

Conference Proceedings

Edited by:
Luciana Duranti and
Elizabeth Shaffer

The Memory of the World in the Digital Age: Digitization and Preservation

An international conference
on permanent access to
digital documentary heritage

Hosted by:



a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA

In collaboration with



UNIVERSITY OF
TORONTO



United Nations
Educational, Scientific and
Cultural Organization



Memory of the World
20th Anniversary

26 to 28 SEPTEMBER 2012

Vancouver, British Columbia, Canada
Sheraton Vancouver Wall Centre



Conference Proceedings

Edited by:
Luciana Duranti and
Elizabeth Shaffer

The Memory of the World in the Digital Age: Digitization and Preservation

An international conference
on permanent access to
digital documentary heritage

26 to 28 SEPTEMBER 2012
Vancouver, British Columbia, Canada
Sheraton Vancouver Wall Centre

UNESCO Memory of the World Programme, Knowledge Societies Division

This book of Proceedings includes most of the papers and posters presented at the International Conference “The Memory of the World in the Digital Age: Digitization and Preservation” held on 26–28 September 2012 in Vancouver, British Columbia, Canada, by the UNESCO Memory of the World Programme, Knowledge Societies Division, and The University of British Columbia in collaboration with the University of Toronto.

The proceedings have been compiled and formatted with minor editing; papers and posters appear as submitted. The authors are responsible for the choice and the presentation of the facts contained in this publication and for the opinions they express, which are not necessarily those of UNESCO and do not commit the Organization.

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The organizers of this UNESCO Memory of the World Programme Conference would like to sincerely thank everyone who contributed to the Conference in Vancouver and to these proceedings.

Published by UNESCO 2013, with the financial support of the Social Sciences and Humanities Research Council of Canada | Conseil de recherches en sciences humaines du Canada (SSHRC) and the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) Project.



Social Sciences and Humanities
Research Council of Canada

Conseil de recherches en
sciences humaines du Canada



InterPARES Project

International Research on Permanent Authentic Records in Electronic Systems

Preface

This publication presents the proceedings of the international conference ‘Memory of the World in the Digital Age: Digitization and Preservation’ which was held in Vancouver, Canada, from 26 to 28 September 2012.

More than 500 experts and other interested persons from all regions of the world participated in this knowledge-sharing and policy-driving event to discuss and exchange opinions on how to protect the world’s documentary heritage. Although this heritage is the record of knowledge, its physical carriers are extremely vulnerable and can easily disappear without a trace. Whether recorded on a clay tablet or an electronic tablet, our methods of sharing content and knowledge need to be protected.

It is impossible to exaggerate the importance of documentary heritage in our lives. It governs our actions whether these relate to creating the basis of mutual respect between different civilizations and communities or building knowledge societies. Documentary heritage provides the foundation of peace, our identity and knowledge.

UNESCO’s interest in this subject matter is as fundamental as its constitution with its mandate to contribute to building peace through the spread of knowledge from improved access to printed and published materials. These core materials, our documentary heritage, have been preserved in archives, libraries and museums for generations.

But while measures needed to maintain access to print materials are globally understood, the newer challenges related to preserving digital information are not keeping pace with technological development. The need for dedicated hardware and software, associated with their rapid obsolescence, hamper our ability to keep invaluable content accessible. Unless timely migration to newer technologies, operating systems and software platforms is assured, we face the risk developing digital Alzheimer’s.

UNESCO’s expectation from this Conference was to obtain a better definition of our expected role, and our contribution to setting a global digital agenda. The UNESCO/UBC Vancouver Declaration sets out specific recommendations which we will be implementing and incorporating into our digital strategy. Likewise, we expect that our Member States, professional organizations and private sector bodies will also implement the recommendations addressed to them.

Only through collaborative strategic alliances can we overcome the major challenges threatening the preservation of digital information. We believe that the presentations featured in this publication provide the basis for a global commitment to preserving the memory of our world in this digital age.

Jānis Kārklīš
Assistant Director-General
for Communication and Information

Contents

Preface	4
Opening Keynotes	
Kenneth Thibodeau <i>Wrestling with Shape-Shifters: Perspectives on Preserving Memory in the Digital Age</i>	15
Luciana Duranti <i>Trust and Conflicting Rights in the Digital Environment.....</i>	24
Anne Thurston <i>Digitization and Preservation: Global Opportunities and Cultural Challenges.....</i>	31
Intellectual Property Infrastructure Initiatives for Digital Heritage	
Heather Christenson and John P. Wilkin <i>Intellectual Property Rights & the HathiTrust Collection.....</i>	39
Elizabeth Townsend Gard <i>The Durationator® Copyright Experiment.....</i>	46
Kate Hennessy <i>The Intangible and the Digital: Participatory Media Production and Local Cultural Property Rights Discourse.....</i>	58
Preservation Infrastructures: Current Models and Potential Alternatives	
Ilaria Pescini and Walter Volpi <i>An Example to Follow: An Infrastructure for Interoperability and Governance in the Tuscan Public System for Digital Preservation</i>	70
Francis G. Mwangi <i>The Road to Providing Access to Kenya’s Information Heritage: Digitization project in the Kenya National Archives and Documentation Service (KNADS)</i>	83
Jeremy York <i>A Preservation Infrastructure Built to Last: Preservation, Community, and HathiTrust.....</i>	92
Hrvoje Stančić, Arian Rajh and Ivor Milošević <i>“Archiving-as-a-Service”: Influence of Cloud Computing on the Archival Theory and Practice.....</i>	108
The CODATA Mission: Preserving Scientific Data for the Future	
Elizabeth Griffin and the CODATA DARTG Team <i>Recovering the Forgettery of the World</i>	127
Patrick C. Caldwell <i>Tide Gauge Data Rescue</i>	134
Stephen Del Greco <i>Environmental Data Through Time: Extending The Climate Record</i>	150
Tracey P. Lauriault and D. R. Fraser Taylor <i>The Map as a Fundamental Source in the Memory of the World.....</i>	160

Preserving Tradition and Performing Arts in Digital Form

Ravi Katikala, Kurt Madsen and Gilberto Mincaye Nenquimo Enqueri

Life at the Edge of the Internet: Preserving the Digital Heritage of Indigenous Cultures..... 190

Lekoko Kenosi

Digital Madness, Archival Theory and the Endangered Sound Archives of Radio Botswana 206

Jørgen Langdalen

Editing Historical Music in the Age of Digitization..... 212

Lauren Sorensen and Tanisha Jones

Developing and Implementing a Digital Video Repository for Legacy Dance Documentation: Dance Heritage Coalition's Secure Media Network..... 217

Beyond Access: Digitization to Preserve Culture

Fernanda Maria Melo Alves, José António Moreira González and José Manuel Matias

Safeguarding of the Portuguese Language Documentary Heritage: The Lusophone Digital Library..... 229

Benoit Ferland et Tristan Müller

Le réseau francophone numérique 236

John Van Oudenaren

The World Digital Library..... 246

Strategies for Building Digital Repositories

Bronwen Sprout and Sarah Romkey

A Persistent Digital Collections Strategy for UBC Library 257

Neil Grindley

Building the Business Case for Digital Preservation..... 269

Kevin Bradley

Requirements of a Remote Repository 278

Digital Forensics for the Preservation of Digital Heritage

Wayne W. Liu

Accountability for Archival Digital Curation in Preserving the Memory of the World..... 288

Christopher A. Lee and Kam Woods

Automated Redaction of Private and Personal Data in Collections: Toward Responsible Stewardship of Digital Heritage..... 298

Corinne Rogers and Jeremy Leighton John

Shared Perspectives, Common Challenges: A History of Digital Forensics & Ancestral Computing for Digital Heritage..... 314

Giving a Permanent Digital Voice to the Silenced

Terry Reilly

For the Children Taken: The Challenge to Truth Commissions in Building digital collections for research and long-term preservation 338

National Strategies as the Foundation of Togetherness

Andris Vilks and Uldis Zariņš

National Planning as the Key for Successful Implementation of Digitization Strategies..... 348

Ivan Chew & Haliza Jailan

Preserving the Crowdsourced Memories of a Nation: The Singapore Memory Project 354

Ernesto C. Bodé

Digital Preservation Policy of The Chamber of Deputies: Methodology for its development..... 366

Web 2.0 Products as Documentary Digital Heritage: Can We Access and Preserve Them?

Jamie Schleser

Unprotected Memory: User-Generated Content and the Unintentional Archive 378

Heather Ryckman

Context 2.0: User Attitudes to the Reliability of Archival Context on the Web 393

Lisa P. Nathan and Elizabeth Shaffer

Preserving Social Media: Opening a Multi-Disciplinary Dialogue..... 410

The Role of Culture in Digitization and Digital Preservation

Fiorella Foscari, Gillian Oliver, Juan Ilerbaig and Kevin Krumrei

Preservation Cultures: Developing a Framework for a Culturally Sensitive Digital Preservation Agenda..... 419

Tukul Sepania Walla Kaiku and Vicky Puipui

Political, Cultural and Professional Challenges for Digitization and Preservation of Government Information in Papua New Guinea: An Overview 431

Xincai Wang and Yunxia Nie

Current Situation, Problems and Prospects of the Digital Preservation of Documentary Heritage in China 439

Open Archival Information System Reference Model: Answer or Inspiration?

Stefano S. Cavaglieri

Digital Archiving Systems Confronted with the OAIS Reference Model..... 451

Saeed Rezaei Sharifabadi, Mansour Tajdaran and Zohreh Rasouli

A Model for Managing Digital Pictures of the National Archives of Iran: Based on the Open Archival Information System Reference Model 457

Collaboration in Digital Preservation or Lack Thereof: What Works

Maria Guercio

Digital Preservation in Europe: Strategic Plans, Research Outputs and Future Implementation. The Weak Role of the Archival Institutions..... 467

Rolf Källman

Models for National Collaboration: Coordination of the Digital Cultural Heritage in Sweden..... 482

Victoria Reich

Building and Preserving Library Digital Collections Through Community Collaboration..... 489

Steve Knight

National Library of New Zealand, Digital Preservation and the Role of UNESCO..... 500

The Economics of Preserving Digital Information

David S. H. Rosenthal, Daniel C. Rosenthal, Ethan L. Miller, Ian F. Adams, Mark W. Storer and Erez Zadok

The Economics of Long-Term Digital Storage 513

Ulla Bøgvad Kejser, Anders Bo Nielsen and Alex Thirifays

Modelling the Costs of Preserving Digital Assets..... 529

L.M. Udaya Prasad Cabral

Economically Easy Method to Digitize Oversized Documents with Special Reference to Ola Leaf Manuscripts in Sri Lanka 540

Patricia Liebetrau

Preserving Our Heritage: An Independent Advantage 549

Is A New Legal Framework Required for Digital Preservation or Will Policy Do?

Tony Sheppard

Is a New Legal Framework Required for Digital Preservation or Will Policy Do? Building a Legal Framework to Facilitate Long-term Preservation of Digital Heritage: A Canadian Perspective 559

Alicia Barnard

Development of Policies and Requirements for Ingesting and Preserving Digital Records Into a Preservation System: Where to start? 570

Jason R. Baron and Simon J. Attfield

Where Light in Darkness Lies: Preservation, Access and Sensemaking Strategies for the Modern Digital Archive 580

Elaine Goh

Strengthening the Regulatory Framework in a Digital Environment: A Review of Archives Legislation 596

Digital Curation: Convergence of Challenges, Institutions and Knowledge

Sarah Higgins

Digital Curation: The Challenge Driving Convergence across Memory Institutions 607

Jackie R. Esposito

Digital Curation: Building an Environment for Success..... 624

Patricia Forget

Célébrations institutionnelles : Événement catalyseur propice à l'implantation d'un projet de conservation du patrimoine numérique permettant de réunir les acteurs d'intérêts divergents 636

Jeannette A. Bastian and Ross Harvey

The Convergence of Cultural Heritage: Practical Experiments and Lessons Learned 650

Digitization and Digital Preservation Experiences in a Developing Country Perspective

Elizabeth F. Watson

The Conservation and Preservation of Heritage in the Caribbean: What Challenges Does Digitization Pose? 661

Richard Marcoux, Laurent Richard and Mamadou Kani Konaté <i>Digital Preservation of Demographic Heritage: Population Censuses and Experiences in Mali and the Democratic Republic of the Congo</i>	672
Brandon Oswald <i>Partnership in Paradise: The Importance of Collaboration for Handling Traditional Cultural Expression Material in the Pacific Islands.....</i>	685
Ensuring That it Won't Happen Again	
Victoria L. Lemieux <i>Financial Records and Their Discontents: Safeguarding the Records of our Financial Systems</i>	700
Myron Groover <i>The White House E-Mail Destruction Scandal of 2007: A Case Study for Digital Heritage.....</i>	713
Kenneth Thibodeau <i>The Perfect Archival Storm: The Transfer of Electronic Records from the G.W. Bush White House to the National Archives of the United States.....</i>	724
Trusting Records	
Lorraine Dong <i>The Ethical and Legal Issues of Historical Mental Health Records as Cultural Heritage</i>	735
Marie Demoulin et Sébastien Soye <i>L'authenticité, de l'original papier à la copie numérique : Les enjeux juridiques et archivistiques de la numérisation</i>	745
Web Archiving as Part of Building the Documentary Heritage of Our Time	
Liu Hua, Yang Menghui, Zhao Guojun and Feng Huiling <i>Chinese Web Archiving and Statistical Analysis on Chinese Web Archives</i>	765
Gustavo Urbano Navarro <i>Implications of the Web Semantization on the Development of Digital Heritage.....</i>	775
Matt Holden <i>Preserving the Web Archive for Future Generations: Practical Experiments with Emulation and Migration Technologies</i>	783
Technology as the Mediator of Heritage and Its Relations with People	
Ian S. King <i>The Turtle At The Bottom: Reflections on Access and Preservation for Information Artefacts.....</i>	797
Erik Borglund <i>Challenges to Capture the Hybrid Heritage: When Activities Take Place in Both Digital and Non-Digital Environments.....</i>	814
Limited Resources or Expertise: Case Studies in Addressing the Issue	
Jean Bosco Ntungirimana <i>La problématique de la préservation de la mémoire collective au Burundi à l'ère des NTIC : Étude de cas menée à la Cour supreme</i>	823
Farah Al-Sabah <i>Digitizing A Survivor's Identity: The Past, Present, and Future of the Kuwait National Museum Archives.....</i>	838

Wayne W. Torborg, Theresa M. Vann and Columba Stewart	
<i>The Challenges of Manuscript Preservation in the Digital Age</i>	851
Plenary 3 Keynotes	
Dietrich Schüller	
<i>Challenges for the Preservation of Audiovisual Documents: A General Overview</i>	863
International Perspectives and Cooperation	
Claudia Nicolai, Rachele Oriente and Fernando Serván	
<i>One Year of Efforts for Digital Preservation at FAO</i>	871
Peter Burnhill, Françoise Pelle, Pierre Godefroy, Fred Guy, Morag Macgregor and Adam Rusbridge	
<i>Archiving the World's E-Journals: The Keepers Registry as Global Monitor</i>	880
The World Audiovisual Memory: Practical Challenges, Theoretical Solutions?	
Jean Gagnon	
<i>Treasures That Sleep: Film Archives in the Digital Era</i>	892
Caroline Frick	
<i>Seeing, Hearing, and Moving Heritage: Issues and Implications for the World's Audiovisual Memory in the Digital Age</i>	896
Edoardo Ceccuti	
<i>The Digitization of Films and Photos of the Istituto Luce</i>	904
Adam Jansen	
<i>Challenges and Triumphs: Preserving HD Video at the UBC School of Journalism</i>	909
Mick Newnham, Trevor Carter, Greg Moss and Rod Butler	
<i>Digital Disaster Recovery for Audiovisual Collections: Testing the Theory</i>	921
Metadata and Formats for Digitization and Digital Preservation	
Joseph T. Tennis	
<i>Data, Documents, and Memory: A Taxonomy of Sources in Relation to Digital Preservation and Authenticity Metadata</i>	933
Adam Rabinowitz, Maria Esteva and Jessica Trelogan	
<i>Ensuring a Future for the Past: Long-term Preservation Strategies for Digital Archaeological Data</i>	941
Giovanni Michetti and Paola Manoni	
<i>It FITS the Cultural Heritage! Formats for Preservation: From Spatial Data to Cultural Resources</i>	955
Lois Enns and Gurp Badesha	
<i>File Viewers: Examining On-the-Fly File Format Conversion</i>	962
Walter Allasia, Fabrizio Falchi, Francesco Gallo and Carlo Meghini	
<i>Autonomic Preservation of "Access Copies" of Digital Contents</i>	976
A Methodology Framework to Ensure Preservation	
Anca Claudia Prodan	
<i>Bias and Balance in the Preservation of Digital Heritage</i>	989

Giovanni Michetti	
<i>Archives Are Not Trees: Hierarchical Representations in Digital Environment.....</i>	1002
Göran Samuelsson	
<i>The New Information Landscape: The Archivist and Architect – Drawing on a Common Map?.....</i>	1011
Shadrack Katuu	
<i>Enterprise Content Management and Digital Curation Applications: Maturity Model Connections.....</i>	1025
Christopher J. Prom	
<i>Facilitating the Aggregation of Dispersed Personal Archives: A Proposed Functional, Technical, and Business Model.....</i>	1042
Digital Objects as Forensic Evidence	
Carsten Rudolph and Nicolai Kuntze	
<i>Constructing and Evaluating Digital Evidence for Processes.....</i>	1057
Aaron Alva, Scott David and Barbara Endicott-Popovsky	
<i>Forensic Barriers: Legal Implications of Storing and Processing Information in the Cloud.....</i>	1064
Michael Losavio, Deborah Keeling and Michael Lemon	
<i>Models in Collaborative and Distributed Digital Investigation: In the World of Ubiquitous Computing and Communication Systems.....</i>	1079
Fabio Marturana and Simone Tacconi	
<i>Cloud Computing Implications to Digital Forensics: A New Methodology Proposal.....</i>	1093
Andrew F. Hay and Gilbert L. Peterson	
<i>Acquiring OS X File Handles Through Forensic Memory Analysis.....</i>	1102
Institutional and Inter-Organizational Initiatives in Digitization	
Anup Kumar Das	
<i>Digitization of Documentary Heritage Collections in Indic Language: Comparative Study of Five Major Digital Library Initiatives in India.....</i>	1126
Ronald Walker	
<i>Digital Heritage Preservation - Economic Realities and Options.....</i>	1139
S. K. Reilly	
<i>Positioning Libraries in the Digital Preservation Landscape.....</i>	1146
Heidi Rosen, Torsten Johansson, Mikael Andersson and Henrik Johansson	
<i>Experiences from Digidaily: Inter-Agency Mass Digitization of Newspapers in Sweden.....</i>	1153
Preserving Images: What Do We Need to Know?	
Adama Aly Pam	
<i>Chemins de la mémoire : Les archives audiovisuelles au secours de l'identité d'une organisation internationale africaine.....</i>	1163
Krystyna K. Matusiak and Tamara K. Johnston	
<i>Digitization as a Preservation Strategy: Saving and Sharing the American Geographical Society Library's Historic Nitrate Negative Images.....</i>	1173
Jessica Bushey	
<i>Born Digital Images: Creation to Preservation.....</i>	1189

Angelina Altobellis

<i>Essential Skills for Digital Preservation: Addressing the Training Needs of Staff in Small Heritage Institutions</i>	1198
---	------

Small and Large Scale Digitization: Towards a Shared Conceptual Model

Peter Botticelli, Patricia Montiel-Overall and Ann Clark

<i>Building Sustainable Digital Cultural Heritage Collections: Towards Best Practices for Small-scale Digital Projects</i>	1205
--	------

Marco de Niet, Titia van der Werf and Vincent Wintermans

<i>Preserving Digital Heritage: The UNESCO Charter and Developments in the Netherlands</i>	1219
--	------

Paul Conway

<i>Validating Quality in Large-Scale Digitization: Findings on the Distribution of Imaging Error</i>	1233
--	------

Lars Björk

<i>Lost in Transit: The Informative Capacity of Digital Reproductions</i>	1252
---	------

Preservation of Audiovisual Material

Mike Casey

<i>The Media Preservation Initiative at Indiana University Bloomington</i>	1266
--	------

George Blood

<i>Video Compression...For Dummies?</i>	1273
---	------

Pio Pellizzari, Álvaro Hegewich

<i>The Ibero-American Preservation Platform of Sound and Audiovisual Heritage</i>	1289
---	------

Trusting Data and Documents Online

Junbin Fang, Zoe Lin Jiang, Mengfei He, S.M. Yiu, Lucas C.K. Hui, K.P. Chow and Gang Zhou

<i>Investigating and Analysing the Web-based Contents on Chinese Shanzhai Mobile Phones</i>	1297
---	------

Junwei Huang, Yinjie Chen, Zhen Ling, Kyungseok Choo and Xinwen Fu

<i>A Framework of Network Forensics and its Application of Locating Suspects in Wireless Crime Scene Investigation</i>	1310
--	------

F.R. Van Staden and H.S. Venter

<i>Implementing Digital Forensic Readiness for Cloud Computing Using Performance Monitoring Tools</i>	1329
---	------

Yongjie Cai and Ping Ji

<i>Security Monitoring for Wireless Network Forensics (SMoWF)</i>	1340
---	------

Workshops

Peter Van Garderen, P. Jordan, T. Hooten, C. Mumma and E. McLellan

<i>The Archivemata Project: Meeting Digital Continuity's Technical Challenges</i>	1349
---	------

Hannes Kulovits, Christoph Becker and Andreas Rauber

<i>Roles and Responsibilities in Digital Preservation Decision Making: Towards Effective Governance</i>	1360
---	------

Posters and Presentations

Collence Takaingehamo Chisita and Amos Bishi

<i>Challenges and Opportunities of Digitizing and Preserving Cultural Heritage in Zimbabwe</i>	1382
--	------

Donna McRostie	
<i>The long and winding road from aspiration to implementation – building an enterprise digitization capability at the University of Melbourne</i>	1384
Asger Svane-Knudsen and Jiří Vnouček	
<i>Retrieving a part of Danish colonial history: From dust to digital copy.....</i>	1386
Mitra Samiee and Saeed Rezaei Sharifabadi	
<i>A Paradigm for the preservation of national digital memory of Iran</i>	1392
Chinyere Otuonye, Tamunoibuomi F. Okajagu, Samuel O. Etatuvie, Emmanuel Orgah, Gift Eyemienbai, Luke Oyovwevotu, Ewoma Borgu, and Janet Ukoha	
<i>Insights on the Digitization of Traditional Medicine Knowledge in Nigeria</i>	1395
Nader Naghshineh and Saeed Nezareh	
<i>Crowd-sourced digital preservation: An Iranian model</i>	1397
Chris Muller	
<i>Data at Risk: The Duty to Find, Rescue, Preserve</i>	1399
Natalia Grincheva	
<i>Digital diplomacy: Providing access to cultural content, engaging audiences on a global scale</i>	1401
Rusnah Johare	
<i>Preserving digital research data</i>	1403
Claudia M. Wanderley	
<i>Multilingualism at the University of Campinas.....</i>	1405
Anne Thurston	
<i>Open government and trustworthy records</i>	1407
Jan Marontate, David Murphy, Megan Robertson, Nathan Clarkson and Maggie Chao	
<i>Canada – Aural memories: A case study of soundscape archives</i>	1421
Na Cai, Leye Yao and Liu Liu	
<i>Creating Social Memories of Major Events in China: A Case study of the 5•12 Wenchuan Earthquake Digital Archive</i>	1423
Addendum	
Howard Besser	
<i>Archiving Large Amounts of Individually-Created Digital Content: Lessons from Archiving the Occupy Movement</i>	1432
Nadja Wallaszkovits	
<i>Digitisation of Small Sound Collections: Problems and Solutions</i>	1440
UNESCO/UBC Vancouver Declaration	
<i>The Memory of the World in the Digital Age: Digitization and Preservation</i>	1452
Sponsors	

The Map as a Fundamental Source in the Memory of the World

Tracey P. Lauriault¹ and D. R. Fraser Taylor²

¹Postdoctoral Fellow, Geomatics and Cartographic Research Centre (GCRC), Carleton University, Canada, tlauriau@gmail.com; ²FRSC, Distinguished Research Professor and Director of the Geomatics and Cartographic Research Centre, Carleton University, Canada; member of the CODATA Data at Risk Task Group, fraser_taylor@carleton.ca

Abstract

Maps and spatial information have been fundamental facets of the memory of societies from all over the world for millennia, and their preservation should be an integral part of government strategies in managing digital data. The digital era in map-making is a relatively recent activity; the first digital maps date from the 1960s. Digital mapping has accelerated very rapidly over the last decade and is now ubiquitous with an increasing amount of spatially referenced information being created by non-governmental organizations, academia, the private sector and government, as well as by social networks and citizen scientists. Unfortunately, despite that explosion of digital mapping little or no attention is being paid to preservation. As a result, the very maps that have been such a fundamental source of scientific and cultural information are now seriously at risk. Already we are losing map information faster than it is being created, and the loss of that central element of the cultural heritage of societies all over the world is a serious concern. There has already been a serious loss of maps; the Canada Land Inventory and the 1986 BBC Domesday Project are only two such casualties, and mapping agencies all over the world are struggling to preserve maps in the new digital era. It is somewhat paradoxical that it is easier to get maps that are hundreds, and in some cases thousands, of years old than maps of the late 20th and early 21st centuries. This paper examines the opportunities and challenges of preserving and accessing digital maps, atlases and geospatial information, all of which are Canada's cultural and scientific knowledge assets.

Authors

Dr. Tracey P. Lauriault is a Postdoctoral Fellow at the Geomatics and Cartographic Research Centre working on the SSHRC Partnership Development Project entitled Mapping the Legal and Policy Boundaries of Digital Cartography in collaboration with the Centre for Law, Technology and Society; the Canadian Internet Public Policy Clinic and Natural Resources Canada. She also participates in activities and represents the GCRC on topics related to the access to and preservation of data. As a citizen she is an open data advocate and is the co-founder of civicaccess.ca and datalibre.ca.

Dr. Professor Taylor is a Distinguished Research Professor of International Affairs and Geography and Environmental Studies at Carleton University, Canada and is also Director of the Geomatics and Cartographic Research Centre. He is a Fellow of the Royal Society of Canada and is widely recognized as one of the world's leading cartographers. Dr. Taylor's main research interests in cartography lie in the application of geographic information processing to the analysis of socio-economic issues and the presentation of the results in the form of cybercartographic atlases. Cybercartography is an innovative new concept which he first introduced in 1997. His current funded research involves working with aboriginal and Inuit communities to empower these communities to express their perceptions of their own environmental and socio-economic reality in new ways utilizing the Cybercartographic Atlas Framework. This includes the innovative open source software called Nunaliit. Dr. Taylor is Chair of the International Steering Committee for Global Mapping (ISCGM) and the International Advisory Group on the Arctic Spatial Data Infrastructure, he is also a member of the Mapping Africa for Africans Working Group of the International Cartographic Association, a member of the Joint Board of the Geospatial Information Societies and a member of the UN initiative on Global Geospatial Information Management.

1. Introduction

Maps do more than help us locate things. Maps and atlases, like books, reports, and archival records, enable us: to understand the territorial evolution of our nations, to see how colonial surveyors drew property boundaries, to visualize the spaces negotiated in treaties, and to assess the distribution of population and national resources. They can model climate change, the economy, or display the spatial relationships between cities, infrastructure and society. Maps inform planning, policy and economic decisions, while also shaping how we imagine spaces. They delight, are repositories of cultural and scientific knowledge and are a key part of our collective geographical memories.

In Canada, all levels of government create maps. At the Federal level, Natural Resources Canada (NRCan) produces, acquires and maintains the largest collection of maps, satellite and radar imagery, air photos, geospatial datasets, and publishes the *Atlas of Canada*. These are disseminated via geospatial data portals, map servers, and special programs such as *GeoGratis*. NRCan is also home to Canada's Geospatial Data Infrastructure (CGDI), which comprises the open and interoperable standards, specifications, practices, technology, framework data, operational policies and institutional environment to disseminate Canada's geospatial data assets on the Internet. Numerous other federal departments also produce geospatial data.¹ Geospatial data also represent the largest collection of freely accessible datasets currently being disseminated in the Treasury Board Secretariat's (TBS) Open Data Pilot portal.² Provincial and territorial governments also produce maps, are responsible for cadastres, and have geospatial databases to help manage transportation networks, watersheds, natural resources and to administer programs such as health and education. At the level of the city geographic information system (GIS) units can be found within information technology (IT) sections to manage infrastructure and other municipal responsibilities. Academic institutions create maps and atlases which are normally funded by government granting agencies, foundations or special programs such as the International Polar Year (IPY). Canadian non-governmental organizations (NGOs) produce maps on a variety of topics such as food security, water quality and population health. The private sector through companies such as Google, Autodesk, and ESRI is a major producer of digital maps and geomatics products in Canada and is a multibillion-dollar industry.³ The public and communities are also engaged in map-making endeavours by contributing volunteered geographic information (VGI)⁴ into Mashups,⁵ building local infrastructures,⁶

¹ For example the Public Health Agency of Canada; Fisheries and Oceans; Environment Canada; Aboriginal Affairs and Northern Development Canada; Citizen and Immigration and many others.

² See TBS Open Data Portal <http://www.data.gc.ca/default.asp?lang=En&n=F9B7A1E3-1> (accessed 29 August 2012).

³ See Statistics Canada, *Surveying and Mapping Services Bulletin*, (2010) Catalogue no. 63-254-X, available at <http://www5.statcan.gc.ca/bsoic/olc-cel/olc-cel?catno=63-254-XIE&lang=eng#formatdisp> (accessed 28 August 2012).

⁴ VGI is georeferenced user-contributed information that is collected and disseminated in such a way that others can use them. For more details see Teresa Scassa, "Legal issues with volunteered geographic information," 2012, *Canadian Geographer*, available at <http://onlinelibrary.wiley.com/doi/10.1111/j.1541-0064.2012.00444.x/abstract> doi:10.1111/j.1541-0064.2012.00444.x.

⁵ The Regroupement activists pour l'inclusion Québec (RAPLIQ), with Montreal Ouvert and OpenNorth collaborated to create a prototype accessibility audit of commercial establishments. People with disabilities collected VGI data and these were mapped using GoogleMap's open API. The Montréal Accessible mashup is available at <http://montrealaccessible.ca/> (accessed 18 August 2012).

⁶ The Canadian Council on Social Development (CCSD) created a membership based consortium called the *Community Data Program* that includes myriad georeferenced demographic and socio-economic data available at <http://communitydata-donneescommunautaires.ca/Home>. The Social Planning Council of Ottawa has created a Community Information and Mapping System (CIMS) as seen here <http://www.cims-scic.ca/> (accessed 29 August 2012).

creating mobile device apps,⁷ and by being engaged in participatory mapping⁸ or citizen science endeavours.⁹ Digital maps and geospatial data are ubiquitous artefacts in the 21st century.

Geospatial data, maps and atlases, have been produced for millennia,¹⁰ and are cultural, historical and scientific records of knowledge and advancement, which makes them a fundamental source in memory of the world. This source of knowledge, in its borne digital form, is however not being systematically preserved. The digital era in map making is a relatively recent phenomenon of the last fifty years¹¹ and has accelerated very rapidly over the last decade as software has become easier to use and are less expensive, geospatial data are more accessible, organizations have opened their application programming interfaces (APIs) and Internet social networking services have proliferated.

Unfortunately, little, or no, attention is being paid to preserving maps, atlases and their related data whether created by authoritative or non-authoritative sources, as a result, these are very much at risk. Already we are losing spatial information faster than it is being created and the loss of this central part of the cultural heritage of societies all over the world is a serious concern.

This paper will consider seven related topics. It will begin by illustrating the problem by discussing two map and atlas rescue and salvage endeavours. Secondly, cybercartographic atlases are now mapping traditional knowledge (TK) by applying participatory research techniques. These are considered by the Indigenous communities involved repositories of local, cultural, scientific and historical knowledge. In creating these atlases preservation is considered from the outset even though an adequate archive to ingest them has still not emerged. Thirdly, while most geospatial data in Canada are not preserved, there are some geospatial data archives and these will be discussed. Fourthly, data management (including preservation) and access has become a key policy issue in Canada and a number of national public consultations on the topic have taken place. Their outcomes will be reviewed. Fifthly, there are Canadian legislation, directives and policies, which mandate the management of government information assets, and a brief overview of those related to spatial data will be examined. Sixthly, geospatial data and maps are also official government records explicitly referred to in Canadian acts and regulation and those for which the Minister of NRCan is responsible are discussed. Seventhly, the preservation of maps, atlases and geospatial data has been examined in Canadian research and some results are showcased. Finally, the paper will conclude by examining the key issues presented and will suggest strategies for the preservation of these cultural and scientific assets.

⁷ Open data initiatives make data accessible to developers who create GeoApps such as *Restonet* (<http://restonet.ca/>) which maps food inspection results; Edmonton's *Historic Buildings Android App* (<http://contest.apps4edmonton.ca/apps/21>) or the BC Apps4ClimateChange contest apps such as *Waterly* (<http://www.waterly.ca/>) which tracks rainfall and combines these with watering restrictions providing users with a watering schedule.

⁸ The Inuit Sea Ice Use and Occupancy Project (ISIUOP), which produced the Inuit siku (sea ice), *Atlas* (<http://sikuatlas.ca/index.html>) which include data contributed and mapped by Inuit hunters and elders.

⁹ The *Water and Environmental Hub* (WeHub) project “is a cloud-based, open source web platform that aggregates, federates, and connects water data and information with users looking to download, analyse, model and interpret water and environmental-based information”. Citizens can upload their data to the system available at <http://www.watereenvironmentalhub.ca/about/project>. Also see Digital Fishers where citizens load observations into a crowdsourced database and analyse results, available at <http://digitalfishers.net/> (accessed 18 August 2012).

¹⁰ For example: the Bedolina and Giadighe petroglyphs at Valcamonica (2500 BC), Çatal Hüyük (6200 BC) wall paintings, the Hecataeus of Miletus *Ges Periodos* maps (c. 550 – 476 BCE), and Ptolemy's World Map (circa 150).

¹¹ The Canada Geographic Information System (CGIS) developed in 1960s is recognized as the first operational GIS. Dr. Roger Tomlinson who was then based Department of Forestry and Rural Development developed it.

2. The Rescue and Salvage of the Canada Land Inventory (CLI) and the 1986 Domesday Project

The Canada Land Inventory (CLI) was initiated in 1960 by the Department of Forestry and Rural Development. In 1963 the Canadian Geographic Information System (CGIS) was developed to automate the data management and mapping of CLI data. The CGIS was the world's first GIS, it was a milestone in the history of geographic and government computerization and it revolutionized mapping. One of the principal driving forces behind the CGIS "was the idea that the CLI maps could be interpreted and analysed in a myriad of ways if the information could be manipulated by computers."¹² It was established "as a joint federal-provincial project to guide the development of policy on the control and management of land-based resources."¹³ The CGIS ultimately grew to contain thousands of maps and unknowingly became a technology that "spawned an industry that today is worth billions of dollars."¹⁴ The CLI was conceived to "classify lands as to their capabilities; to obtain a firm estimate of the extent and location of each land class and to encourage use of CLI data in planning."¹⁵

The CLI was an incredibly ambitious program that mapped 2.6 million square kilometers of Canada and "the original cost of the program was in the order of 100's of millions of dollars in the 1970's."¹⁶ The CGIS was both a set of electronic maps and the "computer programs that allowed users to input, manipulate, analyse, and output those maps."¹⁷

By the late 1980s, the CGIS was no longer being used. Priorities changed, people retired, and institutional memory was fading. Numerous boxes of tapes and racks of documentation were left behind and only a few computers were capable of reading nine track tapes let alone run the programs. In 1995, an informal trans-organizational group of individuals from Statistics Canada, the *National Atlas of Canada*, Archives Canada and the private sector joined forces to restore it. Fortunately members of this salvage team had either formerly worked on the CGIS or had an interest in its preservation, which meant they had the requisite skills, knowledge, and more importantly they recognized the value of these data and knew that these were part of Canada's technological and scientific heritage that could be repurposed into other applications. 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 and eventually the CLI was distributed in the *GeoGratis*¹⁸ portal for free along with metadata and some text to help

¹² Schut, Peter. *Back from the Brink: the story of the remarkable resurrection of the Canada Land Inventory data*, (2000), available at: <http://web.archive.org/web/20011104160045/http://www.igs.net/~schut/cli.html> (accessed 20 August 2012).

¹³ Ahlgren, Dorothy and John McDonald. "The Archival Management of a Geographic Information System," *Archivaria* 13 (1981): 61.

¹⁴ Schut, *Back from the Brink*, (2000).

¹⁵ See the *GeoGratis* Canada Land Inventory description, available at <http://geogratis.cgdi.gc.ca/geogratis/en/collection/detail.do?sessionId=155D0541488771F32F19DE8CDD39A37E?i d=CB6E057B-0D85-9B22-29DE-6351369A8B02> (accessed 20 August 2012).

¹⁶ Wilson, Cameron and Robert A. O'Neil. "GeoGratis: A Canadian Geospatial Data Infrastructure Component that Visualizes and Delivers Free Geospatial Data Sets," in *Proceedings of the International Cartographic Association Conference, Ottawa, 2009*, available from <http://icaci.org/publications/> (accessed 20 August 2012).

¹⁷ Schut, *Back from the Brink*, (2000).

¹⁸ *GeoGratis* disseminates geospatial data at no cost and without restrictions to the public. It is the first open data initiative in Canada and was the first to adopt an unrestricted user licence, it is available at <http://geogratis.cgdi.gc.ca/geogratis/en/index.html> (accessed 21 August 2012).

interpret the content. The CLI “rapidly became their most popular product.”¹⁹ Not all of the CLI data were saved and the cost of the effort described above was very substantial. The CLI has become the common basemap in *Atlas of Canada*²⁰ maps, layers are used to train satellite imagery software to recognize land cover patterns in remotely sensed images²¹ which are then used to inform the managing of forests and agricultural areas among many others resources.

The UK 1986 Domesday project, like the CGIS, was groundbreaking and a milestone in the history of multimedia computing. It was started by the BBC in 1984 to celebrate the 900-year anniversary of the original William of Normandy 1086 Norman Domesday Survey. The 1986 Domesday project, was conceived as an ‘electronic exhibition’ displaying British life much like the Great Exhibition of 1851.²² It was an interactive atlas that relied on crowdsourced VGI enlisting the help of 14,000 British schools and students. School microcomputers were mobilized to collect survey data, neighbourhood stories and pictures. These data “were combined with thousands of maps, still photographs and central statistical, written and visual information.”²³ The project was funded by the BBC and the European Commission.²⁴ A national disc contained interactive national statistics, photos, newspaper and magazine clippings, virtual reality tours, and movies. The community disc, or the ‘people’s database’ was compiled by nearly 1,000,000 people, contained four by three kilometer blocks of land with photos, data and text for 80 cities.²⁵ These multimedia data²⁶ were georeferenced to Ordnance Survey (OS) maps, airphotos and satellite images. The hardware consisted of a BBC microcomputer with floppy disk drives and a tracker-ball; Philips Laservision in Read Only Memory (ROM); a high-resolution colour monitor and a proprietary retrieval and analysis software.²⁷ It was an easy to use, information rich, interactive database driven multimedia second generation GIS. It also put state data into the public domain, stimulated much discussion about the emerging ‘information marketplace’ and the public versus the private sector’s rights to access public data.²⁸

Software and hardware, however, quickly became obsolete. In the late 1990s rescue and salvage work began and, just in time, as only a couple of fully operational systems were still operative. In 1999 a consortium under the name of CAMiLEON was formed to emulate the Domesday software into modern computers and between 2002-2003 they successfully did so. In 2001 a process to reverse engineer the

¹⁹ Schut, *Back from the Brink*, (2000).

²⁰ The online version of the *Atlas of Canada*, available at <http://atlas.nrcan.gc.ca/site/english/index.html> (accessed 20 August 2012).

²¹ To see how the CLI was used in the *Atlas of Canada*, see chapter 5 and the appendices of Lauriault, Tracey. *Data, Infrastructures and Geographical Imaginations*, Ph.D. diss., Carleton University, Ottawa, 2012.

²² See the discussion by the producers of the Domesday Project: Goddard, John and Peter Armstrong, “The 1986 Domesday Project,” *Transactions of the Institute of British Geographers, New Series* 11, no. 3 (1986): 290-295, available at <http://www.jstor.org/stable/621789> (accessed 20 August 2012).

²³ Darlington, Jeffrey, Andy Finney and Adrian Pearce. “Domesday Redux: The Rescue of the BBC Domesday Project Videodiscs,” *Ariadne* 36 (2003), available at <http://www.ariadne.ac.uk/issue36/tna/> (accessed 20 August 2012).

²⁴ For details about software, code, algorithms and hardware refer to the Darlington, Finney and Pearce (2003). For GIS innovation, geospatial data, functionality and development see David Rhind, Peter Armstrong and Stan Openshaw. “The Domesday Machine: A Nationwide Geographical Information System,” *The Geographical Journal* 154, no. 1 (1988): 56-68, available at <http://www.jstor.org/stable/633476> (accessed 20 August 2012).

²⁵ Rhind, Armstrong and Openshaw (1988), “The Domesday Machine: A Nationwide Geographical Information System.”

²⁶ 40,000 photographs, 22,000 maps, full Ordnance Survey Coverage at 1:50 000, see Goddard and Armstrong (1986).

²⁷ See Rhind, Armstrong and Openshaw (1988).

²⁸ Goddard and Armstrong (1986), pp. 294-295.

data began and by 2004, Adrian Pearce was able make them useable and accessible online.²⁹ And, in 2003 high quality copies of the original discs were made and the UK National Archives re-disseminated it.³⁰ Today the 1986 Domesday Project is available on the BBC *Domesday Reloaded*³¹ website while the content of the 1086 Domesday volumes is made available online with locations georeferenced with page citations using Google Maps in a project called the *Open Domesday*.³²

Both the 1960's Canada Land Inventory (CLI) and the 1986 Domesday Project were massive, expensive, national scale projects. They were technologically, scientifically and procedurally innovative marking a turning point in the history of computerization and science.³³ The CLI informed the creation of wildlife reserves, legislation to protect farmland, and were base maps in ecological frameworks which still in use today.³⁴ Both projects contributed to the geographical knowledge base of their respective nations and remain well-used, loved and important scientific records.³⁵ In both cases their data were inseparable from their software while the 1986 Domesday data could only be viewed with BBC proprietary hardware. Hardware, storage media, and software became obsolete and teams of dedicated people were able to restore these historical records and make them available to thousands of users today. These two examples illustrate the importance of considering long-term preservation at the point of creation by implementing good record keeping practices and to have a data management strategy in place as projects evolve.³⁶ Furthermore, they are arguments against the use of proprietary systems, and for the adoption of open specifications, standards and interoperability and for 'proactive archiving'.³⁷

²⁹ For the technical details on the rescue of the 1986 Domesday, see Darlington, Finney and Pearce (2003), "Domesday Redux."

³⁰ See the "Story of the Domesday Project," available on the *Domesday Reloaded* webpage at <http://www.bbc.co.uk/history/domesday/story> (accessed 20 August 2012).

³¹ The website is available at <http://www.bbc.co.uk/history/domesday/using-domesday> (accessed 20 August 2012).

³² *Open Domesday* is available at <http://domesdaymap.co.uk/> (accessed 20 August 2012).

³³ See R. J. Morris, "History and Computing: Expansion and Achievements," *Social Science Computer Review* 9, no. 2 (summer 1991): 215-230.

³⁴ The *Canada Encyclopedia* entry on the CLI provides describes of how it has contributed to the advancement of the Canadian economy and society, available at <http://www.thecanadianencyclopedia.com/articles/canada-land-inventory> (accessed 20 August 2012).

³⁵ There were 45,373 CLI map downloads in 2011-2012 and The Province of Ontario has a very popular CLI product available on this website <http://www.omafra.gov.on.ca/english/landuse/feed-in-tariffprogram.htm#3>, email communication received from Odette Trottier, Natural Resources Canada, Centre for topographic information on 28 August 2012.

³⁶ The InterPARES 2 Case Study on the *Cybercartographic Atlas of Antarctica* demonstrated that it is possible to design these aspects into systems at the point of creation, CS06 is available at http://www.interpares.org/ip2/ip2_case_studies.cfm?study=5. The IPY project mandated that all contributing scientists implement a data management strategy as part of funded projects see the Canada *IPY Data Management* at http://www.api-ipy.gc.ca/pg_IPYAPI_052-eng.html. Others have written about the technological issues of open source, standards and specifications, see Andrew Williamson, "Strategies for managing digital content formats," (2005) *Library Review* 54, no. 9: 508-513, and the Open Geospatial Consortium Data Preservation Working Group at <http://www.opengeospatial.org/projects/groups/preservdwg> (accessed 21 August 2012).

³⁷ In the case of cybercartographic atlases the use of open source, adherence to interoperable and open specifications, metadata as well as storing of multiple copies in geographically dispersed servers ensures ongoing backup future accessibility. See Peter Doorn and Heiko Tjalsma, "Introduction: archiving research data," *Archival Science* 7 (2007): 1-20, doi:10.1007/s10502-007-9054-6. Also, the IP2, CS06 *Cybercartographic Atlas of Antarctica* Case Study co-authored by Tracey P. Lauriault and Yvette Hackett (2005), available at http://www.interpares.org/ip2/ip2_case_studies.cfm?study=5 (accessed 18 August 2012).

3. Cybercartographic atlases as collective memory systems and community archives

While the CLI and the 1986 Domesday project exemplify the rescue and salvage efforts of legacy artefacts, there are countless examples of equally significant and technologically innovative digital maps and atlases currently being created that would be archived if a preservation infrastructure existed. The Geomatics and Cartographic Research Centre (GCRC) for instance has been actively engaged in the creation of cutting edge cybercartographic atlases for over a decade. Cybercartography “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.”³⁸ Cybercartography also offers an unprecedented opportunity for fundamentally rethinking the way we design, produce, disseminate and use maps on the Internet, both theoretically and in practice.

The GCRC has been applying the concepts of cybercartography to map traditional knowledge in Canada’s north, traditional place names with indigenous heritage and educational institutions and geonarratives³⁹ surrounding treaty processes. These atlases embody the collective memories⁴⁰ of those who have contributed to their creation and have become a means to record the historical, geographical, cultural and scientific facts that have been transmitted orally for centuries. These atlases are the first official recordings of this aurally transmitted knowledge and elders and communities have authoritatively endorsed each record. The communities who have contributed to and authorized them regard these atlases as living archives.

The Inuit siku (sea ice) Atlas⁴¹ for example was developed to respond to Inuit elders’ and hunters’ expressions of interest: to share their knowledge with youth; see more Inuit knowledge and northern content in the northern education system and to share their knowledge more broadly with scientists and the general public. The SIKU Atlas was compiled and developed to reflect the knowledge, stories, maps, language, and lessons shared through years of interviews, focus groups, sea ice trips, and workshops with local sea ice experts in Cape Dorset, Igloolik, Pangnirtung, and Clyde River, Nunavut. Interactive atlas features are used to enable students to explore and learn about various sea ice topics, maps, Inuktitut terminology, community-specific information, and project background, including audio, video, pictures, text, and maps.

³⁸ See D. R. F. Taylor, Key Note Address titled “Maps and Mapping in the Information Era,” in *Proceedings 18th International Cartographic Conference*, vol. 1, ed. L. Ottomson Stockholm (Galve: Swedish Cartographic Society, 1997), 1-10; and *The Concept of Cybercartography*, Public Lecture University of Redlands and ESRI Staff. Redlands, (2003). CA, USA.

³⁹ See Sebastien Caquard, “Cartography I: Mapping Narrative Cartography,” *Progress in Human Geography*, November 7, 2011, doi:10.1177/0309132511423796.

⁴⁰ Anthea Josias describes a number of black South African collective memory projects capturing stories in text, audio and video; remaking district maps and communities documenting apartheid histories and the post-apartheid reconciliation processes. See “Toward an understanding of archives as a feature of collective memory,” *Archival Science*, vol. 11, (2011) pp. 95–112, doi:10.1007/s10502-011-9136-3. Elizabeth Nannelli discusses the Comissão de Acolhimento, Verdade e Reconciliação de Timor-Leste (Commission for Reception, Truth and Reconciliation, or CAVR) that recorded the testimonies of East Timorese living under Indonesian rule for 25 years between 1974 and 1999. See the “Records of the Commission for Reception, Truth, and Reconciliation in Timor-Leste,” *Archival Science* 9 (2009): 29-41, doi:10.1007/s10502-009-9103-4. Like cybercartographic atlases, these projects are making records outside of traditional archives and have become key references and the means to transfer knowledge by and for the communities that have created them. Only the CAVR is formally archived.

⁴¹ The Inuit siku (sea ice) Atlas (<http://sikuatlas.ca/>) was developed as part of an International Polar Year project called The Inuit Sea Ice Use and Occupancy Project (ISIUOP), through the collaboration of many northern, academic, government, and private industry contributors. It is a compilation of Inuit sea ice knowledge and use, as documented between 2004 – 2008; it is available at <http://sikuatlas.ca/index.html> (accessed 27 August 2012).



Figure 1. The Inuit Siku (sea ice) Atlas.

The Kitikmeot Heritage Society, Inuit Heritage Trust's Traditional Name Placing Project and the Gwich'in Cultural Society among others have approached the GCRC to help them map their place name databases with their collections of audio recordings of place names and video recordings of elders narrating the stories of those places. These place name atlases⁴² have become important knowledge transmission tools from elder to youth, are cultural geo-linguistic heritage preservation tools, have been incorporated as part of school curricula and are land occupancy records depicting the territorial extent of a community's land use and settlement. Naming places is part of the infrastructure of experience⁴³ and represents social relationships, kinship, historical events and shared cultural memories. These atlases are enabling local communities to replace their histories onto the map and others to see space from a different cultural lens.

The *Cybercartographic Atlas of the Lake Huron Treaty Relationship Process* (CALHTRP)⁴⁴ on the other hand is a retelling and re-mapping of the treaty process whereby historical archival records such as surveyor notebooks and diary entries, and numerous personal histories are geo-transcribed⁴⁵ and

⁴² See the *Kitikmeot Place Name Atlas* <http://www.kitikmeotheritage.ca/atlas.htm>, *The Arctic Bay Atlas* <http://arcticbayatlas.ca/index.html> (accessed 27 August 2012) and the *Gwich'in Goonanh'kak Goonwandak: The Places and Stories of the Gwich'in* (under development).

⁴³ See Paul Dourish and Genevieve Bell, "The Infrastructure of Experience and the Experience of Infrastructure: Meaning and Structure in Everyday Encounters with Space," *Environment and Planning B: Planning and Design* 34 (2005): 414 – 430.

⁴⁴ Available at <https://gcr.ca/confluence/display/GCRCWEB/The+Cybercartographic+Atlas+of+the+Lake+Huron+Treaty+Relationship+Process> (accessed 29 August 2012).

⁴⁵ Stephanie Pyne coined this term during the making of the *Cybercartographic Atlas of Indigenous Perspectives* in, "A "living Atlas" for Geospatial Storytelling: The Cybercartographic Atlas of Indigenous Perspectives and Knowledge of the Great Lakes Region," *Cartographica* 44, no. 2 (2009): 83-100.



Figure 2. Gwichin Place Name Atlas.

aggregated into a database. This database of stories is then mapped onto old and new maps along with georeferenced photographs, historical maps, signed treaties and the historical and archival records surrounding them. In addition, some historical maps were geo-rectified according to contemporary map projections⁴⁶ and layered with the oral history of events accounted by community elders or with the geo-transcribed stories from surveyor and explorer notes. These participatory and counter cartographies are enacting a post colonial mapping of Canada's treaty system. These atlases are therefore 'remapping' the official historical record by reflexively using western geospatial technologies, multimedia and methods in such a way so as not to recolonize.⁴⁷ Furthermore, they are also providing geonarratives to the making of colonial maps, making new records from old oral traditions, and by doing so put into question the 'authenticity' and 'completeness' of the traditional archival treaty record. Also, by geo-transcribing historical records, digitizing oral cultures, and re-purposing old maps in new ways, cybercartographic atlases are giving each of these records a 'secondary provenance',⁴⁸ even though the provenance of the

⁴⁶ Recently historical maps in the David Rumsey Map Collection have been geo-rectified to be viewed along in Google Map and Google Earth, available at <http://rumsey.geogara.com/gmaps.html> (accessed 29 August 2012).

⁴⁷ For a more detailed account of this process see Stephanie Pyne and D.R. Fraser Taylor, "Mapping Indigenous Perspectives in the Making of the Cybercartographic Atlas of the Lake Huron Treaty Relationships Process: A Performative Approach in a Reconciliation Context," *Cartographica* 47, no. 2 (2012): 92-104.

⁴⁸ Lori Podolsky Nordland in "The Concept of "Secondary Provenance": Re-interpreting Ac ko mok ki's Map as Evolving Text" discusses transmedia shifts of pre-Gutenberg archival records and how these digitized records gain new life and acquire new layers of meaning. Furthermore, in the case of the Ac ko mok ki' map, it was argued that it



Figure 3. Cybercartographic Atlas of the Lake Huron Treaty Relationship Process.

original record is provided, it becomes less important as the atlases allow for the re-interpretation of original records.

Aboriginal social memory or traditional knowledge is now being recognized in the courts for evidential purposes,⁴⁹ to countervail archival sources⁵⁰ and there is growing recognition that archival records such as treaties, maps and surveys were modernist legal devices of British and French colonialists. As Raymond Frogner noted, these records represented “sovereignty’s positivisation” and were legal fictions since they purported to have been created among equals,⁵¹ which was mostly not the case. The CALHTRP in particular, puts into question the ideas of treaties as dispositive documents as the transaction did not necessarily involve willful participants.⁵²

should be reinterpreted according to the Siksika world view and their cartographic conventions thus giving it a secondary provenance, *Archivaria* 58 (2004): 147-159.

⁴⁹ A famous Canadian example is the case of the Gitksan and Wet’suwet’en First Nations who used their oral histories, which capture their geo-histories in song, performances and stories in a land rights trial against the Province of British Columbia and the federal government. They also translated their traditional knowledge into a map. Furthermore, the plaintiffs, decided that to make their claim they “had to turn the legal system, its archives, precedents, and process against itself” as they recognized they were playing in a fixed game see p. 47 of Mathew Sparke, “A Map that Roared and an Original Atlas: Canada, Cartography and the Narration of Nation,” *Annals of the Association of American Geographers* 88, no. 3 (1998): 463-495.

⁵⁰ Raymond Frogner, in “Innocent Legal Fiction: Archival Convention and the North Saanich Treaty of 1852” demonstrated “that conventional archival interpretations identify the silences and discrepancies of the textual colonial record. But critics note that conventional archival method remains tied to its textual and sovereign paradigms. It does not address the often vague and uncertain relationship between the record and the manifold power structures, cultures, and traditions that surround the record’s formation and archival disposition,” *Archivaria* 70 (Fall 2010): 45.

⁵¹ Frogner, *Archivaria* 70 (Fall 2010): 47.

⁵² Oral histories surrounding the *North Saanich Treaty* are featured by Frogner (Fall 2010), pp. 81-86 and dispositive documents are discussed on p. 48.

In these three atlas examples Aboriginal elders and communities appraised the content of the maps; and also supplied, made and curated the records; and directed how these were to be cartographically rendered in collaboration with researchers and programmers at the GCRC. These atlases have become a community archive created outside the traditional archive with the use of ‘authoritative’ maps, methods and technologies, archival records, and by georeferencing traditional knowledge from ‘authoritative’ community sources onto new and old maps. In a sense they have put into question the authenticity of what ‘official’ archival documents about these spaces purport to be and have created a counter geo-narrative.

4. The Preservation of Geospatial Data

The Government of Canada has a number of ongoing geospatial data preservation strategies as seen in Table 1.⁵³ These thematic geospatial data archives are part of science-based departments and data are collected as part of the business function and reporting responsibilities of their custodial institutions. Also, these data are collected according to established scientific data methods that are deeply ingrained practices embedded into the tools from which data are sensed, how quality is assessed, and in how these data are organized, described, formatted, disseminated and used.⁵⁴ Of all the initiatives, only the Department of Fisheries and Oceans (DFO) archived data are part of a departmental science strategy. These archives are also collaborative endeavours which contain data accessioned in partnership with others, are collected as part of a cost-recovery arrangement with other institutions, are the holdings of partner institutions, are part of major collaborative research projects, are derived from sensors housed and managed in many institutions, or are derived from a myriad private and public sector satellite networks. All but the NSDB, which contains legacy datasets, include a combination of data from active and inactive sensors, which are continuously being updated, in many cases with near-real time data. They are therefore growing collections of data that have not necessarily been set aside for disposition purposes into a traditional archive and may remain part of ongoing business practices.

Many of the archives examined disseminate and allow for the visualization of their data using specialized software while none mention the preservation of software, hardware and associated specialized file formats. These initiatives may however not be archives in the traditional sense, because they are not “an agency or institution responsible for the preservation and communication of records selected for permanent preservation” or a “place where records selected for permanent preservation are kept.”⁵⁵ It is uncertain if the institutions within which the *Earth Observation Data Services (EODS)*, *The National Soil DataBase (NSDB)* or the *Integrated Science Data Management (ISDM) Wave Data Archive* housed respectively at NRCan, AAFC or DFO, have a data preservation policy in place together with the

⁵³ The following information forms part of a larger study commissioned by GeoConnections entitled *TA 2: Final Report: Geospatial Data Archiving and Preservation* as part of the Science & Technology Policy Research and Analysis Resource Team Prepared for: NRCan, GeoConnections Operational Framework Team by Hicklings, Arthur and Low (HAL), authored by Tracey P. Lauriault and Ed Kennedy, (March 2011).

⁵⁴ The IP2 Scientific Data Portal General Study discussed the specificities of disciplinary and sub-disciplinary scientific normative practices available at http://www.interpares.org/ip2/ip2_case_studies.cfm?study=34. Also see The Focus 2 – The Sciences section of Yvette Hackett, William Underwood, and Philip Eppard, “Part One: Case and General Studies in the Artistic, Scientific and Governmental Sectors Focus Task Force Report,” *InterPARES 2: Experiential, Interactive and Dynamic Records*, ed. Luciana Duranti and Randy Preston (2008), http://www.interpares.org/display_file.cfm?doc=ip2_book_part_1_focus_task_force.pdf (accessed 29 August 2012).

⁵⁵ Def. Archive: InterPARES 2 Terminology Dictionary, http://www.interpares.org/ip2/ip2_terminology_db.cfm.

necessary human and technological resources required to ensure the permanent preservation of these data and databases. These databases may simply be backed up. The cost recovery potential of the *EODS* data would warrant investment in a good storage, retrieval and backup system, while the *NSDB* could simply be a collection of agricultural data considered important by a group of dedicated soil scientists who have the skill and will to be their custodians. The *ISDM Wave Data Archive*, may be the closest to being a ‘traditional archive’ since DFO has a *Management Policy for Scientific Data*⁵⁶ that includes a mandate, model, infrastructure, and human resources dedicated to the archiving and preservation of their data resources. It is presumed that the DFO *IOS/OSD Data Archive* would fall under the same policy.

In terms of a preservation strategy,⁵⁷ these initiatives do share characteristic elements of a preservation repository, such as:

- The means to access the data and metadata;
- Reference and context information;
- Provenance information;
- Licensing and terms of use;
- File format information; and
- In some cases are created within a data management policy (e.g., DFO).

Most geospatial data portals that store and manage the data they disseminate (e.g., *GeoBase* and *GeoGratis*) have similar characteristics. It is harder to assess file transfer mechanisms, preservation strategies such as data migration, emulation, etc., file name conventions, unique identifiers, adherence to standard vocabularies, backup schedules and the technology used from their websites.

While imperfect in terms of a traditional archival perspective, these are nonetheless starting points to build a more comprehensive geospatial data and mapping preservation strategy. These however only represent but a very small fraction of the geospatial data produced at NRCan and the rest of the Government of Canada.⁵⁸ They are also examples of proactive archiving practices—some of which are inherent in geospatial data portals and of databases considered being part of a records management process.

5. Geospatial Data Management Models

The DFO Integrated Science Data Management⁵⁹ and the International Polar Year (IPY) Data and Information Service have developed data management models, which take into consideration preservation.

⁵⁶ DFO Management Policy for Scientific Data available, <http://www.dfo-mpo.gc.ca/science/data-donnees/policy-politique-eng.htm> (accessed 22 August 2012).

⁵⁷ Defined as “a coherent set of objectives and methods for protecting and maintaining (i.e., safeguarding authenticity and ensuring accessibility of) digital components and related information of acquired records over time, and for reproducing the related authentic records and/or archival aggregations,” in Def. Records Preservation Strategy: InterPARES 2 Terminology Database available at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

⁵⁸ For a more detailed analysis of these initiative according to the 10 principles of digital preservation repositories developed by the Center for Research Libraries in Chicago see the HAL (2011) *TA 2: Final Report: Geospatial Data Archiving and Preservation*.

⁵⁹ DFO Integrated Science Data Management (ISDM), <http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/index-eng.html> (accessed 21 August 2012).

The *ISDM Wave Data Archive* (See Table 1) integrates the preservation and archiving of data from multiple organizations into an operational information management environment. The active field program of wave acquisition started in 1971, eventually discontinued in 1996, the process of submitting delayed-mode wave data has continued. In addition, ISDM acquires daily wave data from buoys operated by the Meteorological Service of Canada (MSC) at EC and manages scientific data from a wide range of sources and contains over 10 million hourly sea state and swell measurements from some 500 locations in Canada's lakes and surrounding oceans. While these data are being used in real-time to produce weather and sea state forecasts, its archiving program ensures that the information can be used in a variety of applications requiring data over long timeframes, such as hindcast models of wave climatology used in ocean maritime navigation, engineering and climate change studies.

Table 1. Government of Canada geospatial data preservation initiatives

IOS/OSD Data Archive , Department of Fisheries and Oceans (DFO)	National Climate Data and Information Archive , Environment Canada (EC)	Integrated Science Data Management (ISDM) Wave Data Archive , DFO
Lithoprobe Data Archive , NRCan, Geological Survey of Canada (GSC)	The Canadian Ice Service Archive (CISA) , EC, Canadian Ice Service	National WaveForm Archive (NWFA) , NRCan, Earthquakes Canada
The National Soil DataBase (NSDB) , Agriculture and Agri- Food Canada (AAFC); Canadian Soil Information Service (CANSis)	Earth Observation Data Services (EODS) , Natural Resources Canada (NRCan), Canada Centre for Remote Sensing (CCRS)	National Water Data Archive , EC, Water Survey of Canada (WSC) w/Archived Sediment Data
Geomagnetism Summary Plots: Archives , NRCan, GSC	System of Agents for Forest Observation Research with Advanced Hierarchies (SAFORAH) , Canadian Forest Service (CFS), University of Victoria and other academic and government partners.	

The archiving and preservation commitment extends across all kinds of scientific data for which DFO is responsible, as evidenced by the department's *Management Policy for Scientific Data*,⁶⁰ which came into effect in June 2001. This IM Policy includes several references to the requirement for data archiving and preservation and it adheres to the following principles:

1. DFO scientific data sets... are irreplaceable, and must be protected and managed to ensure long-term availability;
2. ... it is essential that DFO Science/Oceans maintain responsibility for their quality control, management, archiving and dissemination; and
3. ... all scientific data collected by the DFO must be migrated to a 'managed' archive immediately after the data have been processed.

⁶⁰ DFO, *Management Policy for Scientific Data*, (2001), <http://www.dfo-mpo.gc.ca/science/data-donnees/policy-politique-eng.htm> (accessed 21 August 2012).

In terms of archiving, all DFO scientific data must be managed as part of an integrated system accessible through regional, zonal and national data centres and the responsibilities of the integrated system of data centres will be to:

- Ensure long-term accessibility and documentation in the event of organizational changes, retirements, etc.;
- Protect data against loss resulting from error, accident, technological change, degradation of media, etc.

When data are submitted, the DFO stresses the importance of ensuring that data are quickly migrated into a 'managed' environment where they are properly backed up and secured from accidental or circumstantial loss, and where the supporting metadata are integrated with the data to preserve their long-term usefulness. Finally the DFO includes a data rescue program to locate and preserve scientific data that are of value to departmental programs and to identify data at risk.

The International Polar Year (IPY) project, on the other hand, was a large three-year Arctic and Antarctic scientific program, organized through the International Council for Science (ICSU) and the World Meteorological Organization (WMO). The IPY Data and Information Service (IPYDIS)⁶¹ was a global partnership of data centers, archives, and networks working to ensure the proper stewardship of IPY and related data. The National Snow and Ice Data Center (NSIDC) at the University of Colorado is a coordination office for the IPYDIS and it ensures the long-term preservation and access to IPY data. The policy⁶² states that:

...it is essential to ensure long-term preservation and sustained access to IPY data. All IPY data must be archived in their simplest, useful form and be accompanied by a complete metadata description. An IPY Data and Information Service (IPYDIS) should help projects identify appropriate long-term archives and data centers, but it is the responsibility of individual IPY projects to make arrangements with long-term archives to ensure the preservation of their data. It must be recognized that data preservation and access should not be afterthoughts and need to be considered while data collection plans are developed.

The IPY exemplifies a distributed or "virtual" archive of scientific data and proactive archiving. The Service's Web site guides digital scientific data users to a number of portals and centers that are currently providing access to IPY and related data and all data registries and repositories are required to adhere to the IPY Metadata Profile requirements. This profile was based on the Global Change Master Directory (GCMD) Directory Interchange Format, and is compliant with and has been mapped to recognized geospatial metadata standards.⁶³

⁶¹ IPY, *International Polar Year Data and Information Service (IPYDIS)*, Ironically the IPYDIS site is not longer supported but copies of key documents are available from Mark A. Parsons at parsonsm@nsidc.org.

⁶² IPY, (2008), *International Polar Year 2007-2008 Data Policy* is available at http://classic.ipy.org/Subcommittees/final_ipy_data_policy.pdf (accessed 21 August 2012).

⁶³ The standards are: Global Change Master Directory (GCMD) Directory Interchange Format (DIF); the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) (FGDC-STD-001-1998) and Remote Sensing Extensions (RSE) (FGDC-STD-012-2002); and THREDDS Dataset Inventory Catalog Specification Version 1.0. Other international standards are recommended such as the Open Geospatial Consortium (OGC) standards, and the ISO 19115:2003 metadata standard. See IPY Data and Information Management Service, *IPY Metadata Profile version 1.0*, available from Mark A. Parsons parsonsm@nsidc.org.

The DFO and IPY policies, practices and initiatives, like the archiving examples just discussed above, are good places to start when planning a strategy to preserve geospatial data and maps.

6. Canadian digital data and information consultations, studies, reports and initiatives

The above-mentioned geospatial data archives and data management models occurred in spite of the paucity of tangible outcomes from the 12 consultations, studies, reports and initiatives as seen in figure 4 below.

The *Research Data Strategy*⁶⁴ is one of the tangible outcomes of the 2002 *NCSARD* report process that will be discussed shortly. It is a collaborative effort between government and academia aiming to address the challenges and issues surrounding access to and preservation of data created by Canadian researchers. It is a multidisciplinary group of universities, institutes, libraries, granting agencies, and individual researchers, including several officials involved with the creation, management, dissemination and/or funding of geospatial data,⁶⁵ who recognize the need to address data management issues. While the focus is on research data, the ideas are transferable to the preservation and management of geospatial data. The approach and principles to data stewardship recommended by *Research Data Canada* have been echoed in many other initiatives discussed here, namely that data stewardship and management needs to be framed in a life cycle model at all stages of the data production process and making data

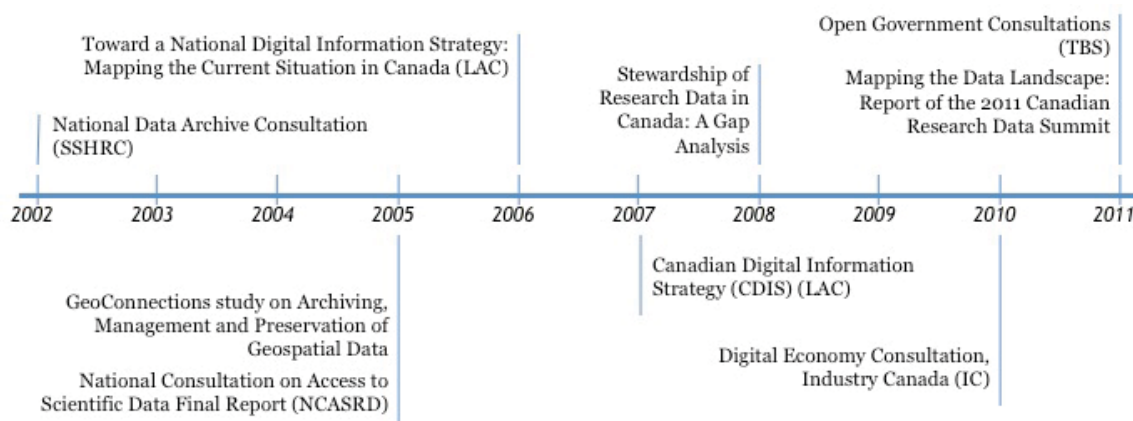


Figure 4. Canadian Digital Data and Information Consultations, Studies, Reports and Initiatives.

preservation a formal institutional responsibility.⁶⁶

Two international groups hold potential for future guidance in geospatial data archiving and preservation. The Committee on Data for Science and Technology (*CODATA*) *Working Group on*

⁶⁴ Available at the Research Data Strategy home page, <http://rds-sdr.cisti-icist.nrc-cnrc.gc.ca/eng/> (accessed 27 August 2012).

⁶⁵ John Broome, the Chair CNC/CODATA and Head, Data Management Policy and Strategy, Earth Sciences Sector, Natural Resources Canada; Scott Tomlinson, Data Management Coordinator, International Polar Year Federal Program and Chuck Humphrey, Data Library Coordinator, Libraries, University of Alberta.

⁶⁶ Research Data Canada, <http://ncasrd-cnads.scitech.gc.ca/eng/about/principles-data-stewardship.html>.

*Archiving Scientific Data*⁶⁷ has been holding symposia and workshops on the topic and the *Canadian National Committee for CODATA*⁶⁸ has been active in documenting and reporting scientific data activities. Members of these include geospatial data experts. The Open Geospatial Consortium (OGC) also has a *Data Preservation WG*,⁶⁹ which aims “to address technical and institutional challenges posed by data preservation, to interface with other OGC working groups that 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.” Reports from this WG are not yet available.

In terms of consultation, only the 2011 TBS *Open Government Consultations*⁷⁰ yielded a report that informed an actual plan⁷¹ that was released at the Open Government Partnership in Brazil⁷² in the spring of 2012. The plan includes better record keeping and the lifting of restrictions to archived material at LAC but it does not include the preservation of data in the TBS *Open Data Pilot Project* portal nor any kind of preservation strategy for maps and data produced by government more broadly. The TBS Consultation report and the Plan also make no reference to preservation and archiving, LAC was not one of the ‘thought leaders’ to make a public submission and the complete list of individuals and organizations who submitted has not been made public.⁷³ Furthermore, the Advisory Panel on Open Government⁷⁴ does not include a representative from LAC, libraries, archives or science although the President and Founder of ESRI Canada is a member. Interestingly, the most popularly requested datasets in the TBS Consultation were geospatial and geospatial datasets predominate in terms of number of datasets in the *Open Data Pilot Project* portal.⁷⁵

The *Mapping the Data Landscape: Report of the 2011 Canadian Research Data Summit*⁷⁶ was the result of National Research Council (NRC) of Canada 2011 Canadian Research Data Summit held in September of 2011 and summit partners included Canadian IPY members.⁷⁷ The report is currently being circulated to discuss the *Research Data Strategy*.

The 2010 IC *Digital Economy Consultation*⁷⁸ did not result in any consolidated report or actions. The IC *Consultation Paper*⁷⁹ circulated to frame the consultation, however, did mention improving access

⁶⁷ *CODATA Preservation of and Access to Scientific and Technical Data in Developing Countries*, 2006, Home Page, <http://www.codata.org/taskgroups/TGpreservation/> (accessed March 2011).

⁶⁸ *Canadian National Committee for CODATA*, <http://dac.cisti.nrc.ca/> (accessed March 2011).

⁶⁹ *OGC, Data Preservation WG*, <http://www.opengeospatial.org/projects/groups/preservwg>) (accessed March 2011).

⁷⁰ Treasury Board Secretariat of Canada, December 6, 2011 to January 16, 2012, *Open Government Consultation Report* available at <http://www.open.gc.ca/consult/wwh-cr-eng.asp> (accessed 1 July 2012).

⁷¹ TBS Open Government Plan, released April 12, 2012, <http://www.tbs-sct.gc.ca/media/nr-cp/2012/0412-eng.asp#back1> (accessed 26 June 2011).

⁷² Available at <http://www.opengovpartnership.org/> (accessed 29 August 2012).

⁷³ Lauriault included preservation in her submission posted on the datalibre.ca blog on January 12, 2012, <http://datalibre.ca/2012/01/16/tbs-open-data-and-open-government-consultation-response/> (accessed 26 June 2011).

⁷⁴ Open Government Advisory Panel, <http://www.open.gc.ca/open-ouvert/bio-bio-eng.asp>.

⁷⁵ When the *Open Data Pilot Project* portal was launched, IASSIST reported that there were “782 general datasets (e.g., not geospatial) and over 260,000 geospatial datasets” available at <http://www.iassistdata.org/resources/canadas-open-data-pilot-project> (accessed 26 June 2011).

⁷⁶ The report is available at rds-sdr.cisti-icist.nrc-cnrc.gc.ca/docs/data_summit-sommet_donnees/Data_Summit_Report.pdf (accessed 27 August 2012).

⁷⁷ See the report background material available at http://rds-sdr.cisti-icist.nrc-cnrc.gc.ca/eng/events/data_summit_2011/index.html (accessed 28 August 2012).

⁷⁸ Industry Canada, (2010), *Digital Economy Consultation*, available at http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/h_00491.html (accessed 27 August 2012).

to research data, data discovery and open access publishing but did not explicitly mention the geomatics sector or geospatial data. A cursory examination of submissions and the idea forum⁸⁰ called for the creation of a digital data and information archive and a network of trusted digital repositories.⁸¹ The summary of key findings⁸² however failed to mention these calls.

The *Stewardship of Research Data in Canada: A Gap Analysis*⁸³ was conducted by members of *Research Data Canada*. The life-cycle model was used as the framework to analyse the current state of research data stewardship in Canada, which includes: data production, dissemination, long-term preservation and data discovery and repurposing. The *Gap Analysis* reviewed: policies; funding; roles and responsibilities; repositories; standards; skill and training; rewards and recognition systems; R&D; access; and preservation. Not surprisingly, the analysis concluded that most research data created in Canada are greatly underutilized and are at a high risk of being lost. Geospatial and research data share similar characteristics, namely: they are complex, are in multiple data formats, use non standardized rendering techniques and specialized open and/or proprietary software, are disseminated in portals, or are shared and visualized in distributed systems, datasets can be very large and data can be modeled. Geospatial initiatives, unlike research data do have uniform metadata and data quality standards are normalized.

The 2007 LAC *Canadian Digital Information Strategy (CDIS)*⁸⁴ report reflects the broad consensus reached at a *National Summit*⁸⁵ held in 2006 which included contributions from more than 200 specialists some of whom are involved in geospatial data creation, use, dissemination and management. There are numerous suggested actions listed, none of which are specific to geospatial data, although GeoConnections, GeoGratis, the Canada Ice Service Data Archive and GeoBase are lauded for some of their initiatives. Environment Canada, Statistics Canada, DFO and NRCan are mentioned as institutions that produce and hold significant collections of government data. Some of the preservation assumptions upon which the *CDIS* was premised mirror contemporary geospatial data productions practices and issues. For example: technological change is constant; a distributed, interoperable, standards based, open and

⁷⁹ Industry Canada, 2010, *Improving Canada's Digital Advantage Strategies for Sustainable Prosperity: Consultation Paper on a Digital Economy Strategy for Canada*, accessed June 26, 2012

http://publications.gc.ca/collections/collection_2010/ic/Iu4-144-2010-eng.pdf.

⁸⁰ The most popularly voted idea was for access to and the preservation of government data assets, see

<http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00172.html> (accessed 29 August 2012).

⁸¹ See the following submissions by Canadian Association of Research Libraries

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00349_4.html, universities

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00357_4.html,

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00402_4.html, two individuals

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/h_00491.html,

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00284_3.html, institutes, federations, associations and council

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00398_4.html,

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00418_4.html,

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/h_00491.html, and

http://www.digitaleconomy.gc.ca/eic/site/028.nsf/eng/00440_4.html, (accessed 21 August 2012).

⁸² Summary of key findings is available at <http://www.ic.gc.ca/eic/site/sittgateway-portailstit.nsf/eng/00055.html> (accessed 28 August 2012).

⁸³ Research Data Canada, (2008), *Stewardship of Research Data in Canada: A Gap Analysis*, available by contacting Bronwen.Woods@nrc-cnrc.gc.ca.

⁸⁴ Library and Archives Canada (2007), *Canadian Digital Information Strategy*, available at <http://www.lac-bac.gc.ca/obj/012033/f2/012033-1000-e.pdf> (accessed February 2011).

⁸⁵ Library and Archives Canada (2006), *Toward a Canadian Digital Information Strategy: National Summit*, accessed February 2011 <http://www.collectionscanada.gc.ca/obj/012033/f2/012033-611-e.pdf>.

accessible strategy is key; stakeholders and communities of practice input are essential. The proposed actions could be incorporated into existing Government of Canada data portals.

The LAC 2006 *Toward a National Digital Information Strategy: Mapping the Current Situation in Canada*⁸⁶ states that “the stewardship of digital information produced in Canada is disparate and uncoordinated” and “the volume, diversity and complexity of digital information is growing exponentially” while the “technologies, standards and practices that will better ensure the ongoing accessibility and integrity of digital information are not yet consistently applied.” The report includes a brief description of the GeoConnections program and lists a number of other geospatial data initiatives at different levels of government in Canada in its Annexes. The report does not provide specific recommendations on how to manage these resources but it does suggest that the:

...preservation of digital information requires a management response, not just a technological response. The response must be active and sustained through time and it must address the preservation of digital information from both the strategic and tactical perspectives and within the context of the management frameworks that govern the businesses of organizations themselves. It must also be comprehensive and address all facets of the required preservation infrastructure: the policies that assign accountability, the standards, practices, procedures and technologies that enable the implementation of preservation strategies, and the people that make it all happen. These issues present a tremendous challenge and must be addressed in an inclusive and collaborative manner.⁸⁷

MacDonald and Shearer also provide a number of next steps toward the creation of a national digital information strategy, which resemble the collaborative strategies used to create the CGDI (e.g., multisectoral, departmental and multidisciplinary collaboration, etc.).

In 2005 GeoConnections conducted a study on *Archiving, Management and Preservation of Geospatial Data*⁸⁸ which provided a well rounded analysis of geospatial data preservation issues, such as: technological obsolescence; formats; storage technologies; temporal management; and metadata.⁸⁹ The study also provided a list of technological preservation solutions with their associated advantages and disadvantages and recommended “as the first step in ensuring the long-term preservation and retention of valuable resources, data producers must adopt an information life cycle management approach, which will ensure that their data will be managed proactively from creation to disposition.”⁹⁰ The recommendations are somewhat outdated in light of institutional changes at NRCan and of the new TBS Directives and Guidelines, irrespective, the study provides the following useful recommendations:

⁸⁶ MacDonald, J. and K. Shearer (2006), *Toward a National Digital Information Strategy: Mapping the Current Situation in Canada* (Ottawa: Library and Archives Canada), accessed February 2011 Summary: <http://www.lac-bac.gc.ca/digital-initiatives/012018-3200-e.html> and full Final Report: <http://www.collectionscanada.gc.ca/obj/012018/f2/012018-3200-e.pdf> (accessed 29 August 2012).

⁸⁷ See MacDonald and Shearer (2006), p. 50.

⁸⁸ GeoConnections (2005). *Archiving, Management and Preservation of Geospatial Data: Summary Report and Recommendations*, produced by David L. Brown, Grace Welch and Christine Cullingworth as an initiative of the Policy Advisory Node.

⁸⁹ The approach adopts the framework created by Bleakly, Denise R. (2002). *Long-Term Spatial Data Preservation and Archiving: What are the Issues?* Sand Report, SAND 2002-0107. Albuquerque, New Mexico: Sandia National Laboratories available at <http://prod.sandia.gov/techlib/access-control.cgi/2002/020107.pdf> (accessed 28 August 2012).

⁹⁰ GeoConnections (2005), *Archiving, Management and Preservation of Geospatial Data*, Page i.

- Organizations should define and implement policies and practices for the creation, use, retention, dissemination, preservation, and disposition of geospatial data. Building a business case for the creation of core geospatial data products is suggested.
- Organizations must establish authoritative responsibility centers that empower individuals with the ability to define and apply the information management principles required to ensure the integrity of an organization's geospatial data holdings. A custodianship model is an option, which provide a means to facilitate data management on the behalf of creators and users, and provide continuity in the delivery of a geospatial data infrastructure.
- Geospatial data preservation issues fall within the realm of a national information policy, and a national data management strategy. Working in partnership with the library and archives communities, government data producers need to standardize and adopt organizational policies and practices to govern the creation, use, retention, dissemination, preservation, and disposition of geospatial data to ensure its authenticity and integrity for as long as it is required for legislation, departmental statutes and other laws and policies. Key geospatial advisory boards, councils, and standards organizations should work with LAC to develop a series of geospatial data policies and practice.
- The Canadian Council on Geomatics (CCOG) should create a task force and invite interested stakeholders from the academic community, private sector and other federal and provincial/territorial agencies that can collaborate to develop priority policy areas identified above.

The 2005 *National Consultation on Access to Scientific Data Final Report (NCASRD)*,⁹¹ developed in partnership with the NRC, the Canada Foundation for Innovation (CFI), Canadian Institutes of Health Research and Natural Sciences and Engineering Research Council (NSERC), expressed the concern that: “no national data preservation organization exists, nor does Canada have any national data access strategy or policies. NCASRD participants expressed considerable concern about the loss of data, both as national assets and definitive longitudinal baselines for the measurement of changes over time.”⁹² The *NCSARD* report provides a comprehensive list of recommendations that include organizing a Data Force and the creation of a Data Canada secretariat.

The 2002 *SSHRC National Data Archive Consultation*⁹³ is *NSCARD*'s predecessor, primarily discusses data in the social sciences and humanities and the preservation of data created in the course of state funded research projects. The *Consultation* identified important institutions, infrastructures, management frameworks and data creators and calls for the creation of a national research data archive.

To date not much has happened in terms of implementation. Currently, LAC does not have a digital geospatial data archive and there is no national network of trusted digital repositories with the explicit mandate to ingest geospatial data, research, government or private sector created data. In light of recent budget cuts at LAC, proactive archiving and records management might be the best interim solution.

⁹¹ F. D. Strong and P. B. Leach (2005). *The Final Report of the National Consultation on Access to Scientific Data*, Ottawa: Government of Canada, it is available by contacting Bronwen.Woods@nrc-cnrc.gc.ca.

⁹² Strong, and Leach (2005), *The Final Report of the NCASRD*, page 2.

⁹³ Social Science and Humanities Research Council (2002), *Final Report: National Consultation on Research Data Archiving, Building Infrastructure for Access to and Preservation of Research Data*, http://www.sshrc-crsh.gc.ca/about-au_sujet/publications/da_finalreport_e.pdf (accessed 29 August 2012).

7. Legislation, Directives and Policies

It is beyond the scope of this paper to discuss all applicable legislation, directives and policies. Tables 2 and 3 list the overarching legislation, regulation, directives and policies which apply to geospatial data and Table 4 includes NRCan legislation that explicitly mention geospatial data, maps, note books, surveys or software that are to be preserved or are considered to be official government records. For a more detailed analyses of these please refer to the 2011 *TA 2: Final Report: Geospatial Data Archiving and Preservation*.

The Library and Archives Act (Table 2) is specific to LAC, while the others direct how Government of Canada information is to be managed and disseminated irrespective of the creating institution. These acts and regulations govern the preservation, management and dissemination of geospatial data, databases and maps, which, irrespective of their form are government information. These acts and regulations provide the context within which TBS and departmental guidelines, directives, and policies related to the management of government information are created.

Table 2. Overarching legislation and regulation that apply to geospatial data.

The Library and Archives of Canada Act (2004, c. 11)	Personal Information Protection and Electronic Documents Act (2000, c. 5)
<i>Copyright Act</i> (R.S., 1985, c. C-42)	Canada Evidence Act (R.S., 1985, c. C-5)
Access to Information Act (R.S., 1985, c. A-1)	Legal Deposit of Publications Regulations (SOR/2006-337)
<i>Privacy Act</i> (R.S., 1985, c. P-21)	Privacy Regulations (SOR/83-508).

Table 3 lists information management policies, directives, guidelines and standards at the federal level in Canada related to data archiving preservation in general and some refer to geospatial information specifically. The primary responsibility for such policies lies with the TBS. The *Standard on Geospatial Data* is the only document that addresses geospatial information domain. Its objective is “to support stewardship and interoperability of information by ensuring that departments access, use and share geospatial data efficiently and effectively to support program and service delivery.” It does not explicitly reference digital information archiving and preservation, although there is a linkage between it and LAC’s *File Format Guidelines for Preservation and Long-term Access*,⁹⁴ which recommends geospatial metadata standards for the preservation of and long-term access to digital geospatial information held by government organizations.

⁹⁴ Library and Archives Canada, *File Format Guidelines for Preservation and Long-term Access*, available at <http://www.collectionscanada.gc.ca/digital-initiatives/012018-2200-e.html> accessed March 2011

Table 3. Information policies, guidelines and standards.

TBS Directive on Recordkeeping ⁹⁵	Policy on Management of Information Technology ⁹⁶
TBS Directive on Information Management Roles and Responsibilities ⁹⁷	TBS Multi-Institutional Disposition Authority (MIDA) ⁹⁸
TBS Standard for Electronic Documents and Records Management Solutions (EDRMS) ⁹⁹	LAC Guidelines: Local Digital Format Registry (LDFR) ¹⁰⁰
TBS Policy on Information Management ¹⁰¹	TBS Standard on Geospatial Data ¹⁰²
TBS Policy Framework for Information and Technology ¹⁰³	TBS Standard on Metadata ¹⁰⁴
Forthcoming Methodology associated with TBS Directive on Recordkeeping, which NRCan Records Management Group contributed.	

Table 4 lists acts and regulations that specifically address geospatial information and are under the responsibility of the Minister of Natural Resources.¹⁰⁵ The *Remote Sensing Space Systems Act* and the *Remote Sensing Space Systems Regulations* were added since they impact how the Canada Centre for Remote Sensing (CCRS) manages its geospatial data sets, while the *Charts and Nautical Publications Regulations* addresses marine geospatial information. Acts and regulation mention raw data, how these are to be processed, the systems used to view and process the data and the use of cryptography. There are also mentions of registrars and catalogs. In some there are specific references to archiving, deposition, data protection plans, the number of years data are to be kept and data curation. Finally, because these geospatial datasets are critical to the management of Canada's boundaries, the health and safety of Canadians, Canada's economy and the attestation of lands, these government records can all be called into evidence to resolve international and national disputes, or any other judicial proceeding. Clearly, the geospatial datasets mentioned here must be well managed once they are created and should be preserved accordingly. The Government of Canada produces geospatial data, databases and maps in many other

⁹⁵ See <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=16552>

⁹⁶ See <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=12755>

⁹⁷ See <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=12754>

⁹⁸ See <http://www.collectionscanada.gc.ca/government/disposition/007007-1062-e.html> and <http://www.collectionscanada.gc.ca/government/disposition/index-e.html>

⁹⁹ See <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18910§ion=text>

¹⁰⁰ See <http://www.collectionscanada.gc.ca/digital-initiatives/012018-2200-e.html>

¹⁰¹ See <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=12742>

¹⁰² See Treasury Board of Canada Secretariat, (2009). *Standard on Geospatial Data*, <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?evttou=X&id=16553§ion=text> (accessed 29 August 2012)

¹⁰³ See <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=12452§ion=text>

¹⁰⁴ Treasury Board of Canada Secretariat, (2010), *Standard on Metadata*, <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18909§ion=text> (accessed 29 August 2011).

¹⁰⁵ Natural Resources Canada, (modified 2010), *List of acts for which the Minister of Natural Resources is responsible*, <http://www.nrcan-rncan.gc.ca/com/resoress/actacte-eng.php>, (accessed 29 August 2012).

departments and agencies and each of these would have to adhere to acts and regulations for which they are responsible along with those in Tables 2 and 3.

Table 4. Acts and regulation governing geospatial data.

Department of Natural Resources Act (1994, c. 41)	Energy Monitoring Act (R.S., 1985, c. E-8)
Resources and Technical Surveys Act (R.S., 1985, c. R-7)	National Energy Board Act (R.S., 1985, c. N-7) and National Energy Board Electricity Regulations (SOR/97-130)
Canada Lands Surveys Act (R.S., 1985, c. L-6)	Northern Pipeline Act (R.S., 1985, c. N-26)
Forestry Act (R.S., 1985, c. F-30) and Regulations Respecting the Report on the State of Canada's Forests (SOR/95-479)	Nuclear Energy Act (R.S., 1985, c. A-16)
Canada Oil and Gas Operations Act (R.S., 1985, c. O-7) and Canada Oil and Gas Geophysical Operations Regulations (SOR/96-117)	Nuclear Fuel Waste Act (2002, c. 23)
Canada Petroleum Resources Act (1985, c. 36 (2 nd Supp.))	Nuclear Safety and Control Act (1997, c. 9)
Canada-Newfoundland Atlantic Accord Implementation Act (1987, c. 3)	Uranium Mines and Mills Regulations (SOR/2000-206)
Canada-Nova Scotia Offshore Petroleum Resources Accord Implementation Act (1988, c. 28)	Remote Sensing Space Systems Act (2005, c. 45) and Remote Sensing Space Systems Regulations (SOR/2007-66)
Energy Efficiency Act (1992, c. 36)	Charts and Nautical Publications Regulations (SOR/95-149)

The legislation reviewed in the HAL 2011 Report, demonstrated the diversity of types of geospatial data and maps, their forms, formats and associated software and systems, how they are managed, the context within which they are created, and how they are considered authentic. Also, many geospatial datasets are very highly regulated. This analysis, combined with an examination of data IP2 science data portals below, reaffirms the heterogeneous nature of geospatial data and maps and the importance of creators to be involved in the process of preserving geospatial data assets.

8. LAC Guidelines pertaining to Geospatial Data and Maps

Few LAC guidelines apply to cartographic material. The 2011 *Managing Cartographic, Architectural and Engineering Records in the Government of Canada*¹⁰⁶ mentions that geomatics records include systems, discs, CD-ROMs and other cartographic material in electronic formats. However, it refers to the

¹⁰⁶ Library and Archives Canada (LAC), *Managing Cartographic, Architectural and Engineering Records in the Government of Canada*, Ottawa: Government of Canada, <http://www.collectionscanada.gc.ca/government/002/007002-2050-e.html> (accessed 29 August 2012).

2001 Canadian Committee on Archival Description (CCAD) *Rules for Archival Description Chapter 5*,¹⁰⁷ which primarily address paper maps although general issues pertaining to digital databases and programs are covered in the 2003 *Chapter 9: Records in Electronic Form*.¹⁰⁸ The LAC Local Digital Format Registry (LDFR) *File Format Guidelines for Preservation and Long-term Access Version 1.0*¹⁰⁹ includes *Geospatial* guidelines and recommends TC 211 ISO 19115 Geographic Information—Metadata and input from digital geospatial data record creators and managers is solicited. These guidelines are, however, inadequate for the kind of complex geospatial artefacts discussed in InterPARES 2 Project (IP2) or for cybercartographic atlases.

9. Research on the preservation of geospatial data in Canada

The following discusses empirical examples from a selection of InterPARES 2 Project¹¹⁰ Case Studies¹¹¹ (Table 5) that study geospatial data, and part of the IP2 General Study 10 (GS10),¹¹² which examined geospatial data portals (Table 6).

The IP2 Case Studies in Table 5 below include observational data in a variety of forms, while five also included computational data or models. All but the Cybercartographic Atlas¹¹³ render or store their data in a proprietary system, and data are stored in a variety of databases or in a searchable data portal while Engineering project data¹¹⁴ (CS19), the Land Registry¹¹⁵ (CS18), and some of the NASA data (CS08) are inseparable from the systems within which they have been created and/or stored. As discussed earlier when examining the CLI and the 1986 Domesday project, these findings are somewhat discouraging.

The IP2 Case Study questions regarding authenticity¹¹⁶ revealed that the prevalent concept of reliability is closely tied to reproducibility and accuracy in the physical sciences. In many cases, reliability is associated with faith in the technological systems in place or security measures related to access to the system. In terms of authenticity, it was revealed that Case Study respondents¹¹⁷ do not think

¹⁰⁷ Canadian Committee on Archival Description, (2001), *Rules for Archival Description Chapter 5 Cartographic Materials*, http://www.cdncouncilarchives.ca/rad_ch5.pdf (accessed February 2011).

¹⁰⁸ Canadian Committee on Archival Description, (2003), *Rules for Archival Description Chapter 9 Records in Electronic Form*, (accessed February 2011). http://www.cdncouncilarchives.ca/RAD_chap9_revised_Aug2003.pdf.

¹⁰⁹ LAC, *File Format Guidelines for Preservation and Long-term Access Version 1.0*, <http://www.collectionscanada.gc.ca/digital-initiatives/012018-2200-e.html>. (accessed February 2011).

¹¹⁰ InterPARES http://www.interpares.org/ip2/ip2_index.cfm (accessed 6 September 2007).

¹¹¹ IP2 Case Studies, http://www.interpares.org/ip2/ip2_case_studies.cfm (accessed 29 August 2012).

¹¹² IP2, General Study 10 Report and associated documents are available at http://www.interpares.org/ip2/ip2_case_studies.cfm?study=34 (accessed 21 August 2012).

¹¹³ Tracey P. Lauriault and Yvette Hackett, *CS06 Cybercartographic Atlas of Antarctica* (Vancouver, 2005).

¹¹⁴ Kenneth Hawkins, *CS19 Authenticating Engineering Objects for Digital Preservation* (Vancouver, 2005).

¹¹⁵ Jean-François Blanchette, Françoise Banat-Berger and Geneviève Shepherd, *CS18 Computerization of Alsace-Moselle's Land Registry* (Vancouver, 2004).

¹¹⁶ See Table 1: Authenticity Statements in the supplementary file on the IP website at http://www.interpares.org/display_file.cfm?doc=Archivaria64_Todays_Data_supplementary_tables.pdf (accessed 29 August 2012).

¹¹⁷ Trust in data sources was found to be important in both the Cybercartographic Atlas of Antarctica (CS06) and the Archeology records study (CS14), while technological integrity, validity procedures and system checks were implemented in the Mars Global Surveyor Data Study (CS08), the Computerization of Alsace-Moselle's Land Registry (CS18) and the Most Satellite Mission project (CS26). Some of the studies controlled access to their datasets via the appointment of responsible agents (VanMap and the Alsace-Moselle Land Registry), or specialists (Cybercartographic Atlas), while peer

of authenticity in the same way that archivists do, as the emphasis leans more towards measures to ensure data quality. The Land Registry (CS18) may be the exception, as the system is specifically designed to ensure that each registration is authenticated in the system.

Table 5. InterPARES 2 Geospatial Data Case Studies.

Case Study Title	Description	Observational Data	Computational Data
<i>CS06 - Cybercartographic Atlas of Antarctica</i>	Online interactive and dynamic, open standards, interoperable multimedia, multisensory, multimodal atlas that renders distributed data from myriad scientific organizations.	Distributed observational data from myriad sources in real time, film, satellite images, tabular data, sounds, etc.	Data are rendered /refined into maps, charts, tables by the Nunaliit Atlas Framework
<i>CS08 - Mars Global Surveyor Data Records in the Planetary Data System</i>	Surveyor mission data records at the Planetary Data System (PDS) Space Science Data Archive.	Level 0 and refined planetary spacecraft mission data stored in a database with attributes and metadata	Software used to access or visualize the data
<i>CS14 - Coalescent Communities in Arizona</i>	Archaeological Records of the American Southwest rendered in a GIS.	Tabulated raw data collected from archaeological digs	Raw data are rendered/refined into maps
<i>CS18 - Computerization of Alsace-Moselle's Land Registry</i>	Electronic registry including digital transcription of 40 000 existing paper registries and new database entries individually signed by a judge using a PKI infrastructure combining biometric access and digital signatures.	Database of digitized paper land registries and attributes	Data and their attributes are accessed via a data base using PKI technology
<i>CS19 - Authenticating Engineering Objects for Digital Preservation</i>	Examines through an engineering experiment the authentication of digital model (CAD) records using a content/message/semantic-based methodology rather than media, bit-count, or static provenancial attribute-based authentication.	CAD solid model files used in the design and manufacturing of mechanical piece-part assemblies	The files are rendered and stored in a proprietary system
<i>CS24 - City of Vancouver Geographic Information System (VanMap)</i>	An enterprise web-based map system maintained by the City of Vancouver's Information Technology Department.	Land use, social statistics, city infrastructure data both internally collected and acquired from external sources etc.	Enterprise GIS renders data into maps, charts, and tables etc.
<i>CS26 – MOST Satellite Mission: Preservation of Space Telescope Data</i>	Repository of the Microvariability & Oscillations of Stars satellite mission data of Canada's first space telescope.	Raw satellite data and their refined counterparts	

review of data was a method used in the Mars Global Surveyor Data study (CS08). In the Engineering Objects Study (CS19), the creators did not believe their records to be authentic as they have no assurance system.

Portals, it was discovered during the IP2 science case studies, provide a framework within which geospatial data archivists can work and from which they can expand with policies, standards and metadata. Especially since, data producers and users have already appraised the data these contain or refer to as being valuable enough to be paid for, collected, described, licensed, endorsed, organized and disseminated. Also, geospatial data are increasingly being discovered and accessed via data portals. Portals make it possible for users to “gather data germane to their own needs more readily, extract data from online and other electronic repositories, develop the information product they need, use the products for decision making, and contribute their locally gathered geoinformation and derived products to libraries or other repositories.”¹¹⁸

As is commonly known in the field of geomatics, 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, analyse and/or manipulate them while the reasons for portal creation are one-stop-shopping, distributed responsibility over data sets, discoverability, and reduction in cost as data are stored or described once and used many times.¹¹⁹

Furthermore, portals are the technical embodiment of data-sharing policies. Individuals within organizations, research projects, or scientific collaborations register their data holdings in portals via an online form organized according to a metadata standard, and then choose to make their data available for free or for sale, viewing or downloading.¹²⁰ Metadata standards “establish the terms and definitions to provide a consistent means to describe the quality and characteristics of geospatial data,”¹²¹ and the ISO 19115 metadata¹²² standard has become an international standard in the field of geomatics as have those mentioned earlier in the case of the IPY. Most of the portals examined in IP2’s General Study 10 include either very detailed metadata or rudimentary header information that contains lineage information.¹²³

The architecture of data portals varies¹²⁴ and portals can be 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). 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 agent¹²⁵ to access or render the data into a map or some other form. Many maps and atlases for example, adhere to interoperable OGC standards and specifications, which allow for distributed mapping. In the case of *Cybercartographic Atlas of Antarctica*, for example, users access a module and a call is made to the British Antarctic Survey data portal in the United Kingdom and these are then rendered into a map in real time by the Atlas Framework in Ottawa and delivered directly to the user’s computer. Other examples of

¹¹⁸ US National Research Council, (1999), Spatial Information Resources, *Distributed Geolibraries* (Washington DC), (<http://www.nap.edu/openbook.php?isbn=0309065402>), p. 36. (accessed 29 August 2011).

¹¹⁹ See Tracey P. Lauriault, (2003). A Geospatial Data Infrastructure is an Infrastructure for Sustainable Development in East Timor, (Master’s Thesis, Carleton University).

¹²⁰ Lauriault (2003).

¹²¹ Nancy Tosta and Michael Domaratz. “The U.S. National Spatial Data Infrastructure,” in *Geographic Information Research: Bridging the Atlantic*, ed. Massimo C. Craglia and Helen Couclelis (London: CRC Press, 1997), 22.

¹²² International Standards Organization (2003). *Geographic information – Metadata* (accessed March 2011), (<http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=26020&ICS1=35&ICS2=240&ICS3=70>).

¹²³ See Table 3 Metadata on the InterPARES 2 website at http://www.interpares.org/display_file.cfm?doc=Archivaria64_Todays_Data_supplementary_tables.pdf.

¹²⁴ National Research Council (1999), Spatial Information Resources, *Distributed Geolibraries*.

¹²⁵ National Research Council (1999).

this type of portal are those, which use Web mapping services, such as the British Antarctic Survey. The GeoConnections Discovery Portal (IP2SF15) is an example of this type of portal. Data portals can be housed in a single physical location (e.g., Statistics Canada), and they may be virtual, housed in a set of physical locations and linked electronically to create a single, coherent collection (e.g., GCMD, International Comprehensive Ocean Atmospheric Dataset). The distinction between centralized, distributed or unified portals may have funding, policy and preservation implications.

There are three functional data collections/portal categories¹²⁶ such as 1) research data collections; 2) resource or community data collections; and 3) reference data collections, and each have sets of characteristics that affect the type of data they contain, data contributors, funding and institutional arrangements and the kind of preservation strategies that may apply.

The GS10 portals provided a rich array of information on the topic of authenticity. Ensuring that data are of good quality and can be trusted is critical to these data portals, or else users would not rely on them. For many of the portals, the process of data control begins at the time the data are ingested into the system, made accessible via Web server sharing protocols, or described in a metadata description form. Some portals only ingest data that are derived from peer review journals; others restrict who can access the portal and contribute data.

Once data are in a particular portal, there are a wide variety of security measures in place to ensure they are not tampered with. Before data are made available, most have validation processes in place to attest to their authenticity and quality. However, there did not seem to be any mechanisms for users to assess if the datasets they downloaded or received are authentic beyond metadata and file headers. Also, the concept of the presumption of authenticity defined as “an inference as to the fact of a record’s authenticity that is drawn from known facts about the manner in which that record has been created and maintained”¹²⁷ is a useful one as the context, practices, associated documentation, validation processes and authentication, and access measures would suggest that the Case Studies discussed and portals discussed here are presumed authentic from an archival perspective.

Table 6. General Study 10 Geospatial Data Portals

Portal Title	Description
Canadian Geospatial Data Infrastructure (CGDI) Access Portal	Geospatial data such as DEM, orthophotos, air photos, satellite and radar imagery, data Services, discovery metadata, documents, Internet Maps, and Atlases.
Statistics Canada	Under the Statistics Act Statistics Canada is required to collect, compile, analyse, abstract and publish statistical information relating to the commercial, industrial, financial, social, economic and general activities and conditions of the people of Canada.
Long-term Ecological Research (LTER)	Data of Holocene barrier island geology; salt marsh ecology, geology, and hydrology; ecology/evolution of insular vertebrates; primary/secondary succession; life-form modeling of succession; online maps, photos, webcams; physical; biological, images and geographic data, software, LTER Network Data; and models.

¹²⁶ For detailed descriptions of these see the American Institute of Physics (AIP), 2001, *AIP Study of Multi-Institutional Collaborations: Final Report. Highlights and Project Documentations* Melville, accessed March 2011, <http://www.aip.org/history/publications.html> and National Science Foundation, 2005, *Report of the National Science Board*, p. 20. The GS10 Report describes this more thoroughly and includes portal examples.

¹²⁷ InterPARES 2, *Terminology and Glossary*, http://interpares.org/ip2/ip2_terminology_db.cfm.

Southern California Earthquake Center (SCEC)	Data sets include seismic waveforms, mostly passive source seismic data collected by broadband, strong-motion and analogue instruments. LARSE I and II seismic survey, data sets from the portable deployments following the Landers and Northridge earthquakes, and data from the Anza network and SCEC borehole stations. Non-seismic data includes “survey-mode” precise GPS measurements made in southern California by various universities. The GPS data archive consists of raw GPS data, RINEX files, indices to the RINEX files, log sheets and site descriptions.
International Comprehensive Ocean Atmosphere Data Set (ICOADS)	Surface marine reports from ships, buoys, and other platform types. Each report contains individual observations of meteorological and oceanographic variables, such as sea surface and air temperatures, wind, pressure, humidity, and cloudiness; monthly summary statistics.
National Geophysical Data Center (NGDC - NOAA)	Geophysical data describing the solid earth, marine, and solar-terrestrial environment, as well as earth observations from space.
Antarctic Digital Database (ADD)	1:200,000/1:250,000 maps and collaborative topographic database compiled from a variety of Antarctic map and satellite image sources.
National Snow and Ice Data Center (NSIDC), NASA	Snow, ice and glaciological Data From satellite images remote sensing instruments, ground measurements, and data models.
U.S. Antarctic Resource Center (USARC)	Maps, orthophotos, satellite imagery, point data, Digital Elevation Models, Digital Raster Graphics, and aerial photography.
British Antarctic Survey (BAS) -Antarctic Environmental Data Centre	Antarctic research results, climate modeling and predictions. Also, oceanic, environmental, atmospheric, geoscience, water, and earth observation data from a variety of sensors and in many forms.
Global Change Master Directory (GCMD) –Global Change Data Center	Data inform climate change research, and datasets include models, instruments and services. It is a workspace as well as a place to store these working operational models. Disciplines include atmospheric science, oceanography, ecology, geology, hydrology, and human dimensions of climate change.
Community Data Portal at NCAR	CDP catalogs observational and computer simulation datasets, sources of data are related to sciences in the areas of: oceanic, atmospheric, space weather, and turbulence.
Earth Systems Grid (ESG) portal	Climate change models include High-resolution, long-duration simulations, data simulations, derived from UCAR/NCAR projects (particularly the Community Climate System Model (CCSM) and Parallel Climate Model (PCM) and the Intergovernmental Panel on Climate Change (IPCC). Also physical” datasets that are generated directly by climate simulations.
USGS Data Portals - GEO-DATA Explorer (GEODE)	Geologic discipline’s programs including coastal and marine geology, earth surface dynamics, earthquake hazards, integrated natural resource sciences, mineral resources, National Cooperative Geologic Mapping, volcano hazards. Also scientific and energy related data.

10. Conclusion

The CLI and the 1986 Domesday rescue and salvage examples demonstrated that preservation as an afterthought is uncertain, will most likely yield only partial results, is expensive and time consuming and is often not 100 percent successful. Also, that maps and atlases designed in closed proprietary systems impede preservation and furthermore, that saving the data is not enough, as entire systems either need to be preserved or recreated to regain intended functionality and aesthetics. For every successful recovery there are many datasets, which have been permanently lost, and many more which are at risk. The two examples given were only saved because of their high profile and demonstrable scientific and cultural importance. The three cybercartographic atlases discussed became community specific archives created outside of the traditional archival system but exemplify proactive archiving in their design. In addition, these are making records by recording oral knowledge and they are contesting the validity of official

treaty archival records, the transmediation of ‘official’ archival records into a digitized form and are creating conditions for secondary provenance. They are re-mapping colonial records with indigenous geonarratives and by doing so are re-interpreting old maps and legal processes according to an indigenous worldview. By examining existing Government of Canada geospatial data archiving initiatives and data management models, it was discovered that complex sensor derived geospatial databases are being proactively preserved although none are being ‘officially’ archived in perpetuity by a recognized archival institution and most of the initiatives examined are not governed by a institutional archival or records management policy. Regardless, these are concrete cases upon which to build. With regards to consultations, studies, reports and initiatives, few address geospatial data explicitly although research data findings and recommendations are mostly transferable to a geospatial context and in some cases recommendations from the public on the topic of archiving were ignored. For the last decade substantial studies have one after the other identified the issues and what needs to be done but none of the key recommendations have been implemented. Also, Canadian experts and citizens make valuable contribution when consulted and government continues to commission studies, which in the main seem to gather dust on shelves or in government servers. The reports and studies are accessible but digital data continue to be at increasing risk and many datasets are lost with limited or no, hope of recovery. People, however, remain committed to see data being preserved and show their commitment by participating in Research Strategy Canada, CODATA and the OGC. In terms of legislation, regulation, directives and policies, the creation of government geospatial data are well covered, but these are rarely implemented and there seems to be no implementation mechanism. There are few existing geospatial data preservation initiatives, and the sorry state of LAC and the lack of trusted digital repositories and records management of data overall is of great concern. It was also seen that laws, rules, directions and policy are irrelevant in Canada unless there are dedicated resources set aside to implement them in the government bureaucracy in real terms. In addition, LAC guidelines for the preservation of geospatial data and maps cannot be effectively used by geospatial data managers, and with recent budget cuts, it is uncertain what LAC’s capacity will be to assist data producers with preservation. It would seem, that as an interim strategy, it is best that data producers and managers incorporate life cycle records management into their processes and build in proactive archiving practices. Regarding research in Canada, the IP2 project was groundbreaking with its Case Studies and General Study 10 and these have provided much insight into geospatial record making practices. These studies provided empirical evidence that geospatial data dissemination and map-making processes generally have many good practices in place, most notably metadata, open source technologies and specifications which make it easier for their products to be archived. Furthermore, data in portals have already been appraised by institutions, and archivists can build archival practices into these. It is hoped that IP3 will implement the findings of IP2 for geospatial data and maps. Also, as seen, GeoConnections commissioned a second study by HAL to examine the preservation of its geospatial data assets, and it was argued that the CGDI is a good model within which preservation can be embedded.

Overall, we are losing digital data, maps and atlases faster than we are producing them. There are weak guidelines, technical and institutional infrastructure is sorely lacking, and limited financial resources have been allocated to operationalize laws and regulation. We have also seen glimmers of hope in some initiatives, and there is tremendous will as witnessed in consultations with experts. Archivists have portals upon which they can build, and some of the examined geospatial data preservation initiatives, portals and all the cybercartographic atlases are proactively building in preservation and records management practices.

Geospatial data, maps and atlases as shown in this paper, are a fundamental source in our memory of the world. They form part of our collective memory system, they help us understand our geonarratives, they counter colonial mappings, are the result of scientific endeavours, represent multiple worldviews, and they inform decisions. We simply need to overcome the many challenges preventing their preservation and build upon existing preservation examples and implement our laws, regulations, directives and policies.

Gold Sponsors



McGill



InterPARES Project

International Research on Permanent Authentic Records in Electronic Systems

Silver Sponsors



Ministry of Education, Culture and
Science

OCUL Ontario Council of
University Libraries



**UNIVERSITY OF ALBERTA
LIBRARIES**



uOttawa
Library

Supporters



CANADIAN COMMISSION FOR UNESCO
COMMISSION CANADIENNE POUR L'UNESCO
www.unesco.ca



Québec



École des sciences de l'information
School of Information Studies