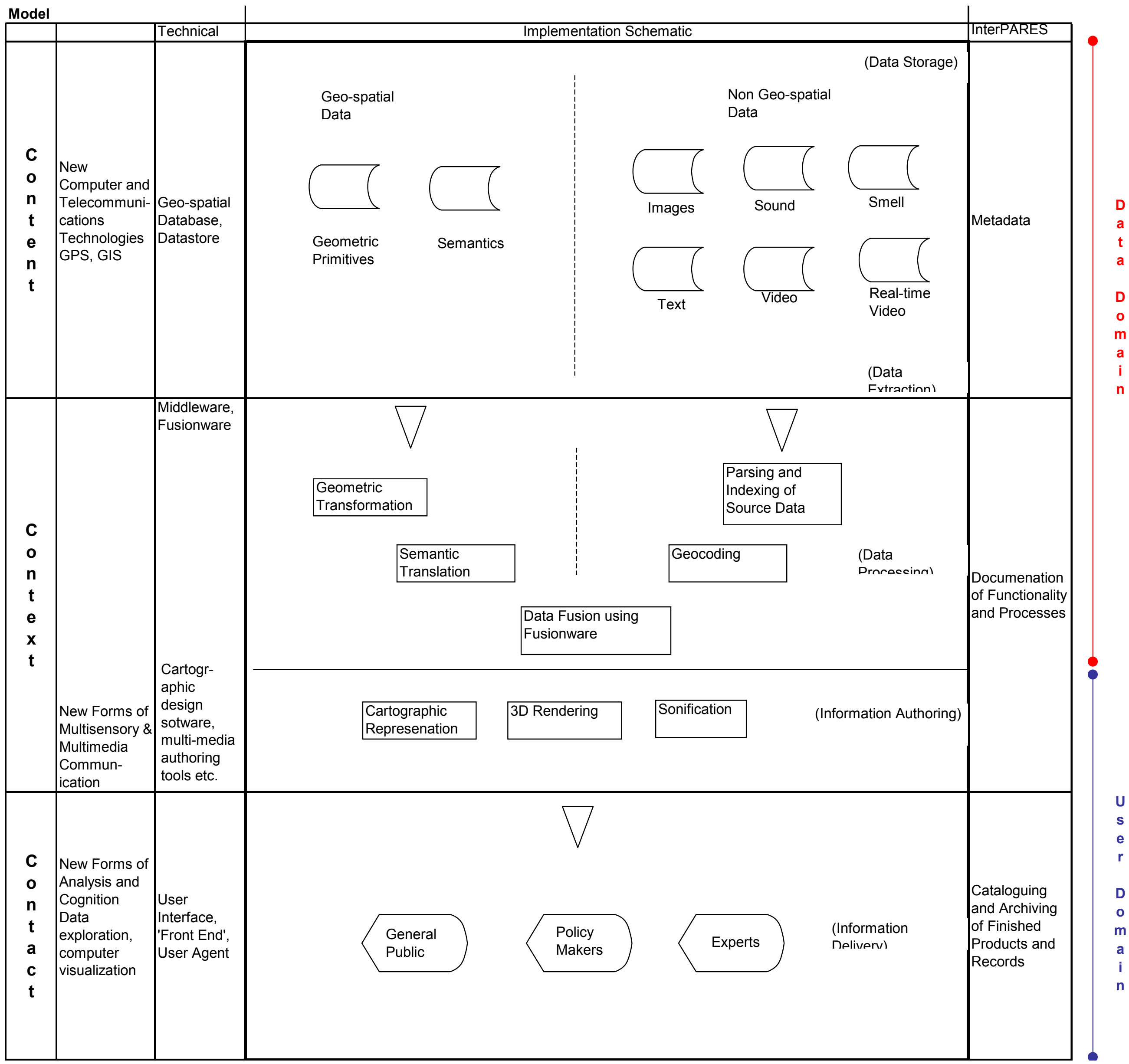

Figure 1.
Preliminary CyberCartography Project Schema

in a recent SSHRC conference in May 2002, indicated that more and more data are available on CD-ROM year there is no clear cataloguing process and the preservation of data in this media is still yet uncertain. The Geological Survey of Canada Ottawa Library has implemented Dublin Core and discovered many pitfalls with cataloging their digital data along the way. The Canadian Geospatial Data Infrastructure (CGDI) Supportive Policy Advisory Node in collaboration with the Geomatics Industry Association of Canada (GIAC) have directed some effort to unravel metadata processes and recognize that there is a long way to go before archiving becomes an operational reality for geospatial data producers and the scientific community. Finally, the

National Archives of Canada recognizes this as a critical gap in the preservation of national heritage. The CyberCartographic Atlas of Antarctica is expected to provide some important answers toward unraveling some of the complicated questions on the preservation of scientific and geospatial digital data.

Geospatial metadata, Internet mapping technology and multimedia tools are well advanced, while the preservation and archiving of digital data in Canada is considered a lost national opportunity. In a recent paper by the Social Sciences and Research Humanities Research Council in Collaboration with the National Archives of Canada identified Canada as one of the few G8 countries that does not have a national facility for managing, preserving and accessing research data. The Canadian Global Change Program and at the OECD Ottawa Workshop, pointed out that Canadian researchers are at a competitive disadvantage both at home and in international research programs without such a facility[8]. The CyberCartographic Atlas of Canada may be the first project of its kind to incorporate at the very beginning of the project the process of preserving the entire life cycle of its data. The project is expected to develop processes, methods, tools and guidelines for archiving and explore the possibility of expanding metadata tools for this purpose. Furthermore, the ability to re-create visualizations will be an integral component of the project.

Methodology

The CyberCartographic Atlas of Antarctica is a project that exemplifies a new and emerging digital environment that is experiential, dynamic, and interactive which abstractly represents scientific research and geospatialy referenced data in a mutisensory fashion. GCRC expects to utilize the data gathering tools developed in InterPARES 1, refer to the best practices and results in InterPARES 1, extend the work of diplomatics and finally identify the best practices in the geomatics and scientific data archiving communities. The work of archiving scientific and geospatial research data in digital form however differs from the work of InterPARES 1 therefore these tools will have to be modified to better suit the processes of map rendering in a distributed Internet environment and the particularities of this type of scientific data. The following is a preliminary and incomplete list of proposed tools for gathering data for this case study, which include questionnaires, templates for analysis, entity and activity models, work flow studies, and the development of methodology statements. It is expected that the research methodology component of this case study will be developed collectively with members of Domain 1, Task Force 2, Working Group 2.1 Records Creation & Maintenance of Scientific Records to ensure that the data gathering processes are comprehensive enough to capture a wide spectrum of activities that exemplify the complexities of the complete project cycle (development to preservation). In particular to identify what a record is in this context, its attributes, record creation, preservation processes and authentication. Tools, processes and guidelines will also have to be developed to test findings. It is also expected that the broader InterPARES 2 working groups will share their expertise to assist with research tool development and to guide the process.

[8] Social Sciences and Humanities Research Council and the National Archives of Canada Press Release (http://mmsd1.mms.nrcan.gc.ca/archives/bg_paper-e.asp).

Possible tools

a) Modified Case Study Interview Protocol (CSIP)
b) Modeling records creation processes tools
c) Collaboratively created business process analytical tools
d) Functional Analysis tools relating records to the function that they support and revealing where in the system records are likely located.
e) Modified appraisal model from InterPARES 1 framework for information gathering
f) Modified Template Element Data Gathering Instrument (TEDGI) for Analysis
g) Replicating system tools for hands-on study of records and to replicate the work on the InterPARES www site.
h) Testing of preservation strategies by transfering a sample to the InterPARES computer to test preservation strategies or to try the InterPARES 1 preservation model or to prototype a preservation system.
i) Modified Documents Mapping to analyze InterPARES Authenticity Requirements against existing standards
j) Modified Survey of Preservation Practices and Plans
k) Literature review on the preservation of Scientific Data, standards, metadata and to identify best practices
l) Other tools to be developed collectively with Task Force WG members.

Objectives

The objectives of the CyberCartographic Atlas of Antarctica as an InterPARES 2 case study are to explore the process of record creation, the function of the record within the activity in which it participates, to preserve authenticity over time along with its characteristics, process of creation and function and purposes for which it is kept by its creators. Additionally to determine which of its features will allow its authenticity to be determined to ensure its records are generated to carry it forward for future generations.

The project aims to identify the process of preserving, authenticating and extending the work of InterPARES 1 diplomatics research to the complete project cycle of the CyberCartographic Atlas of Antarctica. The CyberCartographic Atlas of Antarctica is a dynamic, distributed, multimedia, multisensory and multi-dimensional Internet project that will transform socio-economic, scientific and environmental data into interactive abstracted map representations that allow its users to navigate, explore and understand spatial patterns and relationships in new ways. The case study is proposed for the InterPARES 2 Domain 1, Task Force 2, Working Group 2.1 Records Creation & Maintenance of Scientific Records.

Spatially referenced data have always been integral to map creation and the abstracted rendering of these data have traditionally been in the hands of cartographers. The paper map was the final product or record that was then catalogued and preserved. Generally data used to create the map were lost and not included as part of the complete record set. The mapping paradigm of geographic information systems (GIS) has positioned map data on equal par with the final map as a result geospatial data and their attributes have become records in and of themselves. Today, in a distributed Internet mapping environment data remain the responsibility of the providers and data are dynamically rendered in real time from disparate databases accessed via servers from around the globe in an international and inter-organizational cross domain environment. In this context, the data, rendered map, data sharing protocols, data requests, metadata and the process of rendering are all equally important. The objective of this case study therefore is to understand the complete life cycle of the CyberCartographic Atlas of Antarctica with regards to authentication, and preservation. Additionally to develop best practices, methods and standards to inform the broader scientific research, geomatics and multimedia communities on the topic of digital data preservation.

The following is a series of preliminary unresolved issues and questions regarding the preservation and authentication of scientific spatially referenced digital data, and it is expected that InterPARES Task Force 2.1 working group members will collaboratively develop criteria to asses which of these are of greatest import and develop and modify data gathering tools and processes to resolve these and/or other identified issues.

Preliminary objectives, issues and questions

1. Determine what type of documents are made or received or set aside in the course of scientific activities that are expected to be carried out on during the course of the project and define their purpose? What types of electronic documents are currently being created to accomplish those activities?
2. Have the purposes for which these documents (#58) are created changed?
3. What are the nature and the characteristics of the traditional processes of document creation in each activity? Have they been altered by the use of digital technology and if yes how?
4. What are the formal elements and attributes of the documents generated by these processes in both a traditional and digital environment? What are the function of each element and the significance of each attribute?
5. What is the manifestation of authorship in the records of each activity and its implications for the exercise of intellectual property rights, authoritative and valid sources, and the attribution of responsibilities?
6. Does the definition of a record adopted by InterPARES 1 apply to all or part of the documents generated by these processes? If yes, given the different manifestations of the record's nature in such documents how do we recognize and demonstrate the necessary components that the definition identifies? Or is it possible to change the definition to maintain theoretical consistency in the identification of documents as records across the spectrum of human activities?
7. What other factors identify a record other than those identified by diplomatics?
8. As cartography moves into the a distributed Internet environment and enters into information exchanges based on more dynamic web presentation, is there a neglect to capture adequate documentary evidence of the occurrence of these transactions?
9. Is the move to more dynamic and open-ended exchanges of information blurring the responsibilities and altering the legal liabilities digital scientific data, particularly regarding their spatial accuracy?

10. How do record creators traditionally determine the retention of their records and implement this determination in the context of each activity? How do record retention decisions and practices differ for individual and institution creators? How has the use of digital technology affected their decision practices?
11. Identify and list geospatial and Antarctic science archive stakeholders;
12. Raise geomatics and scientific community awareness of archiving function efforts
13. Identify and document data accessibility and future retrieval issues along with the re-creation of visualizations and map rendering in this context;
14. Select criteria and principles for the long-term archiving and preservation of georeferenced scientific data;
15. Identify the cost implications for long-term preservation of data and different models of funding (e.g. cost implications and models of funding for reproducing original or experimental data vs. derived data)
16. Identify issues related to media and standards for data sharing, data transformations and metadata.
17. Identify specific priority areas (e.g. time series and environmental monitoring).
18. Explore, identify and describe issues, benefits, gaps, functions, roles and best practices from other scientific research data archiving initiatives (e.g. International Council for Science Committee on Data for Science and Technology (CODATA), Woods Hole Oceanographic Institution (WHOI) digital archive, International Council for Scientific and Technical Information (ICSTI), joint UNESCO/International Council of Scientific Unions (ICSU), the Université de Montréal social statistics research facility, University of New Brunswick protocols for indexing and describing databases and digital objects).
19. Explore, identify and describe issues, benefits, gaps and best practices from other multimedia archiving initiatives (e.g. CANARIE)
20. Explore, identify and describe issues, benefits, gaps (e.g. functionality, description) and best practices of current Internet metadata standards (e.g. FGDC, OGC, ISO TC 211, ESRI tools, ANZlic, Access Portal, Open Archival Information Systems (OAIS Standard), University Consortium for Political and Social Research (ICPSR, US) Dublin core 'resource discovery metadata', and etc.)
21. Explore the possibility of expanding current well-developed geospatial metadata tools, processes and standards (e.g. FGDC, OGC, and CGDI Access Portal).
22. Compare metadata tools and applications with other cataloguing processes (e.g. Dublin Core, Dewey Decimal Classification, and other multimedia or music initiatives)
23. Explore the possibility of developing hybrid tools to be embedded into the CyberCartographic Atlas development process that will both capture metadata, and enable digital archiving.
24. Carry out a literature review and develop a comprehensive annotated bibliography of important papers on the subject archiving spatially referenced scientific digital data
25. Capture the lifecycle processes of the CyberCartographic Atlas Data
26. Identify what a record is in the context of CyberCartography, namely is it the data, the metadata catalogue, the rendered map of the process of rendering?
27. Describe context of how data were generated, analyzed, variables constructed, and sharing of data and map rendering in a distributed Internet environment etc.;
28. Describe changes over time - generation of new independent variables, better understanding of substance, new models of analyses; when data are in remotely updated databases

29. Describe saving, storing and maintaining scientific data and dbase management (e.g. dbase size, version, proprietary software, display and visualization program);
30. How should documentation be done, and what should be made available;
31. How should changes to database be saved; particularly in a distributed environment?
32. Should data be saved at every point of time or just an archive of important results; or should only the final rendering be saved or the components of the rendering be saved?
33. How to preserve operating systems, hardware, and storage media;
34. Who pays for data preservation, storing, and archiving particularly in an international inter-organizational project of this kind
35. Identify issues of archiving and preserving individual records vs. entire collection;
36. Reproducible data vs. one-time observations;
37. How will calculated (computed) results be handled;
38. Identify issues related to the interoperability of data (e.g. work of the International, Interoperability, Institute 3i)
39. Archiving intermediate results vs. final results. Should both be archived?
40. Digitizing data for archiving and preservation; in a distributed environment and how to authenticate contributions?
41. Proprietary interests in data archiving and preservation (software interests; policy; data, copyright, licensing, and etc.).
42. Issues related to media
43. Identify strategies for implementing electronic data archiving, preservation and interoperation.
44. Integrate current and historical knowledge base using Internet technologies
45. Present a holistic view of a variety of topics (i.e. climate change, tourism) to a broad range of users
46. Create a framework to help better understand importance of Antarctica in terms of sustainable development
47. Facilitate increased cooperation among agencies
48. Foster information exchange between agencies
49. To focus on the record creation and the management of a multidimensional, multisensory, multidisciplinary, multi-media CyberCartographic Atlas of Antarctica.
50. Develop and test crosswalk metadata tools
51. Test authenticity requirements and appraisal methods models from InterPARES 1
52. Develop a prototypical system for the preservation of the records in question
53. Integrate and make tangible Geomatics, ICT multimedia, and archive thesauri
54. Demystify the terminology and processes of archiving digital data
55. Determine dynamic variables related to cartographic animation – duration, rate of change, order and phase
56. Develop and test standard automated system for transfer of cataloguing data to bibliographic databases
57. Develop selection criteria regarding the presentation of components and process of developing a tested and researched user interface
58. Address the issues of the quality of spatial data their archiving for implementation in operational routines for analysis and visualizations of spatial data sets.
59. Expand are data description methods and assess them against the user needs

60. What criteria are used to assess if a data set meets needs? What are the known quality, availability, and ease of accessibility?
61. What should the data producer include in a spatial data set to enable preservation?
62. How important is the criterion of spatial data quality on limiting map rendering?
63. Develop a typology of functions
64. Outline a mechanism for input and storage of functional metadata -implementation
65. Design a method of distributing and accessing the functional metadata - distribution
66. Develop methods to minimize the amount of manual data input required for database managers
67. Develop easily, intuitive and implementable archiving methods

Description of the team

The initial team will be members of Domain 1, Task Force 2, Working Group 2.1 Records Creation & Maintenance of Scientific Records chaired by D.R.F. Taylor, Chancellors Professor of International Affairs, Geography and Environmental Studies and Director of the Geomatics and Cartographic Research Centre (GCRC) at Carleton University in Ottawa, Canada. Currently, WG 2.1 consists D.R.F.Taylor as chair, Su-Shing Chen, Rick Luakowski, Brad Abbot, Gianni Paolini, Barbara Pernici and GCRC Research Assistant Tracey Lauriault.
The case study will involve members from many sections of the entire InterPARES science domain and will require input from other domain WGs. GCRC partners and collaborations involved in the creation of the CyberCartographic Atlas will also be members of the team with varying degrees of involvement at various of the project.

Timeline

The sequence and timing of activities will be those established for the InterPARES 2 project, the division of labour, tasks, methodologies, logistics, and any budget implications are to be determined by the InterPARES. Existing documentation of experience from InterPARES 1 clearly provides excellent building blocks.

Conclusion

The CyberCartographic Atlas of Antarctica can provide a unique and comprehensive case study for InterPARES 2, as it will involve several science disciplines such as geodesy, geophysics, geology, biology, geo-statistics, geomatics, ecology, and oceanography. The project will also include research on scientific data from myriad sources, in various media shared in a distributed Internet environment. The project is notable for the caliber of its collaborators and partners who are scientists, cartographers, psychologists and visualization experts committed to making scientific research data accessible to a broad range of user communities in many languages. The combination of experts from these fields with the broader InterPARES 2 collaborators and Working Group members will provide an exceptional array of expertise to advance the objectives of ensuring that scientific research digital data withstands the test of time and become a knowledge asset for future generations. The project has been positioned as important to Canada and the International Scientific Community and Antarctica as the ‘continent of science’ reinforces the importance of the creation of the Atlas for broad dissemination but more importantly it highlights the critical need to preserve the vast and rich amount of digital scientific research data. It is a knowledge asset created from historical peaceful agreements for the purpose of scientific research and not exploitation, further highlighting the needs for its preservation, in order to gain new insight from the past, to better plan today and to inform the future decisions.