2 The Preservation and Archiving of Geospatial Digital Data: Challenges and Opportunities for Cartographers

Tracey P. Lauriault, Peter L. Pulsifer, D.R. Fraser Taylor

Geomatics and Cartographic Research Centre (GCRC), Department of Geography and Environmental Studies, Carleton University
tlauriau@gmail.com, ppulsife@connect.carleton.ca,
fraser_taylor@carleton.ca

Abstract

In terms of preserving our digital cartographic heritage, the last quarter of the 20th century has some similarities to the dark ages. In many cases, only fragments or written descriptions of the digital maps exist. In other cases, the original data have disappeared or can no longer be accessed due to changes in technical procedures and tools. Where data has not been lost, as with the Canada Land Inventory, the cost of recovery has been high. Based on experience gained through participation in a major research project focused on preservation, the development of several digital cartographic frameworks, systems and artifacts (e.g. Maps and atlases), multidisciplinary work with archivists, data preservationists, data librarians, public officials and private sector cartographers, the authors discuss possible strategies toward the preservation of maps, geospatial data, and associated technologies – cartographic heritage. The chapter also discusses the findings of two International Research on Permanent Authentic Records in Electronic Systems (InterPARES 2) studies: Case Study 06 The Cybercartographic Atlas of Antarctica and General Study 10 on Preservation Practices of Scientific Data Portals in the natural and geospatial sciences. The chapter concludes with an overview of some of the questions and research opportunities that are emerging from the discussion.

M. Jobst (ed.), *Preservation in Digital Cartography*, Lecture Notes in Geoinformation and Cartography, DOI 10.1007/978-3-642-12733-5_2,

"the mottling, wavy lines, and occasional offset in lines on this map bear evidence to several layers of correction sheets applied sequentially to the original. Over an extended period of time, these maps became *quite weighty palimpsest of urban development in the most literal sense"*

Robert R. Churchill (2004, p.16) describing historical urban maps of Chicago.

Quod non est in actis, non est in mundo - What is not in the records does not exist

Old Latin Proverb

2.1 Introduction

In terms of preserving our digital cartographic heritage, the last quarter of the 20th century has some similarities to the dark ages. In many cases, only fragments or written descriptions of the digital maps exist. In other cases, the original data have disappeared or can no longer be accessed due to changes in technical procedures and tools. As new technologies and approaches to data collection and cartographic production are established, new challenges in preserving and archiving of geospatial digital data and maps are emerging. This chapter examines some of these emerging challenges and presents possible strategies for the preservation and archiving of contemporary digital maps, geospatial data, and associated technologies. Developments in contemporary cartography are presented along with an identification of key issues related to preservation and archiving. This discussion is followed by an elaboration on these issues informed by applied historical review and applied research. The chapter concludes with an overview of some of the questions and research opportunities that are emerging from the discussion.

2.2 Contemporary Cartography

All of the geospatial sciences, both natural and social, are making increasing use of Internet technology. Internet mapping, such as Google Maps and Google Earth, is increasingly popular with the general public. Many other Web 2.0 tools such as online photo, audio and video sharing ser-

vices; blogs and Geowikis (e.g., Visible Past Initiative) are geo-enabling their content by including the ability to add a georeference and/or providing geographically encoded objects (e.g. GeoRSS) capabilities. The United States Library of Congress for instance is now making some of its photo collections available in Flickr, a popular photo sharing service (Library of Congress 2009), allowing these photos to be used as multimedia content in a variety of historical thematic maps and atlases. National mapping organizations (NMOs) are primarily producing geospatial data that are born digital while continuing to digitize and scan older paper maps and airphotos. NMOs are now rendering their data in online maps and atlases (e.g., The Atlas of Canada), distributing frameworks (e.g., GeoBase) and other key datasets using interoperable data services. The data used to render maps not only come from many sources, but are also now being rendered, in some cases in real time from myriad distributed databases via a number of sharing protocols (e.g http://gcrc.carleton.ca/isiuop-atlas).
Concurrently, many datasets are being registered into online discovery portals using international metadata standards (e.g., ISO 19115) and being distributed with new licenses (Wilson and O'Neil 2009, GeoBase, GeoGratis, Science Commons). Cybercartography, a term coined by D. R. Fraser Taylor, is also a new way of approaching the theory and a practice of map making. It is "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 and multidisciplinary format" (Taylor 1997, 2003). Cybercartographic atlases, for example, are increasing the complexity of map making while also enriching the way topics, issues and stories are rendered and conveyed (e.g., Living Cybercartographic Atlas of Indigenous Perspectives and Knowledge, Atlas of Canadian Cinema or the Atlas of the Risk of Homelessness). Geospatial data are also increasingly being included into and are inseparable from models and simulations (e.g., general circulation models). Regardless of who is creating geospatial data and maps the increasing volume of digital maps; and irrespective of their increasingly varied and complex form or functions they all have one thing in common; few of these new products are being effectively preserved and many are being permanently lost. There is generally a naïve belief and practice by digital map creators that back-up and storage techniques are enough to ensure preservation. Nothing could be further from the truth. Without a clear strategy for ensuring that the data will be accessible using future technologies, and

that storage media remain intact, simple backup and storage is not sufficient. We are losing digital spatial data as fast as we are creating them.

2.3 Geospatial Data and Portals

Maps are knowledge representations, and in pragmatic terms, data representations. Cartography cannot exist without data, which come in many forms and are cartographically rendered in an increasing number of new ways such as cybercartography, distributed online mapping, and municipal enterprise systems. Databases are often thought of "as relatively new forms of records, the essential concept of structured information gathering has existed for thousands of years" (Sleeman 2004, p.174). For example, the "Ptolemaic census, written in demotic, unearthed by the archaeologist Flinders Petrie in Rifeh in the early twentieth century…was similar to modern-day census data, tabular in structure with data divided into columns and equally as indecipherable to the naked eye" (Sleeman 2004, p.173). Data are therefore more than just "facts, ideas, or discrete pieces of information" (Pearce-Moses 2007). Geospatial data can also be "numbers, images, video or audio streams, software and software versioning information, algorithms, equations, animations, or models/simulations" (National Science Foundation 2005, p.18) which have a spatial referent. Geospatial data, according to one definition by the US Environmental Protection Agency (EPA):

"identifies, depicts or describes geographic locations, boundaries or characteristics of Earth's inhabitants or natural or human-constructed features. Geospatial data include geographic coordinates (e.g., latitude and longitude) that identify a specific location on the Earth; and data that are linked to geographic locations or have a geospatial component (e.g., socio-economic data, land use records and analyses, land surveys, environmental analyses)" (EPA 2009).

With advancements in scientific methods, computer technology and cartographic innovations, we can infer more from data than ever before. Cumulative sets of data can assist with understanding trends, frequencies and patterns, and can form a baseline upon which we can develop predictions, and the longer the record, the greater the confidence we can have in conclusions derived from them (National Research Council 1995). Thus, preservation of geospatial data is important as these data can provide the raw materials required for future unanticipated uses, especially as technology advances. The assembled record of geospatial and scientific data “has dual value: it is simultaneously a history of events in the natural world and a record of human accomplishments. The history of the physical world is an essential part of our accumulating knowledge, and the underlying data form a significant part of that heritage” (National Research Council 1995, p.11). These data also portray a history of our geographic, scientific and technological development while databases “constitute a critical national resource, one whose value increases as the data become more readily and broadly available” (National Research Council 1995, p.50). It is cost effective to maximize the returns on these investments by preserving them for the future and disseminating them widely. In addition, the costs of preserving and archiving data are relatively small in comparison with the costs of re-acquisition. The cost of repairing a lost and abandoned dataset yields partial results at a significant cost, as seen in the Canada Land Inventory (CLI) example discussed later in this chapter. There is also an argument to be made that publicly funded data should also be accessible data, now and for future generations.

2.4 Maps, Data, Technological Systems and Infrastructures Leave Historical Traces

Maps have been collected in archives and libraries for nearly 2500 years (Ehrenberg 1981, p.23) and today’s digital cartographers and geomatics practitioners may perhaps be unintentionally disrupting that historical practice. Maps, regardless of who created them are “an integral part of the record of a nation's history, and any national archives should include a rich cartographic collection” (Kidd 1981-1982, p.4). Maps and atlases picture a time, they are part of our memory, give and provide evidence of the past while also creating continuity and a sense of belonging. Their role as artifacts is also to “inform and reform” as Churchill reminds us in his historical work on their influence in shaping urban Chicago (2004, p.13). Maps are socially shaped and in turn are also social shapers. They “often are made not on the basis of the territory itself but on some preconceived sense or vision of the territory. Informed by these maps, subsequent actions move the territory toward the vision” (Churchill 2004, p.11) eventually changing the maps themselves. Those visions were consolidated in maps which in turn molded our physical world - today’s cities, boundaries, the nation. They illustrate how we thought, provide an argument with a distinct point of view, and these views provide a context to understand or read representations of the social construct of places. The data used to render those maps are no less important: "legacy data, archived data or data used for independent research is very valuable: valuable in terms of absolute cost of collection but more importantly, as a resource for others to build upon" (Wilson and O'Neil 2009, p.2).

Historical and contemporary maps inform decisions. Newsworthy, cataclysmic and obvious examples are the information resources accumulated to respond to disasters. For example, the information gathered in response to Hurricane Katrina, including geospatial data and maps, is being preserved by the US Federal Emergency Management Agency and the value of this information is recognized: "one of the best resources we have for preparing for the next major event are the lessons and data accumulated from this catastrophic experience. If we do not preserve this data and use it for research purposes, then we have wasted time and energy and done a great disservice to those who will be affected by the next major hurricane" (Curtis et al., 2006a in Warren Mills et al., 2008, p.477). FEMA is not a national mapping organization, nonetheless it requires specialized data and maps for its ongoing work in emergency preparedness, during an emergency and post emergency work.

Maps, beyond being contributions to intellectual material for historians and content for decision makers are intractably tied to the innovative technologies that have created them (Kinniburgh 1981, p.91). They are social-technological systems (Hughes 2004) that represent the evolution of map making and "effort should be made to preserve the apparatus of cartography: the tools and the machinery" (Kinniburgh 1981, p.95) or in present-day terms : code, software, metadata models and systems. Mapping and data related technologies have changed the way we do things and "while datasets can be selected for the important data they hold reflecting government policy and administration, they also represent interesting innovations, either technological or organizational…Computer systems that changed what was possible, rather than just re-implemented manual processes, are of great historical interest" (Sleeman 2004, p.183) and can be learning tools for the future. Today's programmers, especially those involved in the development of open source technologies, are continuously building on previous technological innovations, and there is merit in preserving those. Particularly since these can enable archivists to view today's content tomorrow. Based on the experience of the authors, many archivists are already thinking in this way. They argue that this requires keeping maps and data "in the information systems in which they were created". The 'records continuum' concept is built upon this approach. The basic idea is that records can function both actively in the organization in which they were created and passively as part of an archive" (Doorn and Tjalsma 2007, p.9).

2.5 Maps and Archives

Maps, data, technologies and their related infrastructures, "are a product of society's need for information, and the abundance and circulation of documents reflects the importance placed on information in society. They are the basis for and validation of the stories we tell ourselves, the story-telling narratives that give cohesion and meaning to individuals, groups, and societies" (Schwartz and Cook 2002, p.13). Accordingly, the function of an archive in a society "must deal with two intimately related, but separately conceived themes: 'knowledge and the shaping of archives' and 'archives and the shaping of knowledge" (Schwartz and Cook 2002, p.14). Contemporary cartographers should reflect on their role in ensuring their content is preserved in the archive and subsequently shaping future knowledge while also working with archivists to transform the archive so that it can actually ingest their artifacts. This is not to be taken lightly as "memory, like history, is rooted in archives. Without archives, memory falters, knowledge of accomplishments fades, pride in a shared past dissipates. Archives counter these losses. Archives contain the evidence of what went before. This is particularly germane in the modern world" (Schwartz and Cook, 2002, p.18). The following section provides an account of a 'rescue' process that recovered an important Canadian geospatial dataset. This account provides some insight into many of the issues related to heritage preservation: lack of clear policy and stewards; technical and methodological obsolescence; time and effort required for data rescue. Given the historical and practical importance of the Canada Land Inventory data, this account also highlights that valuable heritage may be at risk.

2.6 The Rescue and Salvage of the Canada Land Inventory

The Canada Land Inventory (CLI) story demonstrates the importance for cartographers of considering preservation as they are creating maps, compiling data and programming systems.

If our cartographic digital heritage is to be preserved it is important that cartographers consider archiving and preservation as an integral part of the life cycle of the creation of new digital geospatial products and maps. The CLI was part of the Canadian Geographic Information System (CGIS) developed in 1963 at the Department of the Environment in Canada. The CGIS was in fact the world's first GIS. It was a revolutionary idea at the time and "one of the principal driving forces behind the Canadian Geographic Information System (CGIS) was the idea that the CLI maps could be interpreted and analyzed in a myriad of ways if the information could be manipulated by computers" (Schut 2000). It was established "as a joint federal-provincial project to guide the development of policy on the control and management of land-based resources" (Ahlgren and McDonald 1981, p.61). The CGIS ultimately grew to contain the equivalent of thousands of maps and unknowingly became a technology that "spawned an industry that today is worth billions of dollars" (Schut 2000). To demonstrate how innovative CGIS was, Library and Archives Canada (then known as the Public Archives) did not hire their first computer systems specialist with the "responsibility to develop the automation requirements for a National Map Collection intellectual and physical control system" until 1977-78 (Kidd 1981-1982, p.17).

The CLI was an incredibly ambitious federal - provincial program that mapped 2.6 million square kilometers of Canada an "the original cost of the program was in the order of 100's of millions of dollars in the 1970's" (Wilson and O'Neil 2009, p.6). The CGIS was both a set of electronic maps and the "computer programs that allowed users to input, manipulate, analyze, and output those maps" (Schut 2000).

By the late 1980s, the CGIS was no longer being used. Priorities changed, people retired, and institutional memory was being lost. Numerous boxes of tapes, and racks of documentation were left behind with only a few computers left in Ottawa that were capable of reading 9 track tapes let alone run the programs (Schut 2000). In 1995, an informal trans-organizational group of individuals from Statistics Canada, the National Atlas of Canada, Archives Canada and a private sector programmer came together to restore the CLI. They knew this was a valuable heritage dataset, needed some of the layers and they had the skills, know how and more importantly the will to restore the CLI. All had either formerly worked on the CGIS or had an interest in its preservation. Their work consisted of converting into a modern coordinate system a technology that encoded each point as a relative offset (distance and direction) from the previous point, that did not use discrete tiles, and where the entire country of Canada was coded as one enormous database (Schut 2000). On June 18, 1998, the agriculturally relevant portions of the CLI were handed over on one CD and it worked flawlessly on the analytical tools built in anticipation of the new format (Schut 2000). The data were eventually distributed on Natural Resources Canada's GeoGratis site for free with documentation and some text to help interpret the content. The CLI "rapidly became their most popular product" (Schut 2000). An updated and the original versions are both also available through the CanSIS website. Not all of the CLI data was saved and the cost of the effort described above was very substantial.

The CLI demonstrates the archival adage that "where the information and form of the record are so tenuously related, archivists must appraise, acquire, preserve, and control whole systems of information within which various physical media may exist" (Ahlgren and McDonald 1981, p.64).

2.7 Evaluating progress in Preserving Cartographic Heritage

The preceding sections have provided an overview of developments in contemporary cartography, issues related to preserving and archiving the resulting artifacts, and the importance of preserving our geospatial information and cartographic heritage. Unfortunately, the technological, institutional and organizational issues related to the long-term preservation of data remain largely unresolved. The basic digital data upon which we depend to inform decisions on planning, health, emergency preparedness, industrial exploration and research are rarely being effectively archived and preserved and, as a result much is being lost, some permanently. John Roeder, a researcher on both International Research on Permanent Authentic Records in Electronic Systems (InterPARES) projects, discovered that one fifth of the data generated by the 1976 Viking (a space probe) exploration of Mars (Cook 1995 and Harvey 2000), the entire 1960 U.S. Census (Waters and Garret 1996) and the works of nearly half of digital music composers (Longton 2005) and one-quarter of digital photographers (Bushey and Brauen 2005) have been lost or threatened by technological obsolescence or inadequate preservation strategies. It has been argued that "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, Lauriault and Pulsifer 2005). It has also been noted that "indeed digital technology is responsible for much of the loss, as storage technology has given a false sense of security against loss and obsolescence (Strong and Leach 2005, p.13) and "an unprecedented firestorm is incinerating Canada's digital research wealth" (SSHRC 2002).

2.8 How Can Today's Maps, Data and Technologies be Preserved?

Researchers from the Geomatics and Cartographic Research Centre (GCRC), participated in the InterPARES 2 Project precisely to try to answer the question posed in the heading above. InterPARES 2 (IP2) was a research initiative led by the University of British Columbia. The goal of IP2 "was to ensure that the portion of society's recorded memory digitally produced in dynamic, experiential, and interactive systems in the course of artistic, scientific and e-government activities can be created in accurate and reliable form, and maintained and preserved in authentic form, both in the short and the long term, for the use of those who created it and of society at large, regardless of digital technology obsolescence and media fragility" (Duranti 2007, p.115). The GCRC led two IP2 studies in the Science Focus: i) a Case Study about the Cybercartographic Atlas of Antarctica and (Lauriault and Hackett 2005) ii) a General Study examining the preservation practices of scientific data portals (Lauriault and Craig 2007; Lauriault, Craig, Pulsifer & Taylor 2008). The following sections provide a summary of the results of these studies with a focus on the challenges faced, possible strategies for meeting these challenges and research opportunities emerging from the studies.

2.9 Multidisciplinary Archival Research

2.9.1 Case Study 06 (CS06) Cybercartographic Atlas of Antarctica (CAA)

The Cybercartographic Atlas of Antarctica (CAA) research project was designed to contribute to developing the theory and practice of cybercartography and emerging forms of geographic information processing. The first phase of the project (completed fall 2007) resulted in the development of a series of chapters or modules that examine and explore topics of interest to both Atlas users and researchers alike. The project was a collaborative effort developed as a project under the Scientific Committee on Antarctic Research's (SCAR) geographic information program. The model used to develop the CAA includes the use of multimedia cartography and distributed data sources. For more information, the reader is directed to https://gcrc.carleton.ca/confluence/x/XAc.

2.9.1.1 CS06 Research Methodology

The primary information-gathering tool for CS06 was the InterPARES 2 case study questionnaire, comprising 23 questions (InterPARES 2 2003). Two sets of semi-structured interviews at two different development stages of the CAA Project were conducted to answer questions of interest to the archival community (Lauriault and Hackett 2005). Concurrently, these interviews helped GCRC researchers make explicit some implicit, tacit assumptions in the production of the CAA. Reflections on production processes identified both shortcomings and strengths in archival terms.

2.9.1.2 CS06 Observations

The archival research inquiry revealed a number of issues that may prevent effective preservation and archiving. First, much of the data used in the project did not have persistent identifiers. While data records used to create maps and other types of representation often had an identification number or 'primary key', there was no long-term strategy to ensure that this value would not change over time; during a database migration, for example. Lack of a persistent identifier may introduce ambiguity that prevents subsequent users or archivists from effectively establishing data provenance. Secondly, custom software was developed to create the atlas. This software uses a markup language combined with processing algorithms to define and integrate data resources. While this software is open source and uses many standard approaches (e.g. 'XML schema'), the software system as a whole was not comprehensively documented. Thus, while archivists and future generations may be able to understand the components of the atlas, establishing an operational version of the CAA may be difficult.

GCRC researchers realized the importance of detailed documentation including the possibility for creating training courses, and capturing processes in the CAA's online forums and WIKIs. However, dedicating resources to these activities in a research environment that places priority on peer-reviewed publication presents a challenge. And while the primary funding for the project stipulated a requirement for preservation, no guidelines existed for how this was to be done and more importantly no funding was available to do it! To mitigate the impact of limited documentation the researchers used, well documented open standards, established software development methods, a source code versioning system (to document the evolution of the software) and open source licensing (allowing others to contribute to documentation).

In considering source data preservation and archiving, the origin and provenance of the data is an important consideration. The Cybercartographic Atlas of Antarctica (CAA) was endorsed by the Scientific Committee on Antarctic Research (SCAR) the most reputable body in Antarctic science. The Atlas uses data from authoritative and reputable scientific organizations and are accompanied by standards compliant metadata. Data reliability, in archival terms, is therefore assured by the quality of the base data used and by the methods applied by content creators. These data were provided using formats based on open standards and so the risk of effectively 'locking' future generations out of the data is small. Some concerns with respect to using proprietary formats for some multimedia objects remain. While the Nunaliit Framework's has an open source license and it generates high performance, standards-based Web applications, some of the multimedia content in the Atlas is encoded in proprietary formats. While the Nunaliit generated application may outlive the multimedia in terms of preservation, it is recognized that at times practical decisions such as the use of video content compressed with proprietary software may result in information loss.

The CAA project included standard intellectual property concerns, although the terms of the Antarctic Treaty system allow much of the data used in the creation of the CAA to be used at little or no cost. The CAA also follows typical license agreements, use rights to objects and data, and copyright while its software, created at the GCRC, is distributed under an open BSD license. The atlas itself includes caveats and disclaimers (e.g. the CAA is intended for information, not navigation purposes) and the project must adhere to the requirements of the funding agency, research clearances and the Nunaliit License (Nunaliit 2006).

The CAA production process was considered to be adequate to meet the challenge of technological obsolescence. The use of open source software was thought by archivists to make the CAA more sustainable than if proprietary products were being used. If, for example a popular open source software project is discontinued, the source code may still be available for use in preservation and archiving activities. Additionally, in the rare case that popular projects are discontinued, emerging communities typically develop backwards compatibility that support access to legacy data. Concurrently, complete and available documentation of proprietary formats is also considered important. The use of XML (an open standard) for the content modules should make the CAA easily translatable (via new compilers) into any future markup languages. The CAA also adheres to other open standards such as the OGC interoperability specifications (2006) and the International Standards Organization 19115 Geomatics Standards (2003). Although a strong foundation for preservation of the CAA exists, effort is required to 'package' these elements in a way that would promote preservation and archiving. While the components of the CAA may be suitable for archiving, a 'map' of the project as a whole does not yet exist. Such a map is required to document the various component of the CAA and the relationships between and among these components.

Working with archivists was beneficial to the CAA production process in terms of considering preservation at the point of creation. Preservation issues were considered early in the development process, thus reducing cost and disruptions in development. The case study revealed that the CAA development processes were adequate for preservation and archival purposes. The focus on interoperability, adherence to open source standards, documentation through metadata creation, use of professional software development practices, establishment of data quality standards and are all strong features of the atlas in terms of its suitability for preservation and archiv-

ing. See Pulsifer et al., (2005, 2008) for additional theoretical and technical details related to the production of the CAA).
The InterPARES 2 project also benefited by gaining an increased understanding of cybercartography and collaborative science, practices and processes. InterPARES 2 researchers also learned that the fields of geomatics and producers of scientific data have very rigorous metadata descriptions, excellent standards, and professional data gathering and maintenance procedures that can be used as models for records created in the arts and in e-government which are other IP2 focus priorities.

2.9.1.3 CAA Preservation Challenges

The greatest challenges limiting the long-term preservation of the CAA are neither technological nor procedural. The greatest roadblock is simply the fact that no Canadian archival institution is currently in a position to ingest the CAA. This is a major problem not only for the Atlas but for the preservation of similar digital products in Canada. The GCRC is having ongoing discussions with members of the Data Library at Carleton University to attempt archive the CAA as required by its funder. The Data Library potentially has both the technological, policy and human resource capacity to archive the CAA but not the technology nor the mandate to do so. Discussions are ongoing with Library and Archives Canada (LAC), the National Research Council and GeoConnections. The CAA project is now complete (CAA 2009), and to date there is no explicit transfer plan in place. The GCRC is fortunate to be located in the Nation's Capital as it has ready access to Canada's top officials in key organizations that can assist with resolving this problem, but alas no obvious solutions have presented themselves. The GCRC waits for the creation of an Institutional Repository (IR) at Carleton University that can ingests more than text based research material, or the creation of a cartographic and geospatial data Trusted Digital Repository, a data archive and/or for LAC to develop the capabilities and mandate to ingest the output of publicly funded research on complex digital mapping in Canada.

2.9.2 General Study 10 (GS10) Preservation Practices of Scientific Data Portals

Geospatial and science data are increasingly being discovered and accessed in data portals (i.e., data repositories, clearinghouses, catalogues, archives, geo-libraries and directories). In this context, a portal can be defined as a user interface that acts as a starting point for finding and accessing geospatial and scientific data. Portals can provide all or some of the following services: search and retrieval of data, item descriptions, display services, data processing, the platform to share models and simulations, and the collection and maintenance of data. Much but not all of the data derived from portals are raw in nature and require the user to interpret, analyze and/or manipulate them. The reasons for their creation are one-stop-shopping, distributed responsibility over data sets, discoverability, and reduction in cost as data are stored once and used many times (Lauriault 2003). Data portals are the technical embodiments of data-sharing policies. Individuals within organizations, research projects, or scientific collaborations register their data holdings in the portal via an online form organized according to a metadata standard, and then choose to make their data available for free, sale, viewing or downloading (Lauriault 2003). Metadata standards "establish the terms and definitions to provide a consistent means to describe the quality and characteristics of geospatial data" (Tosta and Michael Domaratz 1997, p.22) and the ISO 19115 metadata (ISO 2003) standard has become an international standard in the field of geomatics. Thus, portals and the data resources that they connect, can be seen as a collective geospatial information artifacts. The GS10 study examined portals to reveal issues related to preservation and archiving of portals.

2.9.2.1 GS10 Research Methodology

The GS10 study included an extensive literature review of publications from national and international scientific organizations, government and research funding bodies and empirical evidence from a selection of IP2 Case Studies and 32 scientific Data Portals most of which included geospatial data (Lauriault and Craig 2007; Lauriault, Craig, Pulsifer & Taylor 2008). A GS10 Survey was undertaken to collect information about the actual practices, standards, and protocols (Lauriault and Craig 2007).

2.9.2.2 GS10 Observations

The portals selected pertained to different communities of practice in geomatics and other sciences that are thematically heterogeneous, and each adheres to that community's specific methodologies, tool, technologies, practices and norms. As expected, portals are rich repositories of data and information that serve the needs of many types of users. The architecture of data portals varies: some are a single enterprise sponsored portal (like a national library); a network of enterprises (like a federation of libraries) or a loose network connected by protocols (like the Web) (NRC 1999). Distributed data portals have datasets described according to a given standard, and when a request is sent to them by a given site a search is executed by a search agent to access or render the data into a map or some other form GRID[1] portals. Those using use Web Map Services are an example of these. A Collection level catalog/portal identifies a data custodian's holdings and uses them to direct searches (e.g. Z39.50, Ocean Biogeographic Information System – Spatial Ecological Analysis of Megavertebrates Populations). A unified catalogue exists in one place: data custodians submit metadata for each data set to a central site which makes them available for searching, and the record directs the user to the data set (e.g. GeoConnections Discovery Portal). Digital collections/portals can be housed in a single physical location (e.g. Statistics Canada), and they may be virtual (e.g. Earth Systems GRID), housed in a set of physical locations and linked electronically to create a single, coherent collection (e.g. Global Change Master Directory, International Comprehensive Ocean Atmospheric Dataset). The distinction between centralized, distributed or unified portals may have funding, policy and preservation implications. Data collections may also differ because of the unique policies, goals, and structure of the funding agencies.

There are three functional data collections/portal categories: research data collections; resource or community data collections; and reference data collections (AIP 2007). These are not rigid categories. Research data collections portals contain the results of one or more focused research projects and data that are subject to limited processing. Data types are specialized and may or may not conform to community standards, adhere to metadata standards, or to content access policies. Data collections vary in size but are intended to serve a specific scientific group, often limited to immediate participants. These collections are supported by relatively small budgets, often through research grants funding a specific project, and therefore do not have preservation as a priority (e.g., Indiana University Bio Archive, National Virtual Observatory (NVO)). Resource or community data collections serve a single science, geomatics or engineering community.

1 Grid computing refers to the automated sharing and coordination of the collective processing power of many widely scattered, robust computers that are not normally centrally controlled, and that are subject to open standards.

These digital collections are often large enough to establish community-level standards, either by selecting from among pre-existing standards or by bringing the community together to develop new standards where they are absent or inadequate. The CanCore Learnware metadata standard is an example of this type of community standard.

The budgets for resource or community data collections are moderate and often supported by a government agency. Preservation is contingent on departmental or agency priorities and budgets (e.g. Canadian Institute for Health Information (CIHI), Southern California Earthquake Center (SCEC), National Geophysical Data Center (NGDC - NOAA)). Reference data collections are intended to serve large segments of the scientific, geomatics and education community. These digital collections are broad in scope; serve diverse user communities including scientists, students, policy makers, and educators from many disciplines, institutions, and geographical settings. Normally they have well-established and comprehensive standards which often become either de jure or de facto standards, such as the Geomatics ISO 19115 Metadata standards. Budgets supporting these are often large and come from multiple sources in the form of direct, long-term support; and the expectation is that these collections will be maintained indefinitely (e.g. Canadian Geospatial Data Infrastructure (CGDI), Global Change Master Directory – Global Change Data Center)

2.9.2.3 GS10 Conclusions

There are three types of issues relating to portals and data quality: i) those related to the portal's operation and its design, management, and long-term viability; ii) those related to the accuracy of the individual datum and data sets; iii) and those related to the relationship between the portal, its data and services, and the individual or corporate user – essentially those issues that emerge from a history of interaction that builds trust and comfort with the user. The issues that are related to the portal itself are those that are linked to maintaining an authentic memory, especially of the sources of the data, their management or changes over time, and their connections to contributors or sources. Building sites and services that continue to be what they purport to be, and whose changes and transitions over time are visible and knowable to a user build conditions of trust. The InterPARES 1 project developed benchmarks that could be used by portals to ensure that their data continue to be authentic over time.

Science and geomatics are heterogeneous domains, and each field and subfield has its own culture, methods, quality measures and ways of explaining what they do. Formal ontologies are an emerging method used to help mediate the myriad metadata standards and facilitate the production of meaningful ways to represent the world and preserve the data. Data portals reflect the policies, funding agencies and the technologies chosen by the organizations that create and manage them. Organizational, technological, metadata and data quality considerations aspects affect appraisal decisions and provide challenges for archivists.

Science is a collaborative endeavour that is premised on the notion of knowledge sharing, dissemination, reproducibility, verification, and the possibility that new methods will yield new results from old data. Therefore, there is an argument to be made that publicly funded collections of data should be made available to the citizens who paid for them and they should be made available to future generations for the advancement of knowledge.

The IP2 research showed that interoperability is a problem with the rapidly increasing number of digital data bases that need to interact if the challenges knowledge integration are to be met. The Cybercartographic Atlas of Antarctica was faced with the challenge of using information from different databases in different countries and, in order to do this, adopted an open source and open standards approach using OGC specifications. This decision was taken primarily for production reasons but has had beneficial effects in archiving and preservation terms as it helps overcome the problem of technological obsolescence. The IP2 Case studies demonstrate that a lack of interoperability can lead to having data that cannot be archived in the same form the creator had intended. Indeed, it can be argued that interoperability is a key element in archiving all digital data and that an open source standards and specifications approach should be a major facet of any archival strategy.

For scientific and geomatics disciplines, trust will continue to rest on specific norms of scientific work. Trusted repositories, whose data are kept reliable, accurate, and authentic over time, will need to be established, managed, and funded on a continuing basis. The problems are on three levels: organizational stability, data and metadata management processes, and technological hand shaking across generations.

Established archival repositories that are mandated (and funded) to guarantee the continuing availability of scientific data records and information that support administrative, legal, and historical research are needed. Although there are digital repositories for social science data, true digital scientific data archives are few and far between. The IP2 General Study on data portals demonstrated that there are numerous excellent initiatives in place to make data discoverable and accessible. However, few of these data portals archive their data. The few portals that are government funded in the US and simultaneously housed in government departments do have preservation as a mandate or are considered to be government archives, but most portals do not have this type of financial or institutional stability. At risk in particular are the repositories that are distributed and leave issues of data quality to the data custodians or creators. Therefore, much government funded science is not enveloped in any data preservation or archiving processes. This is quite troubling, considering the investment tax payers have made in these endeavours, let alone the loss in knowledge dissemination and building opportunities

All stakeholders, including the scientists who create the information, research managers, major user groups and of course the archivists, should be involved in the appraisal decisions on what is to be archived and by whom. This appraisal should be an ongoing process from the point of creation and is best carried out in a project specific fashion, in collaboration with those most knowledgeable about the data. It is recommended that archivists build on existing data portals and extend these activities with archival policies, techniques and technologies. These data have already been appraised as being worthy or else they would not be in the portals. Also, portal creators and maintainers need to seriously consider adding preservation as part of their mandate, since it is highly unlikely in the immediate future that an archive would be able to ingest these holdings. There is merit in having data preserved at their source where that is feasible.

2.10 What is being done?

There are some promising international initiatives particularly in the European Union (e.g. Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR), Digital Repository Infrastructure Vision for European Research (DRIVER), UK Data Archive (UKDA) and the Data Archiving and Networked Services (DANS) in the Netherlands)) and in the US (e.g. The Cyberinfrastructure Project and National Geospatial Digital Archive), and a number of thematic initiatives (e.g., Sierra Nevada ecosystem Digital Spatial Data Archive, European Digital

Archive of Soil Maps, LSU Geographic Information Systems GIS Clearinghouse Cooperative (LGCC)). And as previously discussed there are a number of portals, data archives and GRID computing systems as part of the GS10 Portal Study. The Open Geospatial Consortium (OGC) Data Preservation Working Group was created in December 2006 to address technical and institutional challenges posed by data preservation, to interface with other OGC working groups, which address technical areas that are affected by the data preservation problem, and to engage in outreach and communication with the preservation and archival information community. This is a very promising initiative, as the OGC is dedicated to interoperability, open standards and open specifications that help overcome many of the issues of platform dependency. The OGC has also done excellent work on the production of the de facto standards of Internet mapping internationally, and this working group is dedicated to developing prototypes and testbeds with software vendors. The CODATA Working Group on Archiving Scientific Data has been holding symposia and workshops on the topic, and the Canadian National Committee for CODATA has been active in documenting and reporting scientific data activities. The Preserving Access to Digital Information (PADI) initiative based in Australia provides excellent practical resources for cartographers and data producers who wish to gain practical information on how to go about the preservation of their resources. The US, the Earth Institute at the Columbia University portal for Geospatial Electronic Records also includes a number of recommendations regarding the management and preservation of geospatial data. Finally, the International Council for Scientific and Technical Information (ICSTI) annual conference Managing Data for Science will be hosted in Ottawa in June 2009 and will focus on issues of data access and preservation.

In Canada there is much discussion but to date very little concrete action. GeoConnections is the Government of Canada agency mandated to create the Canadian Geospatial Data Infrastructure (CGDI). GeoConnections conducted a study on Archiving, Management and Preservation of Geospatial Data which provided a well rounded analysis of preservation issues in the field of cartography such as: technological obsolescence; formats; storage technologies; temporal management; and metadata. The Study also provides a list of technological preservation solutions with their associated advantages and disadvantages, and a list of proposed institutional and national actions. A number of studies, reports and committees have made high level recommendations and provided strategies for improving the archiving of digital data in Canada, as they all recognize the poor state of Canada's digital data resources. The SSHRC National Data Archive Consultation Report discussed the preservation of data created in the course of publicly funded research projects and identified important institutions, infrastructures, management frameworks and data creators and called for the creation of a national research data archive. The report Toward a National Digital Information Strategy: Mapping the Current Situation in Canada indicates that "the stewardship of digital information produced in Canada is disparate and uncoordinated" and "the area of digital preservation, which involves extremely complex processes at both the organizational and technical levels, comprehensive strategies are not yet being employed. Many feel that much of the digital information being created today will be lost forever." The Final Report of the National Consultation on Access to Scientific Data, developed in partnership with the National Research Council Canada (NRC), the Canada Foundation for Innovation (CFI), Canadian Institutes of Health Research and NSERC, expressed concern about "the loss of data, both as national assets and definitive longitudinal baselines for the measurement of changes overtime." This report also provides a comprehensive list of recommendations that include ethics, copyright, human resources and education, reward structures and resources, toward the creation of a national digital data strategy and archive. In December 2006, Library and Archives Canada hosted a National Summit on a Canadian Digital Information Strategy. The challenges of the new Web 2.0 social computing environment, open access, interoperability and licensing among numerous other topics were discussed.

While progress is being made and many discussions are in progress, the implementation of concrete solutions lags far behind the rhetoric The problem is not confined to Canada. Few nations are developing comprehensive digital data (I.e., science, research and geospatial) strategies, let alone preservation strategies. National mapping organizations (NMOs) and governmental geospatial data producers have a head start over the private, academic and not for profit sectors since there are often archival accessioning rules in places. There is no guarantee, however, as seen with the CLI example, that governments will preserve these artifacts, nevertheless there is a framework and some resources to take action are available. Research cartographers and data producers, may be able to rely on institutional repositories (IR) providing of course they exist in their home countries. Such repositories must however be able to ingest more than text based research and the custodians must be willing and have the resources (i.e., technical, skill, financial and mandate) to carry out this function.. An IR is “is a specific kind of digital repository for collecting, preserving, and disseminating -- in digital form -- the intellectual output of an institution, particularly a research institution” (Glick 2009). The Registry of Open Access Repositories keeps a list of open access IRs (ROAR 2007). IRs are however not archives. An “archive is normally understood as being a trusted steward and repository. but note, it is not just for objects that are digital - it may acquire these and many are actually doing so - but its status as an 'archive' and the trust reposed in its work is largely related to its mission, organizational transparency, and of course, that niggling issue of long-term viability” (Craig 2009). Trusted digital repositories (TDRs) are works in progress, and none have yet been certified, these are IRs and archives which could be seen as “an archives of digital objects only” (Craig 2009). A TDR is any kind of digital repository that meets the requirements of the Trusted Repository Audit Checklist (Centre for Research Libraries 2008). The audit checklist “brings together existing best practice and thought about the organizational and technical infrastructure required to be considered trustworthy and capable of certification as trustworthy. It establishes a baseline definition of a trustworthy digital repository and lays out the components that must be considered and evaluated as a part of that determination” (Glick 2009). Additional information can be found in the Records Library Group (RLG) and the Online Computer Library Center (OCLC) Trusted Digital Repositories: Attributes and Responsibilities (OCLC 2002) or the RLG and the US National Archives and Records Administration (NARA) Audit Checklist for Certifying Digital Repositories (2005). Finally, around the world there are some map libraries such as “McGlamery’s networked Map and Geographic Information Center at the University of Connecticut, that ingest digital maps, and cartographers and geospatial data producers are encouraged to spearhead initiatives of their own and begin a conversation with digital map librarians, archivists, content creators and the curators of IRs. Such an approach will be challenging and will require vision and leadership, but it is better than the current state of affairs where our map, data and technological system heritage is disappearing and leaving a large 'blank spot in history' not unlike the one found in the Jedi Archive by Obi-Wan Kenobi (Ketelaar 2002)!

As Information and Communication Technology evolves and new forms of information exchange and computing emerge, new challenges for preservation and archiving will materialize. Of current interest are the implications for using distributed systems such as the networks established by

Spatial Data Infrastructure programs, GRID computing systems, and a more recent development; cloud computing. Cloud computing is a nebulous term used to describe any number of distributed models being established by mainstream industry (as opposed to GRID computing in the science). Some key issues and concerns in relation to GRID and Cloud computing are in establishing information provenance in a highly distributed environment ii) establishing a discrete archival record that can be captured managed over time. Doorn and Tjalsma identify the challenges faced:
"A descriptive way to explain computational grids is by analogy with the electric power grid. The latter provides us with instant access to power, which we use in many different ways without any thought as to the source of that power. A computational grid is expected to function in a similar manner. The end user will have no knowledge of what resource they used to process their data and, in some cases, will not know where the actual data came from. Their only interest will be in the results they can obtain by using the resource. Today, computational grids are being created to provide accessible, dependable, consistent and affordable access to high performance computers and to databases and people across the world. It is anticipated that these new grids will become as influential and pervasive as their electrical counterpart."(Doorn and Tjalsma 2007, p.16)
While these trends present challenges, they simultaneously provide research opportunities. Given the nature of the issue, it is clear that solutions will require multi and interdisciplinary collaboration that can address the technical, cartographic, preservation, archival and larger social issues implicated in preserving these new information and knowledge phenomena.

2.11 Conclusion

If we acknowledge that "remembering (or re-creating) the past through historical research in archival records is not simply the retrieval of stored information, but the putting together of a claim about past states of affairs by means of a framework of shared cultural understanding" (Schwartz and Cook 2002, p.3), then as cartographers and geospatial data producers we need to assure that our artifacts inform that cultural framework. As has been discussed in this chapter, much has been lost. Also, many are beginning to take seriously the fact that "archivists no longer have the luxury of waiting for thirty years to make appraisal decisions. Selection has to be made very near to, if not at, the time the record is created" (Sleeman 2004, p.180). In other words creators will need to work collaboratively with archivists, librarians, technology specialists to design cartographic artifacts that will stand the test of time and build them accordingly (Doorn and Tjalsma 2007; Kinniburgh 1981; Sleeman 2004; Schut 2000; Wilson and O'Neil 2009; Ahlgren and McDonald 1981). The development process of Cybercartographic atlases such as the Antarctic Atlas exemplify this practice. On the positive side existing efforts can be built upon, such as existing science and geospatial data portals where appraisal, cataloguing and issues of data quality have already been addressed. The next step will be to transform these into trusted digital data repositories. Concurrently, we need national digital strategies to indentify "future research needs and the establishment of mechanisms that allow stakeholders to consider the potential gains from cooperation in planning the data resources required to meet these needs" (Doorn and Tjalsma 2007, p.13). Finally, "records are not only a reflection of realities as perceived by the "archiver". They constitute these realities. And they exclude other realities" (Ketelaar 2002, pp.222-223). The map is not the territory, it creates a record of the territory and it occasions it. It is not just a recording: it constitutes the event. Fortunately, many of the 20th century's digital cartographers are still alive and therefore can be a part of the preservation process and can discuss the history of the mapmaking process of their digital artifacts, and they could also be representatives of technology that created these (Kinniburgh 1981). The solution of the problems of the preservation and access to the remarkable explosion of digital cartography and the emerging variety and volume of Geospatial products is one of the greatest challenges of the 21st century. It is hoped that this book will make a significant contribution by cartographers to meeting those challenges.

2.12 Acknowledgements

Much of the content of this chapter is the result of two Canada Social Sciences and Humanities Research Council (SSHRC) funded research projects, InterPARES 2 at the University of British Columbia and Cybercartography and the New Economy at Carleton University.

2.13 References

Agriculture and Agri-Food Canada. 2009. Canadian Soil Information System (CanSIS). [Online] (Updated 27 November 2008) Available at: http://sis-.agr.gc.ca/cansis/ [Accessed 1 April 2009].

Ahlgren, Dorothy & McDonald, John. 1981, The Archival Management of a Geographic In formation System, Archivaria 13, pp. 59-65.

American Institute of Physics (AIP). 2001. AIP Study of Multi-Institutional Collaborations: Final Report. Highlights and Project Documentations [Online] Available at: http://www.aip.org/history/publications.html [Accessed 17 August 2007].

Ballaux, Bart. 2005. CS26 Most Satellite Mission: Preservation of Space Telescope Data Report. Vancouver: InterPARES 2.

Bushey, Jessica & Braun, Marta. 2005. Survey of Record-Keeping Practices of Photographers Using Digital Technology Final Report. Vancouver: InterPARES 2.

Caquard, Sebastien. 2009. Cybercartographic Atlas of Canadian Cinema. [Online] Available at: http://www.atlascine.org/iWeb/Site/atlasen.html [Accessed 1 April 2009]

Centre for Reseach Libraries. 2008. Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist. [Online] Available at http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91 [Accessed April 1 2009].

Churchill, Robert R. 2004. Urban Cartography and the Mapping of Chicago, The Geographical Review, 94(1), pp.1-22.

CODATA. 2009. Canadian National Committee for CODATA. [Online] Available at: http://www.codata.org/canada/ [Accessed April 1 2009].

CODATA. 2009. Preservation of and Access to Scientific and Technical Data in Developing Countries. [Online] Available at: http://www.codata.org/taskgroups/TGpreservation/index.html [Accessed April 1 2009].

Cook, Terry. 1995. It's Ten O'clock, Do You Know Where Your Data Are? Technology Review. 98, pp.48-53

Craig, Barbara. .barbara.craig@utoronto.ca 2009. IR or TDR vs Archive. [E-Mail]. Message to Tracey P. Lauriault (tlauriau@gmail.com). Sent March 3 2009 10:50 AM.

Data Archiving and Networked Services (DANS). 2009. Home Page. [Online] (Updated March 23 2009) Available at: http://www.dans.knaw.nl/en/ [Accessed April 1 2009].

Doorn, Peter and Heiko Tjalsma. 2007. Introduction: Archiving Research Data. Archival Science. 7, pp.1–20.

DRIVER (Digital Repository Infrastructure Vision for European Research). 2009. Digital Repository Infrastructure Vision for European Research Home Page.

[Online] (Udated 2009) Available at: http://www.driver-repository.eu/ [Accessed April 1 2009].

Duranti, Luciana. 2007. Reflections on InterPARES. The InterPARES 2 Project (2002-2007): An Overview. Archivaria 64, pp.113-121.

Earth Institute at Columbia University. 2005. Center for International Earth Science Information Network (CIESIN).Geospatial Electronic Records web site [Online] Available at: http://www.ciesin.columbia.edu/ger/ [Accessed April 1 2009].

Ehrenberg, Ralph E. 1981. Administration of Cartographic Materials in the Library of Congress and National Archives of the United States. Archivaria 13, pp.23-29.

Environmental Protection Agency (EPA), 2009, Environmental Information Exchange Network & Grant Program Glossary, [Online] (Updated September 25th, 2007) Available at: http://www.epa.gov/Networkg/glossary.html#G [Accessed 28 March 2009].

European archive on the soil maps of the World (EuDASM). 2009. Overview of EuDASM. [Online] Available at: http://eusoils.jrc.ec.europa.eu/esdb_archive/EuDASM/EUDASM.htm [Accessed April 1 2009].

European Commission - Sixth Framework Programme. 2009. Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR) User Community! [Online] (Updated 2006) Available at: http://www.casparpreserves.eu/ [Accessed April 1 2009].

GeoConnections. 2005. Archiving, Management and Preservation of Geospatial Data. [Internet] Ottawa: Government of Canada (Published 2005) Available at http://www.geoconnections.org/publications/policyDocs/keyDocs/geospatial_data_mgt_summary_report_20050208_E.pdf [Accessed 17 January 2007].

Geomatics and Cartographic Research Centre (GCRC). 2009. Living Cybercartographic Atlas of Indigenous Perspectives and Knowledge. [Online] Available at: https://gcrc.carleton.ca/confluence/display/GCRCWEB/Living+Cybercartographic+Atlas+of+Indigenous+Perspectives+and+Knowledge [Accessed 1 April 2009].

Glick, Kevin. kevin.glick@yale.edu. 2009. IR or TDR vs Archive. [E-mail]. Message to Tracey P. Lauriault (tlauriau@gmail.com). Sent March 3 2009 5:54 PM.

Gupta, A., B. Ludascher, M.E. Martone. 2002. Registering Scientific Information Sources for Semantic Mediation. Lecture Notes in Computer Science. no. 2503, pp.182-198.

Harvey, Ross. 2000. An Amnesiac Society? Keeping Digital Data for Use in the Future. LIANZA 2000 Conference, New Zealand.

Hughes, Thomas P. 2004. Human Built World: How to Think About Technology and Culture. Chicago: The University of Chicago Press.

International Council for Scientific and Technical Information (ICSTI). 2009. 2009 Conference Managing Scientific Data [Online] Available at: http://www.icsti2009.org/02-program_e.shtml [Accessed April 1 2009].

International Standards Organization (ISO). 2003. Geographic information – Metadata ISO 19115. [Online] Available at: http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=26020 [Accessed 10 January 2006].

InterPares 1. 2002. Authenticity Task Force Report. (InterPARES 1) Vancouver [Internet] Available at http://interpares.org/display_file.cfm?doc=ip1_atf_report.pdf [Accessed 1 April 2009].

InterPARES 2. 2003. 23 Case Study Questions that the researchers should be able to answer at the completion of their investigation, [Online] Available at: http://interpares.org/display_file.cfm?doc=ip2_23_questions.pdf [Accessed September 2006].

InterPARES 2. 2006. International Research on Permanent Authentic Records in Electronic Systems Glossary. [Online] Available at: http://interpares.org/ip2/display_file.cfm?doc=ip2_glossary.pdf&CFID=137165&CFTOKEN=15668692 [Accessed September 2006]

InterPARES 2. 2009. International Research on Permanent Authentic Records in Electronic Systems Internet Site. [Online] Available at: from http://interpares.org/ [Accessed 2 April 2009].

Ketelaar, Eric. 2002. Archival Temples, Archival Prisons: Modes of Power and Protection. Archival Science 2, pp.221-238.

Kidd, Betty. 1981-1982. A Brief History of the National Map Collection at the Public Archives of Canada. Archivaria 13, pp.3-22.

Kinniburgh, Ian A.G. 1981. The Cartography of the Recent Past. Archivaria 13, pp.91-97.

Lauriault, Tracey P. 2003. A Geospatial Data Infrastructure is an Infrastructure for Sustainable Development in East Timor,, Master's Thesis, Ottawa: Carleton University..

Lauriault, Tracey P. & Craig, Barbara. 2007. GS10 - Preservation Practices of Scientific Data Portals. [Online] Available at: http://interpares.org/ip2/ip2_case_studies.cfm?study=34 [Accessed 1 April 2009].

Lauriault, Tracey P. & Hackett, Yvette. 2005, CS06 Cybercartographic Atlas of Antarctica Case Study Final Report. [Online] Available at: http://interpares.org/ip2/ip2_case_studies.cfm?study=5 [accessed October 2006].

Lauriault, Tracey P.; Caquard, Sebastien; Homuth, Christine & Taylor, D. R. Fraser. 2009. Pilot Atlas of the Risk of Homelessness. [Online] Available at: https://gcrc.carleton.ca/confluence/display/GCRCWEB/Pilot+Atlas+of+the+Risk+of+Homelessness [Accessed 1 April 2009].

Lauriault, Tracey P.; Craig, Barbara; Pulsifer, Peter. L. & Taylor, D. R. Fraser. 2008. Today's Data are Part of Tomorrow's Research: Archival Issues in the Sciences. Archivaria 64, pp.123-181.

Library and Archives Canada (LAC). 2007. Toward a Canadian Digital Information Strategy: National Summit (Library and Archives Canada) [Internet] Available at: http://www.collectionscanada.ca/cdis/012033-601-e.html [Accessed 27 January 2007].

Library of Congress. 2009. Library of Congress Flickr pilot. [Online] Available at: http://www.flickr.com/photos/library_of_congress/collections/72157601355524315/ [Accessed 1 April 2009]

Longton, John. 2005. Record Keeping Practices of Composers Survey Report. [Online] Available at: http://www.interpares.org/ip2/ip2_general_studies.cfm?study=27 [Accessed 23 August 2007).

MacDonald, John & Shearer, Kathleen. 2005. Toward a National Digital Information Strategy: Mapping the Current Situation in Canada. (Ottawa: Library and Archives Canada) [Internet] Available at: http://www.collectionscanada.gc.ca/obj/012033/f2/012033-300-e.pdf [Accessed 1 April 2009).

National Digital Information Infrastructure and Preservation Program (NDIIPP). 2007. Digital Preservation: The National Digital Information Infrastructure and Preservation Program. [Online] Available at: http://www.digitalpreservation.gov/index.html [Accessed 27 January 2007].

National Geospatial Data Archive (NGDA). 2009. Home Page. [Online].(Updated March 2 2009) Available at: http://www.ngda.org/ [Accessed April 1 2009].

National Library of Australia. 2009. Preserving Access to Digital Information (PADI). [Online] Available at: http://www.nla.gov.au/padi/about.html [Accessed April 1 2009].

National Research Council (NRC). 1995. Commission on Physical Sciences, Mathematics, and Applications, Preserving Scientific Data on Our Physical Universe: A New Strategy for Archiving the Nation's Scientific Information Resources [Online] Available a: http://www.nap.edu/catalog.php?record_id=4871 [Accessed 23 August 2007].

National Research Council (NRC). 1999. Spatial Information Resources, Distributed Geolibraries. [Online] Available at: http://www.nap.edu/openbook.php?isbn=0309065402 [Accessed 23 August 2007].

National Science Foundation (NSF). 2005. Report of the National Science Board: Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century. [Online] Available at: http://www.nsf.gov/pubs/2005/nsb0540/ [Accessed 23 August 2007].

National Science Foundation (NSF). 2008. NSF Wide Investment Cyberinfrastructure. [Online] (Updated July 10 2008) Available at: http://www.nsf.gov/news/priority_areas/cyberinfrastructure/index.jsp [Accessed April 1 2009]

Natural Resources Canada. 2006. GeoBase Unrestricted Use Licence Agreement. [Online] (Updated 2 may 2006) Available at: http://www.geobase.ca/geobase/en/licence.jsp [Accessed 1 April 2009].

Natural Resources Canada. 2009. Geogratis Splash Page. [Online] Available at: http://geogratis.cgdi.gc.ca/ [Accessed 1 April 2009].

Natural Resources Canada. 2009. The Atlas of Canada. [Online] (Updated 22 May 2007) Available at: http://atlas.nrcan.gc.ca/site/english/index.html [Accessed 1 April 2009].

Nunaliit.2006. The Nunaliit Cybercartographic Framework License. [Online], Available at: http://nunaliit.org/license.html [Accessed 1 April 2009].

Nunaliit.2009. The Nunaliit Cybercartographic Framework Internet Site. [Online] Available at: http://nunaliit.org/index.html [Accessed 1 April 2009].

Open Geospatial Consortium (OGC). 2006. OGC Specifications (Standards). [Online] Available at: http://www.opengeospatial.org/standards [Accessed September 2006]

Open Geospatial Consortium (OGC). 2009. Data Preservation Working Group. [Online] Available at: http://www.opengeospatial.org/projects/groups/preservwg [Accessed April 1 2009].

Preston, Randy. randy@interpares.org. IR or TDR vs Archive. [E-Mail]. Message to Tracey P. Lauriault (tlauriau@gmail.com). Sent March 2 2009 6:44 PM.

Pulsifer, P.L., Parush, A., Lindgaard, G., & Taylor, D. R. F. (2005). The Development of the Cybercartographic Atlas of Antarctica. In D. R. F. Taylor (Ed.), Cybercartography: Theory and Practice (pp. 461-490). Amsterdam: Elsevier.

Pulsifer, P. L., Caquard, S. & Taylor, D.R. F. 2007. Toward a New Generation of Community Atlases - The Cybercartographic Atlas of Antarctica. In Cartwright, W., M. Peterson and G. Gartner eds. Multimedia Cartography, Second Edition, Springer-Verlag. Ch.14.

Pulsifer, P.L., Hayes, A., Fiset, J-P., & Taylor, D. R. F. (2008). An Open Source Development Framework in Support of Cartographic Integration. In Peterson, M. (Ed.). International Perspectives on Maps and the Internet (pp.165-185). Berlin: Springer.

Reagan, Brad. 2006. The Digital Ice Age: The documents of our time are being recorded as bits and bytes with no guarantee of future readability. As technologies change, we may find our files frozen in forgotten formats. Will an entire era of human history be lost? (Popular Mechanics) [internet] (Published 2006) Available at: http://www.popularmechanics.com/technology/industry/4201645.html [Accessed April 1 2009].

Records Library Group (RLG) and the Online Computer Library Center (OCLC). 2002. Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report. (OCLC Pulbication) [Internet] Available at: http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf [Accessed April 1 2009].

Records Library Group (RLG) and the US National Archives and Records Administration (NARA). 2005. Audit Checklist for Certifying Digital Repositories (OCLC RLG) [Internet] Available at http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070511/viewer/file2416.pdf [Accessed April 1 2009].

Registry of Open Access Repositories (ROAR). 2007. Home Page. [Online] Available at: http://roar.eprints.org/index.php [Accessed April 1 2009].

Schut, Peter. 2000. Back from the Brink: the story of the remarkable resurrection of the Canada Land Inventory data. [Online] Available at: http://www.igs.net/~schut/cli.html [Accessed 27 March 27 2009].

Schwartz, Joan M. & Cook, Terry. 2002. Archives, Records, and Power: The Making of Modern Memory. Archival Science, 2, pp.1-19.

Science Commons. 2009. The Science Commons Page. [Online] Available at: http://sciencecommons.org/ [Accessed 1 April 2009].

Sierra Nevada ecosystem Project (SNEP) Digital Spatial Data Archive. 2009. Home Page. [Online] Available at: http://ceres.ca.gov/snep/data.html [Accessed April 1 2009].

Sleeman, Patricia. 2004. It's Public Knowledge: The National Digital Archive of Datasets. Archivaria 58, pp.173-200.

Social Science and Humanities Research Council (SSHRC). 2001. National Data Archive Consultation Report, Phase 1: Needs Assessment Report. (SSHRC Report) [Internet] Ottawa, Available at: http://www.sshrc.ca/site/about-crsh/publications/da_phase1_e.pdf [Accessed April 1 2009].

Social Science and Humanities Research Council (SSHRC). 2002. Final Report of the SSHRC National Consultation on Research Data Archiving, Building Infrastructure for Access to and Preservation of Research Data. [Online] Available at: http://www.sshrc.ca/web/about/publications/da_finalreport_e.pdf [Accessed 23 August 2007].

Social Science and Humanities Research Council (SSHRC). 2002. Research Data Archiving Policy. [Online] Available at: http://www.sshrc.ca/web/apply/policies/edata_e.asp [Accessed 23 August 2007].

Strong, David F. & Leach, Peter B. 2007. The Final Report of the National Consultation on Access to Scientific Data, [Online] Available at: http://ncasrd-cnadrs.scitech.gc.ca/NCASRDReport_e.pdf [Accessed 5 January 2007].

Taylor, D. R. F. 1997. Maps and Mapping in the Information Era, Keynote address at 18th ICA Conference, Stockholm, Sweden. In Ottoson, L. ed., Proceedings, 1, Swedish Cartographic Society.

Taylor, D. R. Fraser, Lauriault, Tracey P. & Peter L. Pulsifer. 2005. Preserving and Adding Value to Scientific Data: The Cybercartographic Atlas of Antarctica. In PV2005: Ensuring Long-Term Preservation and Adding Value to Scientific Technical Data. Edinburgh.

Taylor, D.R. Fraser. 2003. The Concept of Cybercartography. In Peterson, M. ed. Maps and the Inernet, Elsevier. pp.413-418.

Taylor, D. R. F., Pulsifer, Peter L.and Lauriault, Tracey P. 2005. Preserving and Adding Value to Scientific Data: The Cybercartographic Atlas of Antarctica. PV2005, Ensuring Long-term Preservation and Adding Value to Scientific and Technical Data. Edinburgh, Scotland.

Tosta, Nancy & Domaratz, Michael. 1997. The U.S. National Spatial Data Infrastructure. In Geographic Information Research: Bridging the Atlantic, ed. Massimo C. Craglia and Helen Couclelis, London, 1997.

UK Data Archive (UKDA). 2009. About Page. [Online] (Updated April 1 2009) Available at: http://www.data-archive.ac.uk/about/about.asp [Accessed April 1 2009].

Visible Past initiative. 2009. Home Page. [Online] Available at: http://www.visiblepast.net/home/ [Accessed 1 April 2009].

Warren Mills, Jacqueline; Curtis, Andrew; Pine, John C.; Kennedy, Barrett; Jones, Farrell; Ramani, Ramesh & Bausch, Douglas. 2008. The clearinghouse concept: a model for geospatial data centralization and dissemination in a disaster. Disasters, 32(3), pp.467-479.

Waters, Donald & Garrett, John. 1996. Preserving Digital Information, Report of the Task Force on Archiving of Digital Information. [Online] Available at: http://www.rlg.org/en/page.php?Page_ID=114 [Accessed 11 January 2007].

Wilson, C. and Robert A. O'Neil, 2009, GeoGratis: A Canadian Geospatial Data Infrastructure Component that Visualises and Delivers Free Geospatial Data Sets.

Springer
http://www.springer.com/978-3-642-12732-8

Preservation in Digital Cartography
Archiving Aspects
Jobst, M. (Ed.)
2010, XXII, 310 p. 73 illus., Hardcover
ISBN: 978-3-642-12732-8