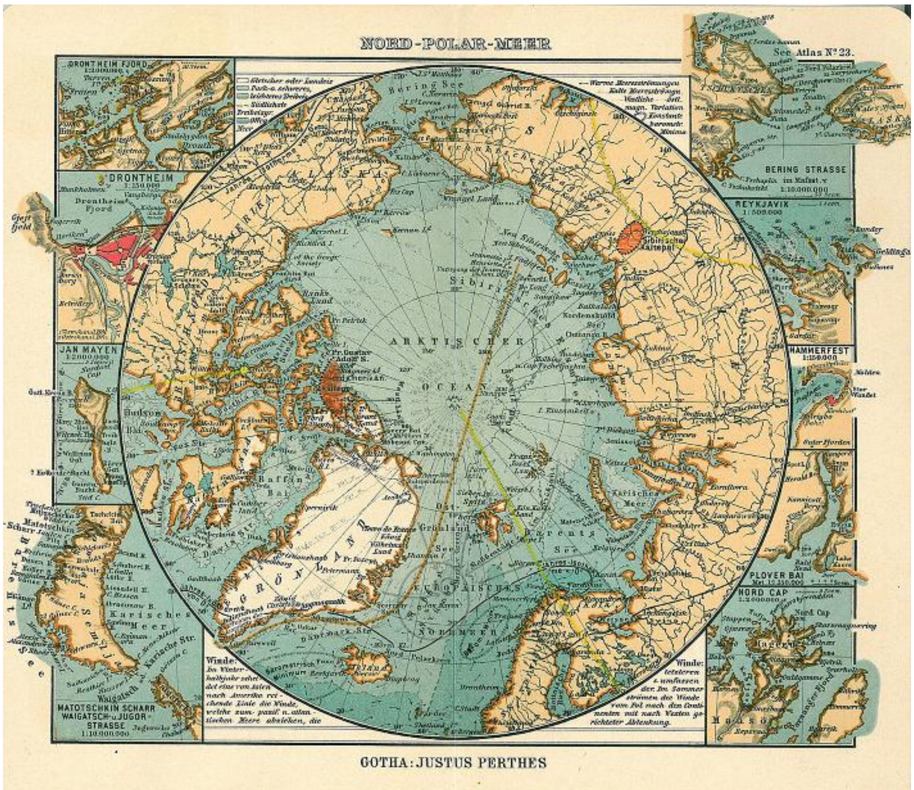
OPEN GEOSPATIAL CONSORTIUM (OGC)

PHASE 1 REPORT: SPATIAL DATA SHARING FOR THE ARCTIC

OGC Arctic Spatial Data Pilot



release: May 10, 2016

OGC Arctic Spatial Data Pilot

Phase 1 Report: Spatial Data Sharing for the Arctic

The Open Geospatial Consortium (OGCő)

35 Main Street, Suite 5 Wayland, MA 01778 Telephone: +1-508-655-5858

Acknowledgements

The OGC expresses its gratefulness to its sponsors US Geological Survey and Natural Resources Canada for supporting this work. OGC further wishes to express its gratitude to the following companies and organizations, who provided excellent contributions that have been used in excerpts in this report.[1]

Organization / Company
US Geological Survey (sponsor)
Natural Resources Canada (sponsors)
ASL Environmental Science Inc
Compusult Limited
Content Interface Corporation
Ecere Corporation
ESRI & ESRI Canada
exactEarth Ltd
Federal Geospatial Platform Project, Federal Committee on Geomatics and Earth Observations (FCGEO)
Fugro Pelagos Inc.
Geographical Names Board of Canada
Government NWT
GRID Corp
Harris
Inuvialuit Regional Corporation
Kivalliq Inuit Association
National Snow and Ice Data Center
NGA Maritime Safety Office
NOAA Coast Survey
Polar Community (for details see Appendix D)
Polar Knowledge Canada
Public Safety Geoscience Program (Earth Sciences Sector, NRCan)
Pyxis
University Carleton, GCRC
University Grenoble-Alpes
University of North Carolina
W3C Maps for HTML Community Group
Yukon Government

TABLE 1: Organizations and companies contributing to this report

[1] To avoid an overload with references, in particular as paragraphs often include parts provided by different companies or organizations, this report does not include local references other than for images.

Abstract

Phase 1 Report: Spatial Data Sharing for the Arctic

OGC Arctic Spatial Data Pilot

by OGC

This report presents the results of a concept development study on *SDI for the Arctic*, sponsored by US Geological Survey and Natural Resources Canada, executed by the Open Geospatial Consortium. The focus of this study was to understand how to best support the development of an SDI for the Arctic, to understand the view and specific requirements of indigenous peoples in the North, and how to make existing implementations i) better known to stakeholders, and ii) better serving stakeholders' needs. The study included an open *Request for Information* (RFI) with the objective to gather external positions and opinions on the optimal setup and design of an SDI for the Arctic. Responses to this RFI have been integrated into this report.

The report discusses the various types of stakeholders of an SDI for the Arctic with their specific needs and requirements on aspects such as data sharing, standards & interoperability, funding and investment, integration with existing systems, architecture and platform as well as security, privacy, and safety. The report further discusses various architecture models with focus on standards required to optimize discovery, usage, and processing of data in an highly heterogeneous network of SDI data and service providers. The report concludes with a number of demonstration scenarios that could be used in the pilot's second phase to demonstrate the value of an SDI for the Arctic to a broad range of stakeholders.

Contents

List of Figures

List of Tables

List of Abbreviations

Arctic-SDI	Arctic Spatial Data Infrastructure
ArcticSDP	Arctic Spatial Data Pilot
AIS	Automatic Identification System (AIS) messages
CAAS	Communication as a Service
CGDI	Canadian Geospatial Data Infrastructure
CGNDB	Canadian Geographical Names Data Base
CSV	Comma Separated Values
CSW	Catalog Service Web
DaaS	Data as a Service
DAP	Data Access Protocol
DCAT	Data Catalog Vocabulary
ENC	Electronic Navigational Charts
EO	Earth Observation
EOWCS	Earth Observation Profile Web Coverage Service
FGDC	Federal Geographic Data Committee
GEO	Group on Earth Observation
GEOSS	Global Earth Observation System of Systems
GeoXACML	Geospatial XACML
GIS	Geographic Information System
GML	Geography Markup Language
HDF	Hierarchical Data Format
HTTP	Hypertext Transfer Protocol
IHO	International Hydrographic Organization
InaaS	Information as a Service
ISO	International Organization for Standardization
ICT	Information and Communication Technology
IT	Information Technology
JSON	JavaScript Object Notation
JSON-LD	JSON Linked Data
KML	Keyhole Markup Language
MOU	Memorandum of Understanding
MSDI	Marine Spatial Data Infrastructure
NASA	National Aeronautics and Space Administration
netCDF	network Common Data Form
NOAA	U.S. National Oceanic and Atmospheric Administration
NRCan	Natural Resources Canada
NSDI	National Spatial Data Infrastructure
NWT	Northwest Territories
NWTCG	The Northwest Territories Centre for Geomatics

OGC	Open Geospatial Consortium
OPeNDAP	Open-source Project for a Network Data Access Protocol
PaaS	Platform as a Service
POI	Points-of-interest
RDF	Resource Description Framework
RFI	Request For Information
RFQ	Request For Quotation
SaaS	Software as a Service
SDI	Spatial Data Infrastructure
SOS	Sensor Observation Service
SPARQL	SPARQL Protocol and RDF Query Language
SWE	Sensor Web Enablement
SWG	Standards Working Group
UN-GGIM	United Nations Committee of Experts on Global Geospatial Information M
U.S.	United States
USGS	U.S. Geological Survey
W3C	World Wide Web Consortium
WCPS	Web Coverage Processing Service
WCS	Web Catalog Service
WFS	Web Feature Service
WMS	Web Mapping Service
WMTS	Web Mapping Tile Service
WPS	Web Processing Service
WS	Web Service
WSDL	Web Services Description Language
XACML	eXtensible Access Control Markup Language

Chapter 1

Introduction

The polar regions are of increasing interest to the whole world as a result of their linkage to global climate systems, opportunities for economic development, geo-political strategic importance, and their environmental importance as homes to Indigenous populations and other residents and sensitive ecosystems. Polar data are required by the scientific community and residents to support research on topics such as climate, atmosphere, land, oceans, ecosystems, ice and snow, permafrost, and social systems; and by the operations community to support impact assessments, engineering design, safe navigation and operations, risk management, emergency response, weather forecasting, and climate change adaptation. These activities contribute to environmental protection, heritage preservation, economic development, safety of life and property, and national sovereignty.

The polar data community is well organized and is pursuing activities to improve data management for all of the diverse members of the polar community. Polar data infrastructure is evolving from a system where data are discovered in data catalogues and downloaded to the local machines of users, to a system of distributed data made interoperable using standards and providing users with storage and computational capacity close to large repositories of data.

Interoperability and open standards are core to any spatial data infrastructure for the Arctic, as they enable the exchange of data and the use of processing, visualization, and representation services in distributed systems most efficiently. The economic benefits of building systems based on standards has been shown in many studies (e.g. DIN 2011). This pilot will demonstrate the value of standards in an environment that is principally built as a system of systems, i.e. an Arctic Spatial Data Infrastructure that integrates a number of existing systems as well as individual services and data repositories.

Indigenous Peoples of the Arctic and their representative organizations are increasingly active in using information and communications technologies (ICT) to access data and share their information and knowledge. This includes the establishment of protocols for ethical and culturally appropriate development of ICT.

This report presents the results of a concept development study on *SDI for the Arctic*, sponsored by US Geological Survey and Natural Resources Canada, executed by the Open Geospatial Consortium. The focus of this study was to understand how to best support the development of an SDI for the Arctic and how to make existing implementations i) better known to stakeholders, and ii) better serving stakeholders' needs. The study included an open *Request for Information* (RFI) with the objective to gather external positions and opinions on the optimal setup and design of an SDI for the Arctic. Responses to this RFI have been integrated into this report.

The report discusses the various types of stakeholders of an SDI for the Arctic (chapter 2) with their specific needs and requirements on aspects such as data sharing, standards & interoperability, funding and investment, integration with existing systems, architecture and platform as well as security, privacy, and safety (chapter 3). The report further discusses governance goals (chapter 4) before concentrating on more technical aspects such as architecture models (chapter 5), data (chapter 6), and standards & interoperability aspects (chapter 7) to optimize discovery, usage, and processing of data in an highly heterogeneous network of SDI data and service providers. The report concludes with a number of demonstration applications scenarios that could be used in the pilot's second phase to demonstrate the value of an SDI for the Arctic to a broad range of stakeholders (chapters 8 and 9).

1.1 Pilot Participants and Goals

The Arctic Spatial Data Pilot is sponsored by Natural Resources Canada and the United States Geological Survey (USGS) with a North American focus, yet is scalable to the circumpolar community. It supports the evolvement of the Canadian Geospatial Data Infrastructure and National Spatial Data Infrastructure in the Arctic (i.e. Canadian Territories, Northern Quebec, Northern Labrador and Alaska). To be successful, the Arctic Spatial Data Pilot has to take particular requirements into account, including responding to priorities of Northerners and Indigenous Communities, working in zero/low bandwidth regions and considering the realities of frontier economies.

The Arctic SDI will play a key role in a range of complex issues where geospatial data are necessary, such as responsible resource development, environmental management and assessment, regulatory reviews, and safety and security. The Arctic SDI aims to make geospatial information available in a standardized way to the public, academic institutions, the private sector and others who are involved in to conducting research or produce value-added products and applications, driving innovation and stimulating economic development. Geospatial data, services and applications accessed through Arctic SDI will help agencies understand the impact of climate change and human activities in the Arctic, facilitating monitoring, management, emergency preparedness and decision making. Moreover, Arctic SDI will support Canadian research on climate change and enable decision-makers to take effective decisions and implement mitigation strategies.

Organization managing the RFI

The Open Geospatial Consortium (OGC) is an international consortium of more than 500 companies, government agencies, research organizations, and universities participating in a consensus process to develop publicly available geospatial standards. OGC standards support interoperable solutions that "geo-enable" the Web, wireless and location-based services, and mainstream IT. OGC standards empower technology developers to make geospatial information and services accessible and useful with any application that needs to be geospatially enabled.

Sponsors

Natural Resources Canada (NRCan) seeks to enhance the responsible development and use of Canadas natural resources and the competitiveness of Canadas natural resources products. We are an established leader in science and technology in the fields of energy, forests, and minerals and metals and use our expertise in earth sciences to build and maintain an up-to-date knowledge base of our landmass. NRCan develops policies and programs that enhance the contribution of the natural resources sector to the economy and improve the quality of life for all Canadians.

As the largest water, earth, and biological science and civilian mapping agency of the United States, the U.S. Geological Survey (USGS) collects, monitors, and analyzes data and information, and provides scientific understanding about natural resource conditions, issues, and problems. The diversity of its scientific expertise enables USGS to carry out large-scale, multi-disciplinary investigations and provide impartial scientific information to resource managers, planners, and other customers.

1.2 ArcticSDP: CGDI, NSDI and Arctic SDI Linkages

There are a number of SDI initiatives addressing the Arctic region, such as the US National Spatial Data Infrastructure (NSDI) or the Canadian Geospatial Data Infrastructure (CGDI). Other initiatives, such as the Arctic Council endorsed Arctic SDI address the entire Arctic region based on a cooperation between the mapping agencies of the eight bordering states. From a Canadian perspective, Arctic SDI addresses all SDI elements of the Canadian Geospatial Data Infrastructure primarily north of 60 degrees latitude (i.e. Canadian Territories, Northern Quebec and Northern Labrador), which assist regional, national and international SDIs. The US definition addresses the compatibility of USGS and the FGDC role in supporting national and international SDIs. Both the NSDI and CGDI (or the USGS and NRCan as organizations respectively) are contributing to the development and sustainability of the Arctic Council endorsed Arctic SDI.

For this RFI, the term *Arctic SDI* is also used generically for an SDI that serves Arctic data and services. Arctic SDI is further used for an SDI of several flavors. It is part of national SDIs that address the Arctic region and it is a cooperation between eight national mapping agencies with their internet platform featuring an Arctic data portal and Web services. Thus, it combines the various perspectives based on the rationale that there is a reciprocal influencing and dependency process between national (and even sub-national) and international initiatives addressing the Arctic region.

In a reciprocal process, evolving an SDI for the Arctic helps to generate a better understanding of how the national spatial data infrastructures can be developed and applied to support Arctic priorities. By implementing standardized ways to share and process geographic data among all users, costs for collecting and using data can be significantly reduced while decision-making is enhanced.

1.2.1 National Spatial Data Infrastructure (NSDI)

> "The NSDI has come to be seen as the technology, policies, criteria, standards and people necessary to promote geospatial data sharing throughout all levels of government, the private and non-profit sectors, and academia. It provides a base or structure of practices and relationships among data producers and users that facilitates data sharing and use. It is a set of actions and new ways of accessing, sharing and using geographic data that enables far more comprehensive analysis of data to help decision-makers choose the best course(s) of action." FGDC

The vision of the NSDI is to leverage investments in people, technology, data, and procedures to create and provide the geospatial knowledge required to understand, protect, and promote national and global interests.

According to the National Spatial Data Infrastructure Strategic Plan 2014-2016, the Arctic Spatial Data Pilot addresses a number of desires the NSDI shall fulfill in the near future:

- Provide government, businesses, and citizens with a way to visualize and explore data to derive information and knowledge.
- Create a network of resources and services for the seamless integration of location-based information into broader information assets to serve the needs of government, the business community, and citizens.
- Serve as an enabling resource for discovery, access, integration, and application of location information for a growing body of users.
- Leverage shared and open standards-based services and focus on applied information for improved decision-making.
- Promote place-based business intelligence and smart, shared applications.
- Include a core set of information layers that interface with other non-spatial data being generated.
- Use real-time data feeds and sensor webs for improved monitoring, control, situational awareness, and decision-making.
- Facilitate access to and use of multi-temporal information linked to place.

- Integrate and use advanced technologies and their associated standards and best practices.
- Facilitate use of community-driven open standards with multiple implementations.

1.2.2 Canadian Geospatial Data Infrastructure (CGDI)

The Government of Canada creates an environment for the development of the Canadian Geospatial Data Infrastructure (CGDI). Natural Resources Canada's responsibility is to foster an environment for the development and use of spatial data infrastructures (SDI) both within Canada and internationally. Spatial data infrastructures are highly distributed components that allow data integration based on international standards, industry consortia, operational policies, governance mechanisms and application development.

The CGDI is an on-line network of resources that improves the sharing, use and integration of information tied to geographic locations in Canada. It helps decision-makers from all levels of government, the private sector, nongovernment organizations and academia make better decisions on social, economic and environmental priorities. The infrastructure itself consists of data, standards, policies, technologies and partnerships that are in place to allow the sharing and visualization of information on the Internet. Primarily north of 60 degrees latitude, the CGDI is referred as the Canadian Arctic Spatial Data Infrastructure.

Governance of geospatial information management in Canada is based on a cooperative approach between the federal, provincial and territorial governments, industry, academia and the public. Numerous committees and organisations support the evolution of the CGDI including the Federal Committee on Geomatics and Earth Observation, Canadian Council on Geomatics and the GeoAlliance.

The Arctic Spatial Data Pilot will contribute to realize key actions listed in the 2015 Assessment of the Canadian Geospatial Data Infrastructure, the Canadian Geospatial Data Infrastructure Vision, Mission and Roadmap - The Way Forward, the GeoAlliance Canada Strategy and Action Plan, and the 2008 *Aboriginal Community Land and Resource Management: Geospatial Data Needs Assessment and Data Identification and Analysis*:

- Communicate and promote the benefits of open data sharing.
- Improve mechanisms for active collaboration.
- Develop mechanisms that acknowledge the contribution of data.
- Build collaboration tools to create, maintain and improve the quality of geospatial data
- Establish tools related to volunteered geographic information (VGI), cloud computing and data sharing.
- Establish mechanisms to make standards and policy processes easily and widely communicated.
- Continue to promote open data, with an emphasis on efficiency of data accessibility.
- Foster standards and operational policies that will facilitate and increase the use, sharing and non-duplication of geospatial data.

1.2.3 Arctic SDI

> The Arctic Council endorsed Arctic SDI is a cooperation between eight Arctic National Mapping Agencies whose Board includes the Mapping Directors of those Agencies [including NRCan and USGS]. The purpose of the Arctic SDI is to support the Arctic Council and other relevant stakeholders in meeting their goals and objectives by using reliable and interoperable geospatial reference data of the Arctic, accessible via the Arctic SDI Geoportal. The decision-making body of the Arctic SDI cooperation is the Arctic SDI Board. The Board consists of one Director General or deputy Director General from each of the MOU signatories which countries are members of the Arctic Council. (Arctic Council, 2015).

The Arctic SDI geoportal is available online. One goal of this pilot is to evaluate the geoportal's ease of use and to potentially complement it with additional clients, tools, and applications that allow efficient use of Arctic SDI data and processing of resources. Additionally, the pilot supports the Technical and Interoperability Objective of the Arctic SDI Strategic Plan 2015-2020.

1.2.4 Relationship of Arctic SDI and the Arctic Spatial Data Pilot Activity

The goal of the Arctic Spatial Data Pilot is to make the Arctic SDI better understood and more attractive to its various stakeholders. As part of the pilot, communication with existing stakeholders will be intensified and new stakeholders will be approached. The pilot supports future Arctic SDI development by

- gathering requirements on the Arctic SDI,
- explaining the Arctic SDI concept, technology and its application to new stakeholders (video(s))
- developing additional components of the Arctic SDI and making more data available

1.3 Arctic Spatial Data Pilot Activity

The Arctic Spatial Data Pilot Activity is sponsored by US Geological Survey and Natural Resources Canada. Kicked-off on December 3, 2015, the goal is to demonstrate the diversity, richness and value of Spatial Data Infrastructure (SDI) Web services to Arctic SDI stakeholders. The diversity of Arctic data available via OGC specifications shall be assessed within the context of domestic, continental and international requirements.

The project is being executed in two phases. The first phase is organized as an OGC concept development study. The second phase will be an OGC pilot initiative with active involvement of a number of OGC member organizations. Funding will be made available for this purpose.

Both the OGC Concept Development Study and Pilot are conducted in accordance with the OGC Interoperability Program Policy and Procedures. Phase one develops an inventory of available geospatial Web services across the Arctic, defines the core components of the Arctic SDI architecture, and defines use-cases and scenarios for future implementations as part of phase two. These activities were complemented by the request for information (RFI) in order to capture the various perspectives, requirements, and opinions by Arctic SDI stakeholders and contributors. Further beyond, the Arctic Spatial Data Pilot phase one serves as direct input for the large-scale interoperability program initiative Testbed-12, where additional support may be made available to test and further develop components identified by the Arctic Spatial Data Pilot.

The goal of Phase two is to articulate the value of interoperability and to demonstrate the usefulness of standards. This will be done by implementing the recommended Arctic SDI architecture and developing a video that will tell the story of the scenario(s) and showcase incorporation of the services into Arctic SDI Geoportal and other applications. Funding will be made available as part of phase 2.

1.3.1 Relationship between Arctic Spatial Data Pilot and OGC Testbed-12

The Arctic Spatial Data Pilot and concept study runs in parallel to the OGC Testbed-12 activity as illustrated in figure 4. Testbed-12 is an OGC activity where sponsors have defined requirements on selected topics. The OGC has released an RFQ to solicit participants interest to implement and explore the various components and work items. Further details about Testbed-12 can be found online.

Both the Testbed-12 as well as the Arctic Spatial Data Pilot have components identified that play a role for Arctic SDI. The Arctic Spatial Data Pilot defines the scenarios and use cases and further implements Arctic SDI components to elaborate and demonstrate the usefulness and value of the Arctic SDI. By definition, Testbeds explore components with a lower level of component maturity and experiment with new technologies. Testbeds identify work items that need further exploration, which then could be explored in OGC pilot activities such as the Arctic Spatial Data Pilot.

Chapter 2

Stakeholders

2.1 Types of Stakeholders

The number and types of stakeholders is changing with the changing Arctic climate and environment. Being more ice free during summers allows increased vessel traffic following new routes, increased surveying and research work, increased exploration work, or increased tourism among other things. With the generally increasing human activity, chances of disasters and emergencies also increase, making first responders a key user group of an SDI for the Arctic. At the same time, the Arctic is home of indigenous people who should be considered key stakeholders (on indigenous peoples see also Appendix F).

There are multiple, orthogonal ways to describe relevant stakeholders in the context of an SDI for the Arctic. Here, we have differentiated the range of stakeholders into five classes. The stakeholders summarized under each class often have some level of influence on each other, illustrated by the circular arrows connecting the classes as illustrated in figure 2.1. The classes are not mutually exclusive, and many organizations or individuals are members of more than one class. The wide class of **end-users** includes all consumers of products provided by the other classes such as e.g. data and services, products in the form of reports and statistics, policies and regulations etc. The second class aggregates all **data producers** or creators, **data providers**, **data brokers**, and **value added re-sellers**. This large group is of particular relevance for this pilot, as it is responsible for one of the main goods of the Arctic SDI, the data. The third group covers data processors such as GIS or mapping experts. These experts create products such as analyses, reports, statistics, or maps using data provided by the previous group. The fourth group is somewhat orthogonal to the previous three. The **hardware**, **storage- and computing service providers** provide the necessary infrastructure for data exchange and processing. The last group again is orthogonal to the ones described before. **Policy makers** lay out the necessary rules and guidelines for a successful operation and governance of an Arctic SDI.

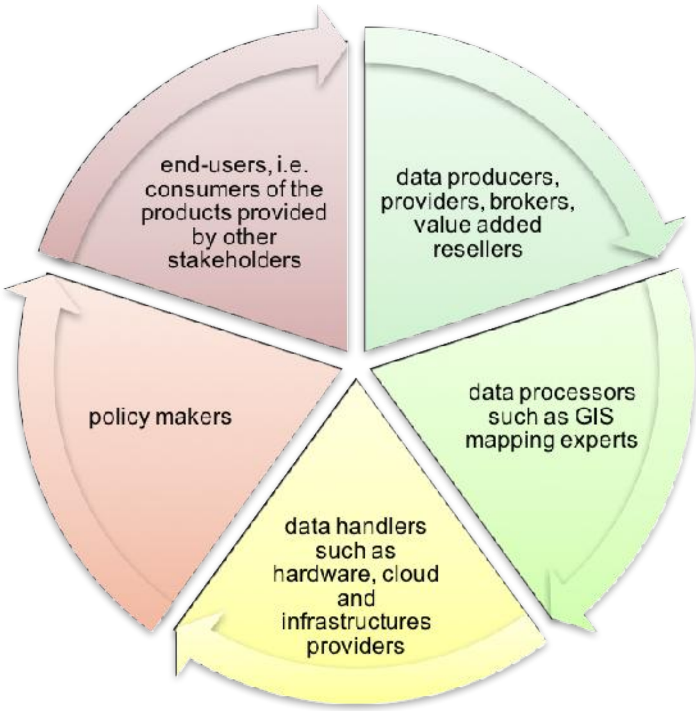


FIGURE 2.1: Types of Stakeholders

The stakeholders classified in one or many of these five classes come from a wide range of organizations. An already long, though still non-exclusive, list is provided in table 2.1.

Stakeholders
Federal, state, provincial, local, or territorial governments
Indigenous peoples governments and organizations
Arctic Council and related boards and groups
Co-management boards
Land and water boards
Land use planners and surveyors
Non-Governmental Organizations (NGOs)
Councils such as Arctic Council or Inuit Circumpolar Council
Standards Developing Organizations
Software developers
Native corporations
Oil and gas or mining companies
Utility companies/organizations
Shipping and cruise ship companies
Fishing companies
Port managers and harbormasters
Insurance companies
Search and rescue officials
Researchers from various fields such as climate, conservation
Archaeology, marine, hydrology, ecology, and geological science
Academic and educational institutions
K-12 programs
Mapping and GIS experts
Diplomatic and national security officials

TABLE 2.1: Overview of the Arctic Data Stakeholders

A number of organizations from this long list has been identified as potential collaboration partners for the Arctic Spatial Data Pilot or have been emphasized by responders to the ArcticSDP RFI as particularly relevant. Still, the ArcticSDP welcomes any organization or individual willing to support the goals and objectives of this pilot. These include:

- Indigenous agencies, Metis, Inuit, First Nations
- Organizations such as Government of Nunavut, Nunavut Geoscience, Nunavut Planning Commission, Nunavut Impact Review Board, Nunavut water board, and Nunavut Tunngavik Incorporated and Regional Inuit Associations, and other organizations to ensure the full engagement with Arctic Indigenous Peoples as emphasized by the recent Arctic Observing Summit (ISAC, 2016).

- Indigenous and Northern Affairs Canada, Department of Natural Resources Canada, Canadian Space Agency, Canadian Circumpolar Institute
- US Geological Survey (USGS), Federal Geographic Data Committee (FGDC), and the US National Aeronautical and Space Agency (NASA)
- Arctic Regional Hydrographic Commission working group on Arctic Marine Spatial Data Infrastructure
- International Arctic Science The Arctic Data Committee (ADC) of the International Arctic Science Committee (IASC), Sustaining Arctic Observing Networks (SAON), and the Arctic Portal.
- International Antarctic Science The Standing Committee on Antarctic Data Management (SCADM) of the Scientific Committee on Antarctic Research (SCAR); and the Southern Ocean Observing System (SOOS).
- International Cryosphere Science Climate and Cryosphere (CliC)
- International Polar Operations The International Ice Charting Working Group (IICWG) and Polar View Earth Observation.
- United States The National Snow and Ice Data Center (NSIDC), the Interagency Arctic Research Policy Committee (IARPC) Arctic Data Coordination Team, the Alaska Data Integration Working Group (ADIwg), the NSF-funded Antarctic and Arctic Data Consortium (a2dc), the Arctic Research Mapping Application, the Arctic Observing Viewer, and the Barrow Area Information Database.
- Canada Polar Knowledge Canada, the Canadian Cryospheric Information Network (CCIN), the Geomatics and Cartographic Research Centre at Carleton University, Polar Data Catalog, and the Canadian Consortium for Arctic Data Interoperability (CCADI).
- Europe EU-PolarNet and the European Space Agency (ESA).
- Asia Japans National Institute of Polar Research (NIPR)
- Canadian Consortium for Arctic Data Interoperability

A list of individuals that can help facilitate contact and engagement of a number of organizations is provided in Annex A.

2.2 Business needs of Stakeholders

Though the businesses of the various stakeholders vary considerably, there is substantial overlap in terms of business needs among most stakeholders, at least from the first three groups described in figure 2.1. Generally speaking, the business needs include aspects such as easily discover, access, download and analyze arctic spatial data on the data consumer side, and the ability to publish, integrate, aggregate and analyze geospatial data and related non-geospatial data on the data producer, provider and processor side. Focus shall be on ease-of-use and effectiveness, as data producers will not publish data voluntarily to multiple locations and try to minimize data integration efforts on the other side. Integrated systems, possibly in a system-of-systems or network-of-networks approach with the ability to harvest data from existing solutions in a secure, reliable manner should be supported.

In addition, there is a need for certain data sets with further requirements on real-time or archived availability, data and system IPR (intellectual property rights), reuse and indemnification rules and regulations, security and privacy settings, as well as costs. As Northern projects often take place in small subsets and consolidation is not always possible, metadata plays an essential role to understand which methodologies or standards have been applied during data acquisition, processing and preparation work.

On the system side, it is essential that systems are operational and reliable with clear life cycle costs to providers and users. Stakeholders require robust, but intuitive easy-to-use tools to access, visualize and contribute data and information in a manner that allows for ingestion into organizations to support policy development and decision making. The underlying systems have to cater for various types consumer capacities. While some of the stakeholders may have very limited internal geospatial capacity or solutions, others are far more advanced.

In terms of data sets, a detailed analysis of both provided and required data sets is discussed in chapter 6. Therefore, the following list only contains aspects that are relevant to the nature and representation of data, independently of the data type and domain; have umbrella importance for other data sets in terms of referencing or geolocation; or represent overarching groups of data:

- Consolidated authoritative data sets

- Geographical names capable of handling indigenous language characters, sound and/or video clips, spatial delineations, images, and naming decision documents for published names
- Geolocations for community infrastructure elements (due to lack of street names and addressing in small communities)
- Traditional knowledge that originates from experiences or oral traditions and that cannot be captured by sensors or processes
- Northern statistical information
- A baseline knowledge base of physical environmental data for a region is necessary for pre-FEED, FEED (front end engineering design), environmental assessment, and engineering design.

The following figures illustrate the interlocking of data and applications. Even though only a small subset of all data types, stakeholders, and applications is provided, the tables still provide a valuable insight in the richness of applications in the Arctic. The tables are based on input provided by Fugro Pelagos. They show how new data sets made available online often produce many new users that have not been in focus at the phase of publications, but found the data to be useful in contexts not intended before. The tables do not provide an exhaustive list of data and all its possible usages.

Data Type	Stakeholder	Application
Hydrographic mapping	Federal, state, local agencies	• Nautical charting for navigation safety • National defense • Emergency response • Infrastructure planning/development • Fisheries management • Legal boundary determinations • Flood planning • Baseline habitat mapping • Environmental baseline monitoring • Sovereignty
	Oil and gas companies	• Safe navigation • Engineering activities during exploration, development, and production • Environmental responsibilities related to sustainable development and protection of biodiversity
	Mining companies	• Safe navigation • Engineering activities during exploration, development, and production • Environmental responsibilities related to sustainable development and protection of biodiversity
	Utility companies	• Engineering activities during exploration, development, and production • Environmental responsibilities related to sustainable development and protection of biodiversity
	Shipping and cruise ship companies	• Safe navigation • Trip planning • Route planning
	Commercial fishing companies	• Safe navigation • Fishing sources geolocation
	Geospatial community	• Survey methodologies, technology manufacturing, software development
	General public	• Safe navigation, subsistence activities, recreational boating and recreational fishing
	Insurance companies	• Safe navigation • Route planning • Environmental baseline monitoring

FIGURE 2.2: Business needs examples, part 1

Data Type	Stakeholder	Application
Coastal mapping	Federal, state, local agencies	• Maritime safety • Emergency response (natural disasters, etc.) • Offshore development regulation • Scientific research • Coastal monitoring (change analysis) • Coastal flooding modeling, analysis, mitigation • Earthquake/tsunami assessment, mitigation • Regional sediment management • Infrastructure development/maintenance • Fisheries management • Environmental baseline monitoring
	Engineers	• Infrastructure development/maintenance • Flood planning • Environmental baseline monitoring • Coastal monitoring (change analysis)
	Insurance companies, real estate companies, lenders	• Flood risk information • Tsunami inundation • Erosion studies
	General public	• Flood risk information • Erosion studies
	Geospatial community	• Survey methodologies, technology manufacturing, software development
Ice thickness mapping (sea ice)	Federal, state, local agencies	• Navigation safety during in-ice operations • Ice load information for infrastructure design/engineering • Establishment of shipping lanes
	Oil and gas companies	• Navigation safety during in-ice operations • Environmental protection during in-ice operations • Ice load information for infrastructure design/engineering
	Shipping companies Cruise ship companies	• Navigation safety during in-ice operations
	Insurance companies	• Risk assessment for vessel operations in Arctic waters
	Researchers	• Tracking icefields over time as a measure of climate change • Correlation to ice gouge mapping and monitoring
	Geospatial community	• Survey methodologies, technology manufacturing, software development

FIGURE 2.3: Business needs examples, part 2

Data Type	Stakeholder	Application
Ice gouge data	Federal, state, and local government agencies	• Impact on navigation safety
	Oil and gas companies Mining companies Utility companies Port managers and harbor masters	• Infrastructure siting, planning, routing and protection • Impact on navigation safety
	Engineers	• Infrastructure siting, planning, routing and protection • Engineering design, construction, and maintenance • Impact on navigation safety
	Researchers	• Relationship of size and depth with age • Correlation to ice thickness mapping (sea ice)
	Geospatial community	• Survey methodologies • Technology manufacturing • Software development
Mid-resolution DEM	Federal, state, and local government agencies	• Aviation safety • Natural resource management • Infrastructure development • Change analysis • Public safety • Forestry • Emergency response
	Native corporations	• Land management
	Oil and gas companies Mining companies Utility companies	• Infrastructure siting and routing • Logistics planning
	Geospatial community	• Survey methodologies, technology manufacturing, software development

FIGURE 2.4: Business needs examples, part 3

Data Type	Stakeholder	Application
High-resolution DEM	Federal, state, and local government agencies	• Land-use planning • Forestry • Public safety • Emergency response
	Native corporations Oil and gas companies	• Infrastructure engineering, design, monitoring, and maintenance
	Mining companies Utility companies	• Infrastructure engineering, design, monitoring, and maintenance • Volumetric surveys
	Geospatial community	• Survey methodologies, technology manufacturing, software development
Multispectral imagery (onshore)	Federal, state, and local government agencies	• Land-use/land cover planning, analysis, and change detection • Forestry • Forest fire management • Wildlife management
Ocean current imaging	Federal government agencies	• Maritime safety • Search and rescue • Environmental protection • Coastal monitoring (change analysis) • Erosion studies
	Oil and gas companies	• Engineering design and large structure building projects • Search and rescue • Environmental protection • Spill management planning

FIGURE 2.5: Business needs examples, part 4

2.3 Analysis of Stakeholders

The engagement of stakeholders and the awareness raising of the Arctic SDI among potential stakeholders are key goals of the Arctic Spatial Data Pilot. First and foremost, the best way to get stakeholders in the Arctic region involved and well served is to meet their needs. This requires making Arctic data easy to find, use, and understand. This report covers guidelines and experiences from a large number of Arctic data providing and handling experts to identify the best way to achieve these essential requirements. In addition to this indirect stakeholder integration aspect that results from technology ease of use, reliability, and completeness, there are further dimensions that can be actively pursued. In this context, the recent conference statement from the 2016 Arctic Observing Summit shall be referenced, which emphasizes the importance of full engagement with Arctic Indigenous Peoples (ISAC, 2016). The following three subsections identify aspects that need to be addressed in order to improve the integration of stakeholders.

2.3.1 Coordination and Planning of activities

Coordination of SDI related activities and collaboration among the various organizations involved is a critical success factor for an SDI in the Arctic. A successful shared SDI would be a stepping stone to other collaboration activities that could focus on increased data collection, introduction of robust monitoring programs and ideally reduced duplication of effort. Fostering early coordination and planning and encouraging transparency within the public sector so that collection priorities and data requirements are clearly stated and the most efficient approach can be applied ensure that end user needs are met. In particular, the following aspects shall be considered:

- Include indigenous and First Nations consultation and consultation with indigenous and First Nations governments. Contact and request responses directly from GIS/IT staff who support land claim focused corporation (or data managers) in each of Canadas four Inuit regions collectively known as Inuit Nunangat, Inuit Tapiriit Kanatami and the Inuit Circumpolar Council. These groups should be considered primary stakeholders in the Canadian arctic.

- Work closely with indigenous representative organizations such as the Permanent Participants of the Arctic Council and national organizations such as Inuit Quajisarvingat/The Inuit Knowledge Centre at Inuit Tapiriit Kanatami in Canada. They can, in turn, facilitate connection to regional and community level organizations and individuals.
- Consult land and water boards, land use planning units, northern industry (through economic development offices)
- Involve the Government of the NWT Departments (the NWTCG could assist with this) and Alaska
- Consider the role of co-management boards established under land claims or other legislation (Wildlife Management Boards for example)
- Allow key entities to play a role in the management and ongoing governance to ensure ongoing local relevance and participation.
- Integrate multiple technologies during offshore data collection to speed the pace of acquisition, increase safety, and benefit multiple stakeholders with a variety of datasets meeting a varied level of needs.
- Connect with international organizations such as the International Hydrographic Organization (IHO), which supports the development of Marine Spatial Data Infrastructures (MSDIs) and through its MSDI Work Group aims to identify and promote national and regional best practices, assesses existing and new standards in the provision of marine components of spatial data infrastructures, promotes MSDI training and education, and facilitates (external) MSDI communication.

2.3.2 Outreach and Awareness

Outreach and awareness activities help attracting new stakeholders and reassure the importance of Arctic SDI among stakeholders already involved or at least aware of the relevance of an SDI for the Arctic. Combined with early coordination activities, outreach and awareness activities across stakeholders help to maximize efficiency and transparency, which are crucial components leading to acceptance and eventually success of an Arctic SDI. Suggested activities include:

- Outreach includes social media, story maps, press releases, conference presentations, websites, on-line and in-classroom training classes, books, etc. All this material is important for an SDI community to thrive.
- Promote the idea of crowd-sourced data, especially in offshore environments.
- Consider developing a White Paper for discussion and comment at both ministerial and senior management level across all stakeholders
- Publicize projects to help making the average citizen care about the Arctic.
- Improve collaboration between the public and private sectors to share lessons learned, establish best practices, and keep abreast of technology advancements.
- Participation in the trade shows, mining symposiums, and conferences conducted in areas north of 60 degrees

2.3.3 Technology Ease of Use and Data Availability

Technology ease of use coupled with reliability greatly impacts stakeholder adoption rates as well as ensuring users are successful. Thus, the best outreach is probably achieved by word of mouth, triggered by an excellent implementation of an SDI for the Arctic serving all stakeholders needs. Another approach to improve outreach is to implement it embedded in technology. In this case, outreach material is shipped with software or directly part of Web portals. Further on, outreach embedded in technology can provide a base set of data in tools out of the box without requiring substantial download of data at start-up time, an approach that simplifies the usage of software components.

Another aspect that needs to be careful revisited is the integration of an SDI for the Arctic with existing regional and national Spatial Data Infrastructures, such as e.g. Yukon Government SDI or US NSDI. Further attention shall be given to the integration of Arctic data and apps (applications that use the data) into widely deployed and used platforms. This is in addition to any stand-alone focused Arctic Portals. Simply put, some stakeholders are better served by integrating data and apps into the tools they use. For geospatial scientists it means being tightly integrated into their GIS; for policy stakeholders it would mean simple story maps, creating dashboard using Arctic statistical and geospatial data tied to policy questions; and for scientist it would mean integration of Arctic datasets with tools such as R. Additionally, stand-alone Arctic Portals must be designed for ease of use and must be interoperable with each other and be reliably available and secure. To achieve this level of integration, standards defining generic data containers or Web service interfaces for easy data access are of overall importance.

Any successful Arctic SDI needs to take into account the particular situation in the far North. The SDI needs to find ways to incorporate traditional knowledge together with additional socio-cultural aspects and potential language barriers.

Chapter 3

Requirements and Constraints

"One cannot learn Indigenous Knowledge by pointing and clicking on the Internet this must be stressed. It is only learned through relationships and learning with people who have learned it as it should be and who have lived it. Knowledge is not from a book or from a webpage; it is from experience." (P. Pulsifer, 2015)

The stakeholder business needs discussed in chapter 2 result in a number of requirements and constraints on an optimal SDI for the Arctic in terms of data sharing, standards and interoperability, funding and investment, integration with existing systems, agility and adaptability, and security, privacy and safety. Figure 3.1 illustrates these categories.

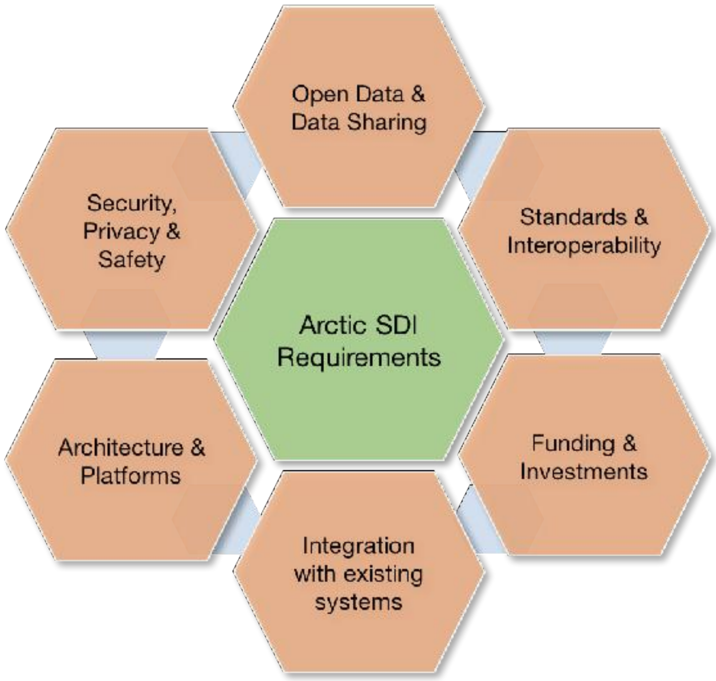


FIGURE 3.1: High level requirement categories

The following sections will briefly discuss more details on the various categories to ensure a robust baseline for the development of an SDI for the Arctic architecture and governance as discussed in chapters 4 and 5.

3.1 Open Data & Data Sharing

Open Data & Data sharing addresses both legal as well as technical aspects such as how to enable data sharing among disparate and heterogeneous endpoints and systems using common data models and schemas. Open data is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control. Open Data & Data Sharing further addresses organizational aspects such as how to encourage data sharing with social or economic incentives and enforcement of rules.

- Any SDI with Arctic data shall broker the delivery of government and non-governmental Arctic observations
- Integration of near real-time observations from both satellites and in-situ sensors is key. Traditionally, this has not been easily achieved due to the

proprietary nature of the sensor interfaces. New technologies such as SensorThings API shall be implemented. Also, auto-registry of sensors is a key requirement.

- The architecture shall support creation and exchange of research-oriented synthesized data sets (i.e. simulation model outputs)
- All data shall be accompanied by metadata. As this requirement is extremely tough to implement, new ways shall be explored to minimize the need for manually generated metadata.

3.2 Standards and Interoperability

Standards and Interoperability addresses mechanisms and agreements to ensure that components being part of or that are loosely connected to an SDI for the Arctic can communicate with each other.

- Interoperability of SDI components across platforms is of overall importance
- Data shall be served at standardized Web interfaces using standardized encodings.
- Standards-based Web GIS integrates and leverages all the investments that have already been made in GIS standards, data, and technologies. Any SDI for the Arctic should benefit from these investments and should be based on Web GIS patterns.
- Detailed compliance tests shall be available to ensure interoperability across components

3.3 Funding and Investments

The operation and maintenance of a successful Arctic SDI generates substantial costs that need to be covered by funding agencies or invested by companies with the goal to generate proportionate profit in the future. In terms of business needs, the following aspects need to be considered under *Funding and Investments*:

- Adequate funding from the various public sectors; at least initially
- Issues araise if benefits to be gained from are not understood in many areas of the Arctic and beyond
- Development of relevant applications in the private sector to generate desirable return on investment
- Recognition of geospatial data as an investment rather than a cost, which is possible through geospatial consortia making the data interoperable between different users to be utilized in an interoperable manner
- Any SDI for the Arctic shall consider not only one-time costs associated with implementing the solution but the ongoing requirements to support, maintain and enhance the solution over its lifecycle to ensure it continues to deliver value and meet stakeholder needs
- Individual management objectives, priorities, planning cycles and investment capacity are all constraints that will affect an organizations ability to participate in the development of an SDI
- Any SDI for the Arctic shall be prepared for eroding national or organizational technical infrastructures
- Cost efficiency is key and provided as much as possible out-of-the-box - meaning using existing cloud hosting and geospatial solutions and without the added expense of in-house software development

3.4 Integration with existing systems

Integration with existing systems is a critical aspect to ensure neat integration of data hosted in external systems and the protection of investments in other SDIs or platforms that shall be conserved. Therefore, SDIs for the Arctic shall

- Coordinate with National Mapping authorities that provide data for the Arctic
- Coordinate with international SDIs such as INSPIRE or GEOSS
- Integrate with national and regional SDIs such as CGSI, US NSDI, Alaska SDI, or Yukon SDI without replicating already available resources

- Integrate with and support widely deployed geographic information systems (GIS)
- Not be perceived as a competitor to local, regional, or national SDIs
- Need to integrate data platforms operated by national space agencies or other organizations providing satellite-derived data products

3.5 Architecture and Platforms

Architecture and Platform aspects play a key role in distributed spatial data collection, exploration, and processing environments; and need to ensure that the targeted SDI for the Arctic can keep pace with changing technologies and Internet trends. The following high level requirements have been identified:

- Development efforts for any SDI could be constrained by how prescriptive the architectural design is at the outset. To benefit from rapidly improving technology, an SDI for the Arctic needs to remain agile. Architectural decisions affect costs to the participants and the ability to benefit as technology changes. Early architectural decision can translate into constraints if they are too rigid in their approach. Therefore, questions such as these must be addressed: Will the SDI for the Arctic be a loose confederation of portals and platforms discoverable by open specifications and standards allowing as-is communities to participate? Or will the SDI for the Arctic be highly architected and all data and apps services be available as hosted/re-hosted services (cloud and/or on-premise)? Or any combination of the two approaches?
- Multi-linguism and technical language requirements should be taken into account.
- Technical knowledge and availability of skills is often a limiting factor in stakeholders adopting technical solutions, or in continuing efforts to maintain solutions already in place. The architecture has to cater for greatly varying paces at which organizations adapt new technology and will have to bridge a wide variety of technical solutions of differing ages and platforms.
- An SDI for the Arctic shall be very dynamic (in contrast to many other SDIs, that tend to be static), because change is occurring at a very high

rate. New data sets are added and the huge number of monitoring data sets are updated constantly.

- An SDI for the Arctic should be also designed for no- or low-bandwidth areas where the Internet is not readily available, as bandwidth and limited connectivity is the single biggest limiter to information in many Arctic areas. Arctic SDI designers must decide if they will provide infrastructure as well as data and apps. Examples of using data appliances that are loaded with data, software, and apps shall be explored.
- Intuitive site structure/navigation with best practices to lower the entry barrier to SDIs
- Efficient search functionality and fast download rates
- Reasonable access and download speeds for northern stakeholders, interoperable with existing geospatial systems, clear cataloging and tagging system to promote ease of discovery and reasonable metadata requirements.
- The architecture shall allow for future extensions and allow the integration of upcoming new patterns to handle e.g. Big Data or semantic annotation

3.6 Security, Privacy, Safety

Security, Privacy and Safety includes aspects such as vulnerability to attacks, acceptance and assurance of privacy concerns, secure and reliable access, protection of intellectual property rights, and assurance of system availability in critical situations, e.g. emergency responses or major crises. Additional items mentioned in RFI responses include:

- Many data sets are access-protected for good reasons (e.g. security implications or commercial or government interests). Though these reasons are fully acknowledged, SDI design should provide for obtaining information about how to access datasets that are not open but may be accessed through some other means. For example, industrial stakeholders who procure their own data collection programs often are protective of the data set but are willing to share them under certain circumstances. The necessary brokering has to be addressed.

- Foundational data shall be provided as license-free data (public sector)
- Individual logins, firewall protection and a secure server connections capable of transferring and storing highly sensitive data need to be available.

Chapter 4

Governance Objectives

The Arctic Spatial Data Pilot had in phase one an emphasis on the northern portions of Canada and the United States. This can be considered as a starting point and certainly acknowledges the significant work that is being done on polar data management in other countries. If possible, one can even strive for synergies with polar data management in the Antarctic.

The term "governance" is often used in the context of SDIs without a clear definition what is actually meant by it, which leads to misunderstandings regarding nature, scope, and challenges of "SDI governance" aspects (Paul Box and Rajabifard, 2009). To overcome issues caused by the polysemous nature of the term, this report follows a definition provided by Box:

"*Governance provides an enabling decision-making and accountability framework within which a community cooperates to achieve collective goals. SDIs which address the goal of sharing, accessing and using geospatial resources are rapidly developing around such communities, based upon interoperability standards and service-oriented architectural patterns. These communities vary greatly in thematic and geographic scope, level of mandate and resources, and technical capacity. With increased social and technical complexities and inter-relatedness of SDI initiatives, the design of effective governance becomes a significant challenge.*"(Paul Box, 2013)

In slightly broader terms, governance is about collective decision making, which is inline with the definition provided by the Commission on Global Governance (Commission on Global Governance, 1995): "*The sum of many ways in which individuals, institutions, public and private, manage their common affairs. It is the continuing process through which conflicting or diverse interests maybe accommodated and cooperative action taken*". Transferred to SDIs, the collective decision making process often gets blurred by the big number of stakeholders that participate in an SDI and may implement very little overlapping and thus collective decision making processes. Provided the goals of this pilot being the demonstration of the usability of an SDI for the Arctic for a large range or stakeholders rather than defining a fully featured SDI from scratch, one can avoid the differentiation of different classes of governance as suggested by (Paul Box and Rajabifard, 2009) or tiers of governance as suggested by (Atkinson and P Box, 2008) and concentrate on high level governance goals instead. These reflect to some extent established governance best practices. Nevertheless, for an optimal performance of any SDI, it is highly recommended to study the lessons learned as documented for example in (Paul Box and Rajabifard, 2009), which groups the most pertinent governance lessons learned into institutional, business, data and service categories and ranks them based on a number of case studies' results (for details see Annex B).

4.1 High Level Governance Goals

In the following a number of high level governance goals are briefly introduced that are of ample importance in any collaborative decision support system that focus on the Arctic.

- **Interoperability** Interoperability, the ability to easily share data across systems and users, is one of the most important priorities identified by the

FIGURE 4.1: High level governance goals

polar data community. An interoperable system must enable data access that can support many different users. This may require visualization or other mediation such as translating vocabularies to make data usable by different communities. Achieving interoperability will require adequate resources, a certain level of standardization, and a connected community.

- **Standards and Specifications** The overarching purpose of the polar data management community is to promote and facilitate international collaboration towards the goal of free, ethically open, sustained, and timely access to polar data through useful, usable, and interoperable systems. This includes facilitating the adoption, implementation and development (where necessary) of standards that will enable free, open and timely access to data.

- **Metadata** Building on the Polar Metadata Profile developed during the International Polar Year, the objective of this activity is to develop recommendations on a common set of metadata elements relevant across polar sciences, to facilitate interoperability and sharing between polar data repositories and online portals. To start, this effort will focus on identifying Arctic data centers or initiatives that have established a metadata template, schema, or profile. Initially, a limited set of disciplines or focus areas will be identified to make the scope manageable. Wherever possible and practical, the effort will build on and/or contribute to other related initiatives.

- **Data Publication** The objective of this activity is to provide a report and guide on data publication and citation for polar researchers. This would provide the polar community with a resource to help them to understand developments in this area, including assignment of DOIs (Digital Object Identifiers) to published data sets.

- **Including Arctic Indigenous Perspectives, Knowledge and Information** In this time of change, Indigenous knowledge and the underlying observations of Arctic peoples are more important than ever. Along with the knowledge of non-Indigenous local inhabitants, this knowledge is being increasingly documented and represented as digital data, but the nuances of these data are not well understood by the broader data management and science community. The perspectives of Indigenous people and other northern residents must be heard directly. This will enhance understanding of how Indigenous and local knowledge and observations can be used appropriately.

- **Community building** Improved polar data sharing that is part of a broader global system will require community building, collaboration, and coordination of efforts. To do this we need to better understand the nature of the polar data community (who is doing the work, where, what systems, etc.) across many scales and what we are collectively trying to achieve. Through the established bodies discussed, improved communication, outreach, and coordination within the polar community is needed while we recognize the importance of engaging with broader global initiatives including OGC and GEO.

- **Data Preservation and Rescue** Continuous re-use and re-purpose of past observations is key to increase their current understanding. Therefore, data, indigenous knowledge, and all the necessary descriptive information must be preserved. Too often, preservation is forgotten and data managers must pursue data rescue activities. Even current data are at risk of loss. Now, only seven years after the International Polar Year (IPY), there needs to be a data rescue campaign for much valuable IPY data because inadequate preservation support was being developed at that time and was limited in scope. Strategic data rescue programs must be developed, and preservation must be prioritized as a long-term investment and cost-saving measure.

- **Adequate Resources** Making progress will require adequate financial, technical, and human resources. More focus is needed on the training of early career scientists and youth to ensure that they have the necessary data literacy to engage in intensive research while contributing to and benefiting from an open, interoperable system.

4.2 Recommendations for Collaboration

At this stage, two items shall be further discussed to help attracting indigenous people as stakeholders and to have starting points for further collaboration on the organizational level. This list serves as a starting point and does not make any claims of being complete.

- It is recommended that this pilot reaches out to and collaborate with the International Hydrographic Organization's Marine Spatial Data Infrastructure (IHO MSDIWG) and the Arctic Regional Hydrographic Commission (ARHC) Marine Spatial Data Infrastructure Working Group (MSDIWG).

- Interaction with indigenous people cannot be over-emphasized. A key lesson learned is that Indigenous people and their representative organizations must be an integral part of the design and development process if the SDI is to serve the needs of Indigenous peoples and broader society and benefit from Indigenous knowledge and observations. This process can start by working closely with Indigenous representative organizations such as the Permanent Participants of the Arctic Council and national organizations such as Inuit Quajisarvingat/The Inuit Knowledge Centre at Inuit Tapiriit Kanatami in Canada. They can, in turn, facilitate connection to regional and community level organizations and individuals. Further important stakeholders in this context are the Government of Nunavut, Nunavut Geoscience, Nunavut Planning Commission, Nunavut Impact Review Board, Nunavut water board, Municipal and local governments, Nunavut Tunngavik Incorporated and Regional Inuit Associations, Department of Indian and Northern affair, and the Department of Natural Resources Canada. Stakeholders should further include arctic researchers, the research licensing institutes (Aurora Research Institute), federal departments with continued interest/governance in the Arctic, Hunters and Trappers Committees, and Territorial governments and Infrastructure managers.

Chapter 5

Architecture

The architecture of an SDI is a multi-dimensional concept, including software, hardware, deployments, networks, operations, federations and many others. Figure 5.1 identifies a large number of aspects that play a role in architecture design and definition.

FIGURE 5.1: Architectural aspects that need consideration

The main goal of this Arctic Spatial Data Pilot is to demonstrate the value of an SDI for the Arctic to a broad range of stakeholders. As discussed in chapter 2.3.3, one of the best approaches to demonstrate value and increase stakeholders' adoption rates is an excellent implementation of an Arctic SDI serving all stakeholders needs. For this reason, this chapter discusses architecture perspectives and concentrates on a number of key aspects to support the future development and extension of any Arctic SDI without necessarily being a blueprint ready for implementation.

5.1 Data Infrastructure Evolution

There is already a considerable wealth of polar data available on the Internet through portals that vary considerably in function, scope, capability, and content. Appendix C provides an overview of some of these existing polar data portals. The polar data community is aware that there are many opportunities for improvement in how polar data are stored, managed, discovered, and delivered to users, and they are working collaboratively, with limited resources, to improve the situation.

The development of polar data infrastructure is occurring within a context of rapid growth in the provision of polar data and change in user expectations about access to and use of such data. The data available on the state of the planet is growing in precision, volume, velocity, variety, and value, increasing the complexity of scenarios for data exploitation, as well as the resources required by the communities using the data. A number of groups are developing innovative approaches to the creation of data platforms. These approaches share some common characteristics:

- Individual parameters by themselves are not nearly as valuable as integrated data sets. Therefore, the trend is to provide data platform users with access to a wide range of data types that they can be exploited together.
- With the explosion of the data that are available, data discovery and analysis is becoming increasingly challenging. As a result, the trend is to include sophisticated data visualization tools to enable data platform users to easily see and understand both the data they can utilize and the results of their analysis of that data.
- The quantity of data available, especially earth observation (EO) data, means that it is often not practical for each user to download the data they need to their local environment. Rather, the trend is to bring the algorithms to the data and only download the results of their calculations.
- Working with such large data sets is often computationally intensive. This means that modern data platforms need to provide users with highly capable information and communication (ICT) technology infrastructure for data processing, storage, and networking.

- Research is increasingly collaborative. Therefore, the trend is to combine data and computation capabilities with the tools required for such collaboration and the ensuing dissemination of research results.
- The increasing diversity of data sources and the need for scientific and operational communities to access data unfamiliar to them makes it essential that usable data quality information is available for all products.
- There is an aversion to lock-in with any one technology or supplier. Therefore, many data platforms use open source software where possible and are platform independent, often hosted in the cloud.

5.2 SDI Architecture Concepts

The ideal Arctic SDI architecture includes many facets that will need to be addressed in detail in the next phase of the Arctic Spatial Data Pilot. To avoid restricting too much the exploration activities planned for the second part of this pilot, this document will concentrate here on a number of rather high level views on the architecture, including key components of a future ideal SDI architecture, knowledge generation views, and technical perspectives demonstrating the current state of the art in terms of existing SDI components and installations to reflect real world component renewing cycles and operational realities.

5.2.1 Key Infrastructure Components

Taking the characteristics discussed in chapter 5.1 into account, it can be summarized that modern spatial data platforms are going far beyond traditional data portals by combining multiple functionalities and making them available (often in the cloud). The components of a modern data platform are shown in Figure 5.2, representing an ideal high level architecture of an integrated Arctic information system for observing, research, and community applications.

This high level architecture contains the following major components:

- **Community as a Service** (CaaS: Collaborative tools for users to publish, share and discuss their results, information, data and software/code on the platform. Social networking makes a new level of online collaboration among communities of practice possible.

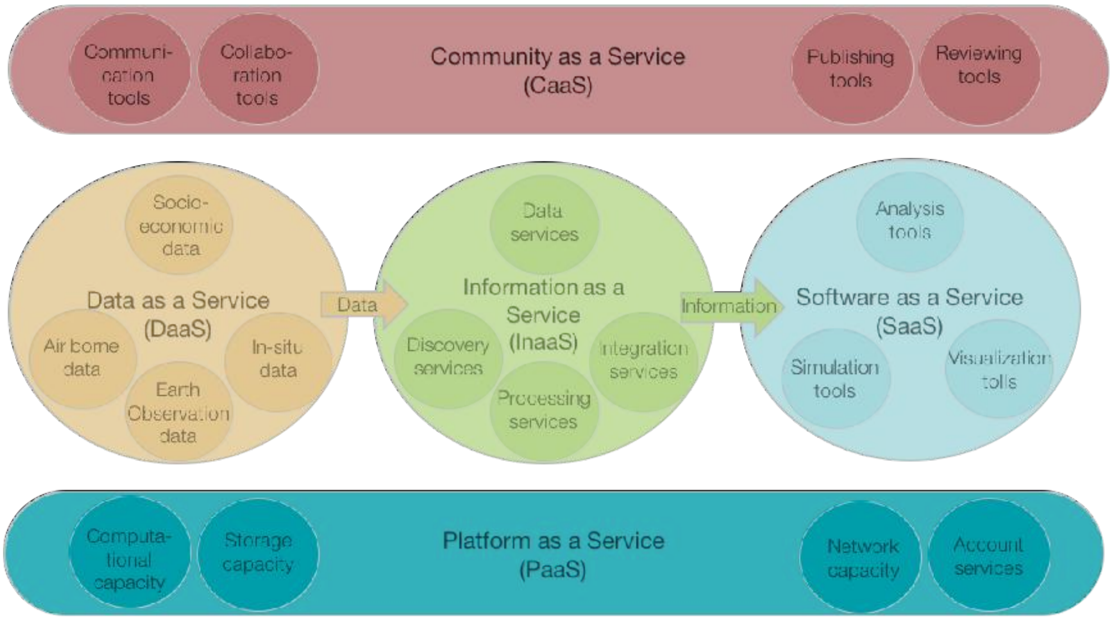


FIGURE 5.2: Components of a modern data platform, source: Polar Community (modified)

- **Data as a Service** (DaaS): On-demand data sharing through discovery, access, and transportation. Data sets can cover earth observation, air-borne and in-situ sensors, as well as other socio-economic data. The emergence of service-oriented architecture has rendered the actual platform on which the data resides less relevant.
- **Information as a Service** (InaaS): The ability to provide standardized and secure methods to create, manage, exchange, and extract meaningful information from all available data in the right format at the right time.
- **Software as a Service** (SaaS): Delivery and management of applications and tools by the platform or its users that are used remotely on the platform. Provides users with the capability to deploy user-created or acquired applications.
- **Infrastructure as a Service** (IaaS): The provision of computing resources, complemented by storage and networking capabilities, as shared resources, scalable on-demand, and enabling cost efficiencies.

Systems implementing these components integrate a number of functionalities that are crucial for modern spatial data infrastructures. Almost all elements are provided by services, which allows to serve the full spectrum from raw data access to highly customer-tailored products. This approach improves the ease of use for a large group of heterogeneous stakeholders using different platforms, including field workers with almost no Internet connectivity to scientists with fiber optics and supercomputers, or infrastructure novices to data processing experts. They reflect the growing complexity of research and analysis situations and provide the necessary communication infrastructure to connect distributed stakeholders. They take into account that data cannot be transferred to customers in all situations, but needs to be processed close to the physical data stores to minimize transport issues, which are either caused by limited physical bandwidth or by sheer amount of data that needs to be transferred. Cloud technologies support further evolvement of the infrastructure as resources can be added on demand.

5.2.2 Knowledge Chain Perspective

The first perspective discussed in 5.2.1 above illustrated a high level future architecture. A different position is taken in this chapter to bridge to the technical architecture as it can be realistically implemented within the next few years, while at the same time being extended step-wise and evolutionary with modern concepts as they reach production level maturity.

From a knowledge generation perspective, an SDI for the Arctic should ideally consider the full data value chain that includes connecting to earth observation and in-situ sensor networks, providing mechanisms for storing and hosting data (when hosting is not possible at the data source), make the data discoverable and enable use of the data in different media and accounting for both online and offline use. From an SDI for the Arctic, content may be disseminated to other global or national networks such as GEOSS, United States Geospatial Platform, and others. This increases visibility of the Arctic data and information products. Figure 5.3 illustrates this concept. The term *Portal* needs to be read as data access web-based API, not as a simple Web portal in the form of a Website.

At the same time, data might be served from storage and server components being part of other SDIs or data portal APIs. As illustrated in chapter 6 and further detailed in appendix C, a large number of data portals exist and will remain active in the future. It is one of the goals of this pilot to demonstrate the value of an SDI as being part of an application scenario that involves data registered and served at other SDI or portal API instances.

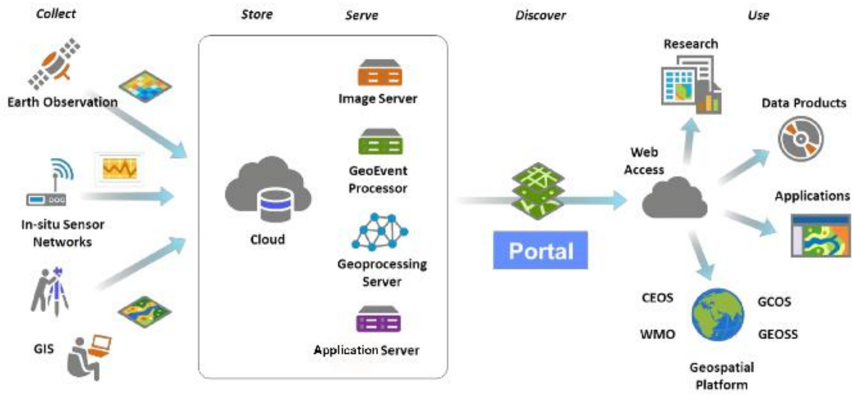


FIGURE 5.3: A knowledge generation perspective, source: ESRI (modified)

5.2.3 Classical SDI Technology Perspective

When it comes to SDI design, two important approaches have to be differentiated. They are not mutually exclusive and a chosen approach can still be complemented by the other. In fact, both approaches represent the two extremes of a given continuum, with most implementations featuring some level of middle course. Nevertheless, the architecture design differs depending on the preferred approach. The first approach focuses on the SDI for the Arctic as a closely architected infrastructure that provides data and apps as services. Thus, the defined architecture caters for a defined set of services (includes re-hosted services) that are operated and maintained by an SDI control board, i.e. a group with control over the individual components. The second approach focuses on infrastructures, platforms, and geoportals as they currently exist and emphasizes their integration into a loose confederation. Here, emphasis is on discoverability and integration based on open standards. The first puts more control into the hands of the control board, whereas the second provides more flexibility and distributed responsibilities. Key to both approaches is the strong adherence to standards to avoid vendor lock-in with limited flexibility and extensibility. It should be emphasized that both approaches can complement each other, i.e. they do not necessarily act in isolation, but support interfaces to allow mutual usage.

Closely Architected Approach

The first, closely architected approach is illustrated in figure 5.4. The platform itself consists of the infrastructure, the content, any number of APIs and SDKs, and application and content management tools. The actual applications or usually provided as external components or as web-based thin clients. Key here is the fact that the entire system focuses on the single platform concept, which means that the individual layers and implemented aspects are not particular characteristics of the closely architected approach. It is the way they are implemented and linked with each other.

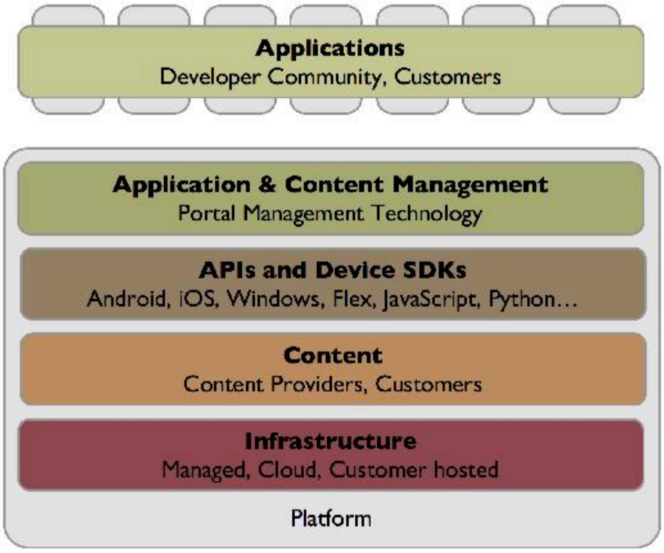


FIGURE 5.4: Closely architected approach, source: ESRI (modified)

The **infrastructure** includes the hardware and software needed to operate an SDI for the Arctic. The infrastructure design will need to take into consideration the different user scenarios, data sources (either managed by the SDI or coming from third party sources), appropriateness of cloud technologies, current and future IT policies, and existing hosting capacity. The SDI will need to account for offline use situations. To mitigate these, the SDI could consider using data appliances, container formats such as GeoPackage, or programs such as Geonetcast.

The **content** aspect of the Arctic SDI can be broken down into the following:

- The **geospatial data management** includes the technologies and workflows for managing vector and raster data that will be managed and used in the Arctic SDI. Following the best practices defined by the Federal Geographic Data Committee (FGDC) for the National Spatial Data Infrastructure (NSDI), the Arctic SDI would define the key spatial and non-spatial data layers that support the needs of the use cases of the Arctic SDI. For these data layers, data management and portfolio management policies and procedures need to be defined. This includes but is not limited to data models, data update frequencies, conflation of multi-source data, data quality assurance, and availability assurances.

- The **real-time data management** includes the technologies and workflows for ingesting and using real-time data feeds such as sensor feeds, vessel tracking, news feeds, and feeds from other systems relevant for the Arctic SDI.
- **Data integration with 3rd-party systems** allows to feed or consume data from the Arctic SDI. For this, a Web services approach using common service interface specifications that build on international standards from the World Wide Web Consortium (W3C) and the Open Geospatial Consortium (OGC) and others are recommended.

API's and SDK's: If data is the fuel of an SDI, Application Programming Interfaces (API) and Software Development Kits (SDK) form the engine that powers the applications and integration with 3rd-party components. Whatever platform is selected, it needs to offer an effective way to create and manage geospatial applications to developers. The offered APIs and SDKs shall support building web, mobile, and desktop apps that incorporate mapping, visualization, analysis, and more.

The **Application and Content Management** component provides the tools and concepts that allow for organizing the content in the Arctic SDI in logical and easy to understand groups of thematic or organizational structures. Content Management is typically done through portals.

Applications: The entire platform will be accessed through a number of applications that are tailored to the specific user audiences of the SDI. This component may include map applications for viewing, editing, analyzing, and collecting content. The applications may vary from templates that are used to tell stories around specific issues in the Arctic to advanced desktop GIS that connects to the metadata catalog and discovers web services and other content to consume. The important realization is that not all users will engage with an

SDI for the Arctic through the portal or through the applications managed as part of the SDI.

Loose Confederation Approach

The second approach is illustrated in figure 5.5. This approach, shown here with focus on service interfaces and encodings, identifies four main components, visualized using different background colors. The dark components at the bottom represent data sources such as geospatial feature data, geospatial raster data, map, sensor, and other data. This data is served by a number of services that belong to different classes, such as data access services, processing services, sensor web services, discovery services, or other services. These services make use of standardized data models and encodings. Visualization and decision support tools and applications make use the data provided by the various services in standardized formats.

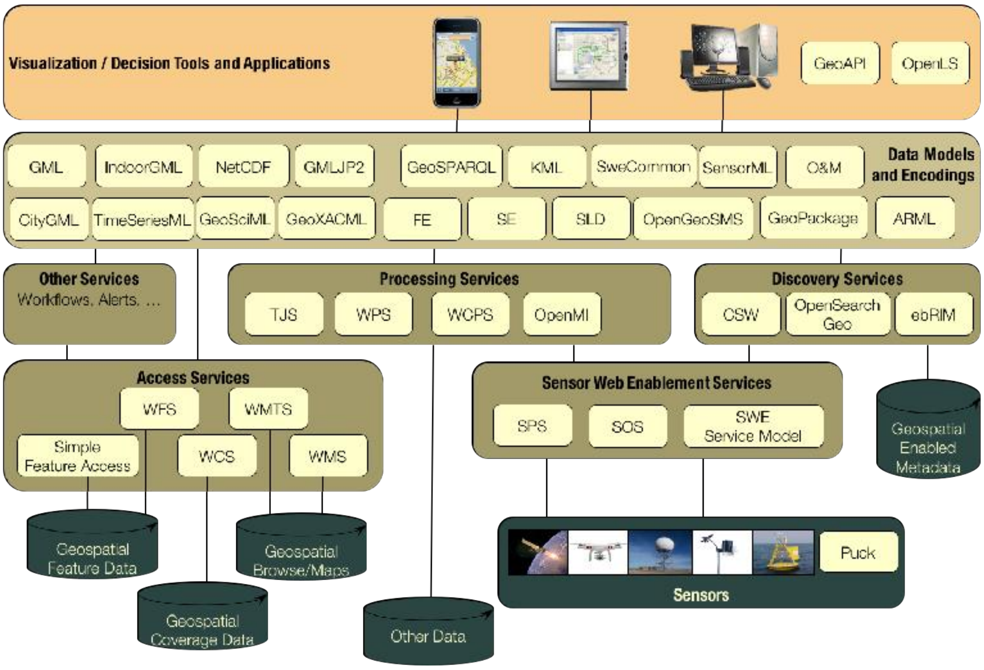


FIGURE 5.5: Loose confederation approach, source: OGC

This approach concentrates on service interfaces and encodings. It allows an entirely decoupled and loosely federated infrastructure with minimized necessary a-priori knowledge required to interact with the various components. This approach leaves aspects such as maintenance, service configuration etc. to the service operators, i.e. this functionality is not part of the architecture view, as it is irrelevant for the actual SDI. This is in contrast to the closely architected SDI concept, where management tools and content tools allow control over more than a single SDI component. The environment illustrated here needs to be enriched with security settings, which usually require some sort of higher level organization if features such as single-sign on shall be supported (otherwise service consumer would need to register with every services, which works in principle, but is not very practical).

Ontology Driven Approach

A third approach specializes on particular aspects and can be used to illustrate the concept and power of semantic annotations based on well-defined ontologies. The concept, illustrated here with some example tools, query languages and runtime environments, would allow for enhanced searches for associations within an SDI for the Arctic. The ontology driven approach is only introduced here briefly. For more details, please be referred to Appendix E.

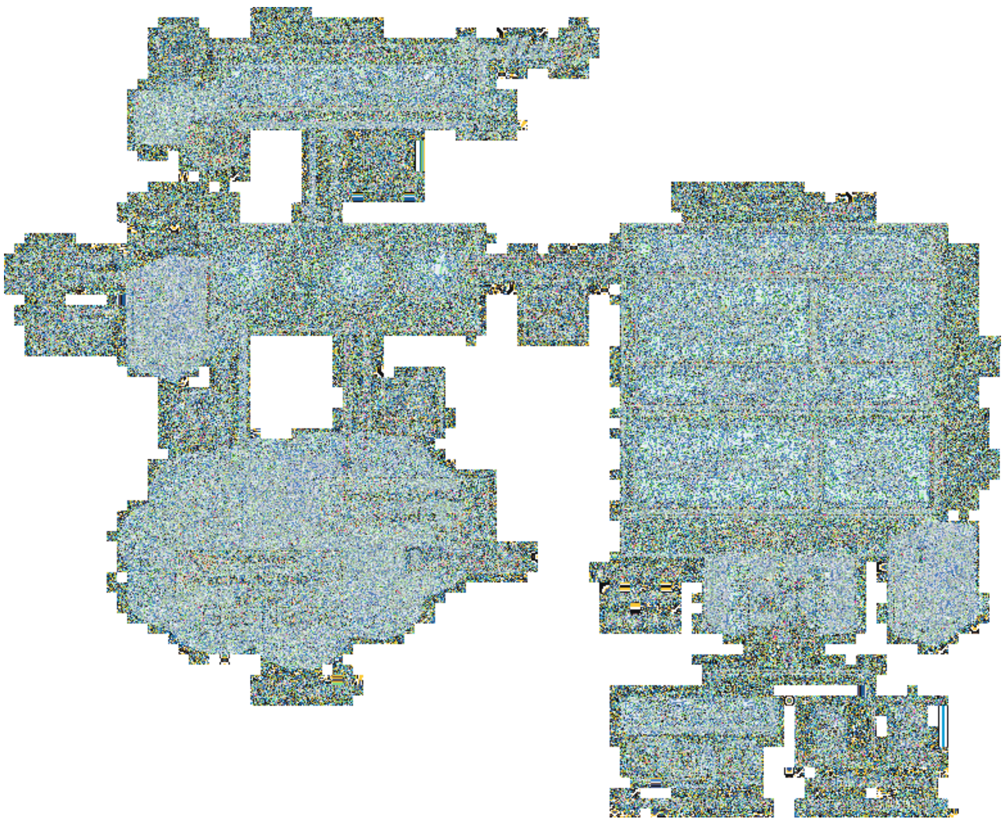


FIGURE 5.6: Ontology and triple store based approach, source: University Grenoble Alpes

This approach allows to visualize associations between stakeholders, data, and processing services, which would allow valuable insights for other stakeholders with similar requirements, as they could consider copying or adapting workflows executed by others.

5.2.4 Architecture Requirements

Independently on the chosen approach, a number of aspects have been repeatedly identified as being relevant for a successful SDI. These are usually complemented with the standing request for *openness* as illustrated in figure 5.7. Openness usually refers a number of aspects that circle around the fact that an element is openly (in the sense of publicly and royalty free) available and reusable, developed in an open process, accessible at minimum costs (in terms of data pure reproduction costs or even no costs).

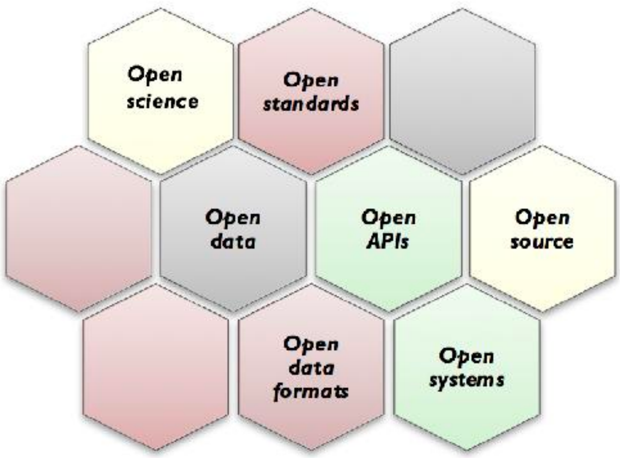


FIGURE 5.7: Aspects of openness

Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society, amateur or professional. Open systems includes open source work and Github resources, choices in hardware, operating systems, Cloud, databases, developer tools, direct links to non-GIS systems such as CAD and BIM, etc.. Open standards include standards as provided by OGC, IHO, ISO TC/211, DGIWG, GWG, CGSB, FGDC, OASIS, W3C, ASPRS, etc.) and open specification (widely used but not yet adopted by SDOs openly published technology such as GeoJSON, Geoservices REST API, etc.).

In addition to these general requirements (that are purely stated here, not judged, as we fully acknowledge that some data cannot not be openly available but needs to be protected to ensure privacy constraints), an ideal SDI architecture shall be designed to provide for:

- **Registry and discovery**
  - rapidly discover and access information, products and data;
  - Architecture shall support search mechanism that go beyond metadata based key-word search, as metadata is never complete and often hard to maintain
  - Auto-registry for sensors (both remote and in-situ)
  - Search engine for finding and browsing data, services, and metadata. It should be adaptable to allow for basic quick searches through detailed searches using multiple criteria including: geography, time, organization, physical parameter.
  - Non-mapped search results (e.g. technical reports, multimedia) should be associated with mapped search results and viewable in the web browser.
- **New functionality and extension**
  - Easily publish/reference information, products and data into the SDI.
  - Integrate new functionality;
  - The vision for SDI deployment is to allow users to discover (search), view, assemble and obtain desired data and services for a particular area of interest without needing to know the details of how the data and services are stored and maintained by independent agencies, organizations and data custodians.
  - Find the right balance between being prescriptive while remaining agile to allow for easy integration of upcoming technologies. This is particular important as the SDI for the Arctic needs to allow for remote communities to participate. This requires the SDI for the Arctic to be implemented as a lose confederation of portals and platforms discoverable by open specifications and standards rather than being a highly architected infrastructure with data and applications being available as hosted/re-hosted services.

  - An SDI for the Arctic is by nature federated. It should be as transparent as reasonable to an end-user as to where the information being accessed is sourced from within the federation.
  - Consider adding an online information network with an ontology-based interface on top to visualize databases and information sources content. The ontology-based approach would allow for efficient searches once all data and operation concepts are annotated. The ontology-based approach would allow for illustration and processing of stakeholder-data relationships, or stakeholder-processing relationships, which could provide valuable insight for other stakeholders with similar requirements, as process could be copied or adapted more easily.

- **Low bandwidth and offline usage**
  - Support both online and offline use
  - One of its key uses within an SDI for the Arctic is the ability to exploit low cost mobile devices, such as Android tablets, in the field for monitoring, gathering and updating data in areas that have no, or poor data communications. Field workers preload data they may need on to their mobile device using a GeoPackager application. They can then go into the field and add or update current data on the device. When they return to an area with data communication, the geoPackage will synchronize with the original data.
  - Should allow for proxies that optimize data for transport over limited bandwidth connections or other specific purpose tools
  - Large datasets need to be made available in very efficient ways to support low bandwidth situations
  - Ideally, data could even be shipped on hard drives to customers with very limited internet connectivity. These datasets shall be made available as being directly served from a Web server, i.e. data storage is transparent to the end user. At least end user experiences shall differ minimally.
  - Downloadable datasets in standard formats

- **External systems and formats**

  - Content may be disseminated to other global or national networks such as the Group on Earth Observation, United States Geospatial Platform, and others. This increases visibility of the Arctic SDI data and information products.
  - Provide connectivity to legacy/heritage systems;
  - Support for scanned documents that provide valuable historic data sets, including maps, forms, other tabular data (both machine and hand written), or hand-drawn sketches
  - Support for documentary videos, oral histories, and other sources beyond purely numerical data
  - Enable visualization of information in a geospatial, data analysis presentation environment and temporal context;
- **Tailoring**
  - Work within the desktop or mobile environment.
  - The architecture has to allow for multi-linguism and appropriate character sets.
  - Target users of the SDI Geoportal will come from a diversity of backgrounds. The efficacy of the portal to accessing information by the uninitiated, the *man-of-the-street* has been proven is key to a successful, i.e. well-used SDI.
- **Key service functionality**
  - Mapping interface showing search results. Map should be interactive: pannable, zoomable, changeable projection. Mapped items should be interactive: obtain metadata by clicking/hovering, get data values by clicking/hovering.
  - Basic analysis and visualization tools, e.g. navigating long time-series, statistical analysis on selected data sets or subsets.
  - Allow for GEONETCast integration. GEONETCast is a global network of satellite based data dissemination systems providing environmental data to a world-wide user community.

Chapter 6

Data

The Arctic provides a unique environment when it comes to data acquisition and collection. Due to ice, snow, and sun angle, the data collection season is usually short. Within this season, complex logistical challenges such as remote locations with little access to supplies, technical assistance, medical facilities have to be mastered, usually yielding in high mobilization and demobilization costs. Data collection campaigns need to handle fast changing, extreme weather conditions in harsh environments, which requires sophisticated health, safety, and environment (HSE) training and expertise. Not least because of potential interactions with marine mammals (haul out and migrations).

This chapter addresses data aspects from two perspectives. The first (6.1) is from the data consumer side, who has particular needs on the type and format of the data; with further requirements on update rates, reliability, quality and other characteristics. The second (6.2) part addresses data sets that are already available. This part can only highlight a number of data sets that are ideally available at standardized interfaces. A more complete inventory of available data sets shall be developed in the course of this pilot project, with catalog services being used as service and data registries that become integral part of the planned demonstrations.

6.1 Data Requirements

The requirements for information in the polar regions are being driven by a broad range of scientific, operational, and societal imperatives. Researchers, practitioners and residents are involved in a host of studies on changes taking place across many domains, including climate, oceans, atmosphere, and ecosystems, which have significant impacts in the regions and, through complex earth system connections, worldwide. The drivers include both national and international science policies, strategies, and programmes that contribute to an understanding of the changes taking place in the polar regions and shape policy responses. To better understand the following requirements on data sets, figures 6.1 and 6.2 provide some examples of polar science activities.

Operations in the polar regions take place in some of the most difficult conditions on Earth. Those involved in these operations, such as shipping and fisheries companies, offshore oil and gas operators, research organizations, coast guards, and local communities, require access to reliable and often near real-time information to plan and undertake their activities. Drivers of information requirements include a range of regulations, standards, and policies (such as the new Polar Code) aimed at ensuring safety of life and mitigating negative environmental impacts. Examples of polar operational activities are contained in figure 6.3.

More information on polar data requirements and sources of information can be found in the Polaris study of the European Space Agency [1].

[1]The Polaris Study reviewed user requirements for polar environmental information, considered current and proposed sources of such information from space-based and in-situ sensors, evaluated the information gaps and the impact of filling those gaps with new integrated products and services, and provided a preliminary discussion of the considerations that will shape new satellite missions to fill the gaps.

FIGURE 6.1: Examples of polar scientific activities that drive information requirements, part 1, source: Polar Community

FIGURE 6.2: Examples of polar scientific activities that drive information requirements, part 2, source: Polar Community

Theme	Examples of Types of Activities
Environmental Impact Assessment	▪ Supporting the responsible development of major infrastructure or resource development projects ▪ Assessing and mitigating the operation of such projects
Engineering Design	▪ Design of buildings and structures for installation in changing permafrost conditions ▪ Design of offshore drilling and production platforms for safe and effective deployment in ice-covered waters
Safe Navigation and Operations	▪ Navigation of vessels through hazardous ice-covered waters ▪ Avoiding collisions with icebergs in operation of offshore oil and gas exploration and production platforms ▪ Navigation to and along the sea ice edge for traditional northern hunting and fishing
Risk Management	▪ Assessing the risks of subsidence around buildings, pipelines and structures in permafrost areas ▪ Assessing and mitigating the risks of flooding due to ice-jammed rivers
Emergency Response	▪ Developing and maintaining a common operating picture (COP) between response organizations ▪ Expeditious movement of responders and their equipment from bases of operation to the emergency site
Weather Forecasting	▪ Observing and modelling weather patterns to improve short-term weather predictions in support of operations in the polar regions
Climate Change Adaptation	▪ Establishing new regulations and standards, investing in new infrastructure, and enhancing operational capabilities in reaction to changes in the polar climate and its impact on southern latitudes

FIGURE 6.3: Examples of polar operational activities that drive information requirements, source: Polar Community

There are a number of general requirements that apply to many data types, such as real-time data availability, integration of local knowledge with physical environmental data (which is a difficult problem but one that represents a rich source of information), or data quality, coverage and resolution. Engaging local communities and academia in a multi stakeholder process aimed at defining a roadmap for Arctic data development has been suggested. Arctic countries could jointly pull resources and work cooperatively to start filling the most important data gaps. Based on various gap-analyses as well as on the scientific and operational activities mentioned before, datasets for the following key data categories have been identified to be required within an effective SDI for the Arctic at minimum. This is not an exhaustive list and the list is in no particular order:

- **Hydrographic data**
  - Hydrographic data should be an integral part of any SDI as it relates to the navigational or other water bodies within a given country or region and represents a key element of the marine component of a National SDI (NSDI) - if the legal situations allow. M(arine)SDI is the component of an SDI that encompasses marine geographic and business information in its widest sense. This would typically include seabed topography (bathymetry), geology, marine infrastructure (e.g. wrecks, offshore installations, pipelines and cables), administrative and legal boundaries, and areas of conservation, marine habitats and oceanography. Much of this information resides in the NOAA as source data (e.g. dense bathymetric data) and/or product data (e.g. ENC data, digital nautical publications, digital elevation models) complete with metadata (data about data).
  - Sea ice presence, thickness, and velocity
  - Ocean currents, waves
  - Water properties including water temperature, salinity, fluorescence, turbidity, dissolved oxygen, chlorophyll, suspended material, chromophoric dissolved organic matter
  - Crowdsourced hydrographic data as collected by shipping and cruise ship companies
  - Nautical charting and seabed data

- Horizontal and vertical datum, maritime baseline, seabed character, marine boundaries, offshore minerals, shoreline, seabed infrastructure, oceanographic features
- Iceberg data

- **Land and coast data**
  - Topographic base maps (high and medium resolution DEM) and coastal mapping
  - Land cover, offshore cadaster, land ownership, flood hazards, and gazetteer.
  - Optical and radar imagery with long term historic imagery to provide valuable insights into changes and near real-time imagery to monitor the region
  - High resolution map of Arctic permafrost (e.g. for diverse areas such as infrastructure, home building, residential and commercial insurance, oil and gas distribution, or disaster preparedness and response)
- **Cryosphere** data such as areas of snow, ice and frozen ground to support research on warming permafrost, reduction in snow cover extent and duration, reduction in summer sea ice, increased loss of glaciers and the break-up of ice shelves
- **Vessel Tracking**. With the dramatic increase in ship traffic in the north, it will be necessary to ensure that there is continual near real-time monitoring of traffic via AIS or remote sensing together with historic data on vessels' voyages. Captured (real-time) data should include vessel position, static, voyage, and historical track and position information.
- **Terrestrial ecosystems data** to better understand extent and speed of changes in the Arctic
- **Wildlife** data to understand changes in wildlife populations such as habitat and migrations data
- **Communities**. Permafrost, reduced sea ice and unpredictable weather patterns affect the use of traditional knowledge for hunting and traveling. These communities can provide valuable input into the SDI as they are the closest to changes that may be occurring.

- **Human Health**. A shift away from a traditional lifestyle has led to a number of health problems for Inuit relating to nutrition and chronic illness. Data associated with health of local communities needs to be available.
- **Meteorology and Climate** including wind velocity and direction, air temperature, humidity, and atmospheric pressure as well as climate parameters and indices

In general, all data should be available in or transformable to different projections using different datums for efficient map productions or integrated processing and analysis operations. Important projections and corresponding datums for the North American part of the Arctic include Yukon Albers using NAD83, UTM Zone 7N to 10N using NAD83, GCS CSRS using NAD83. An earlier study by FGDC identified that most Web mapping services for the Arctic countries support EPSG:4326, with EPSG:4269 and EPSG:4267 being also popular projections and EPSG:32633 being well supported by services from Norway and Sweden. The use of Web Mercator (Auxiliary Sphere) WGS84, though often used in Web applications, has some serious precision implications and should be avoided.

Northern communities have valuable local data and in many places have established digital data centres to enable sharing and to preserve knowledge, however they are not necessarily adhering international standards, producing metadata or following best practice for data management. As part of this pilot project to test interoperability, it is recommended to:

- Identify community based data repositories in the Arctic;
- Inform and provide training to data managers on standards and Web services, as well as best practices of data management;
- Test out the abilities to incorporate Indigenous Knowledge (IK) into standard database services;
- Depending on capabilities, include their data hub as a component in the pilot study.
- Identify and discuss map projections important for local as well as pan-arctic projects.

6.2 Data Identified in RFI Responses

Despite many areas are still lacking sufficient data coverage (or available data is of coarse resolution or low quality), quite a number of data sets are already available. The following list provides an overview of data sets that have been identified in RFI responses, direct communication with Arctic data experts, described by external publications, or referenced in literature. In addition, there is a long list of data portals online (see Appendix B for details). Others have been assembled by a just recently finished yet unpublished study by EU-PolarNet [2]. Further initiatives have used crawlers to detect standardized Web services that provide Arctic data. The following sections provide an overview in alphabetical order.

6.2.1 Esri

A sample of datasets, developed by Esri and/or partner agencies, that could become part of the Arctic SDI include:

- **Living Atlas**: The Living Atlas of the World is a dynamic collection of thousands of maps, data, imagery, tools, and apps produced by ArcGIS users worldwide, as well as by Esri and its partners. It is a curated subset of ArcGIS Online information items contributed and maintained by the ArcGIS community. More than 40 datasets contained in the Atlas cover the Arctic region.
- **Dynamic Landsat coverage for the Arctic**: Esri has access to thousands of Landsat scenes covering the Arctic and plans to make them available for viewing and analysis as part of a forthcoming app developed in cooperation with the USGS and the Polar Geospatial Center (PGC).
- **Digital Elevation Models**: Esri is working with several partners to deploy a DEM geoprocessing app for the Arctic that leverages a gamut of DEMs from third parties. (see Appendix B)
- Natural Resources Canada has **detailed topographic information** for the North available for download from the GeoGratis website. Much of this

[2] EU-PolarNet is the world's largest consortium of expertise and infrastructure for polar research. Seventeen countries are represented by 22 of Europe's internationally-respected multidisciplinary research institutions.

data is included in the Esri World Topographic Map via the Canadian Community Map program.

- **Fisheries and Oceans Canada** has bathymetric data available for Arctic areas that is included in the Esri World Bathymetric map via the Canadian Community Map program.
- **Arctic Ocean Base Map - Beta**: The Arctic Ocean Basemap is designed to be used as a base map for the Arctic Ocean by marine GIS professionals and as a reference map by anyone interested in ocean data for the Arctic region.

6.2.2 exactEarth

exactEarth provide access to data and information services through certified Open Geospatial Consortium (OGC) compliant Web Map Services (WMS) and Web Feature Services (WFS) which allow on- demand access to exactAISő data to retrieve vessel position, static, voyage, and historical track and position information. The WMS services generated from exactEarth allow users to get requested images of exactAISő data in a variety of formats including PNG, JPEG, KML, etc. The WFS services generated from exactEarth allow users to get features (data) in a variety of standard formats including CSV, GML, GeoJSON, and Shapefile. Particularly for predictive analytics, exactEarth explores OGC WPS (Web Processing Services) as additional options to the current WMS/WFS delivery framework.

Accessing Web Services: Geospatial Web Service (GWS) is available from the exactEarth Data Processing Centre (DPC). GWS uses HTTPS and user authentication via user token or user name password to access information and is available online.

Data Sets: Data sets returned by GWS in response to requests made by end users contain fields in a specific sequential order. Data ingest methods that depend on field order, specifically commas separated values (.csv) format, need to take this into account. Descriptions on data content, including field descriptions and schemas are available for each of the three (3) services available; Latest Vessel Information, Historical Vessel Tracks, and Historical Vessel Positions.

6.2.3 Federal Geospatial Platform Project

Data and services identified by the Federal Geospatial Platform Project are described in figure 6.4 on the following page. The links listed in the figure are repeated here to allow direct follow up:

Migratory Bird Species ECCC Data Mart OGP Catalogue

First Nation, Inuit Community, INAC ESRI REST

Offshore Oil and Gas Rights INAC ESRI REST

Northern Major Projects NRCan ESRI REST

North American Cooperation on Energy NRCan ESRI REST OGP Catalogue

GeoScience data NRCan WMS NRCan Surficial Geology WMS

Canadian Geochronology Knowledgebase NRCan ESRI REST

National Marine NRCan NMCA WMS NRCan NP WMS

National Pollutant Release Inventory ECCC ESRI REST

Canadian Environmental Sustainability Indicators ECCC ESRI REST

Title	Description	Link
Migratory Bird Species at Risk Range Maps for the Northwest Territories and Nunavut	Range Maps for all Migratory Bird Species At Risk within the Northwest Territories and Nunavut. Extensive point observation data collected by the Canadian Wildlife Service helped to identify breeding ranges for Migratory Bird Species at Risk in the Northwest Territories and Nunavut.	ECCC Data Mart OGP Catalogue
First Nation, Inuit Community, Tribal Council Locations	This dataset is Indigenous and Northern Affairs Canada (INAC) official source for First Nation, Inuit Community and Tribal Council geographic location on maps.	INAC ESRI REST
Offshore Oil and Gas Rights	The Oil and Gas Rights dataset contains the digital boundaries for existing exploration licences, significant discovery licences, production licences, former permits, former leases and the Norman Wells Proven Area.	INAC ESRI REST
Northern Major Projects	Location of Major Infrastructure, Resource, Oil and Gas Projects in Nunavut, Northwest Territories and Yukon from the Canadian Northern Economic Development Agency.	NRCan ESRI REST
North American Cooperation on Energy Information, Mapping Data	Mapping Resources on energy facilities implemented as part of a North American trilateral cooperation on energy information between the Department of Energy of the United States of America, the Department of Natural Resources of Canada, and the Ministry of Energy of the United Mexican States.	NRCan ESRI REST OGP Catalogue
GeoScience data	Various Geoscience datasets from the Geological Survey of Canada. • Airborne Gamma-Ray Spectrometry (AGRS) data • Canadian Gravity Anomaly Data Base • Aeromagnetic surveys • Surficial Geology	NRCan WMS NRCan Surficial Geology WMS
Canadian Geochronology Knowledgebase	A compilation of publicly available reports of geochronological information from the Geological Survey of Canada.	NRCan ESRI REST
National Marine Conservation Area and National Parks	Parks Canada manages more than 43 NPs and is responsible to protect NMCAs	NRCan NMCA WMS NRCan NP WMS
National Pollutant Release Inventory	The National Pollutant Release Inventory (NPRI) is Canada's public inventory of pollutant releases (to air, water and land), disposals and transfers for recycling. Under the authority of the Canadian Environmental Protection Act, 1999 (CEPA 1999), owners or operators of facilities in Canada that meet the published reporting requirements are required to report to the NPRI. Reported pollutants include toxic substances, air pollutants and other substances of concern.	ECCC ESRI REST
Canadian Environmental Sustainability Indicators	The Canadian Environmental Sustainability Indicators (CESI) program provides data and information to track Canada's performance on key environmental sustainability issues.	ECCC ESRI REST

FIGURE 6.4: Selected Web Services relevant to the Arctic Spatial Data Pilot, source: Federal Geospatial Platform Project

6.2.4 Geographical Names

The Geographical Names Board of Canada (GNBC) is Canadas national coordinating body responsible for standards and policies on place names. Natural Resources Canada (NRCan) provides infrastructure and support for the Canadian Geographical Names Data Base (CGNDB) as the authoritative national database of place names, and a key component of the Canadian Geospatial Data Infrastructure (CGDI) as a fundamental API requirement. NRCan consolidates geographical names data, spatial delineations of features, and new naming decisions provided by the naming authorities of the GNBC in the national database. The geographical names product files covering all Canadian provinces and territories may be downloaded in Shapefile, KML, GML, and CSV format from NRCans GeoGratis portal as well as from the Government of Canadas Open Data Portal. In addition, NRCan offers an Application Programming Interface (API) as a means for direct access to the CGNDB for customized searches.

On the US side, the US Board on Geographic Names provides Web service access to geographical names via the GNIS system. The Geographic Names Information System (GNIS) is the US federal and national standard for geographic nomenclature. The U.S. Geological Survey developed the GNIS in support of the U.S. Board on Geographic Names as the official repository of domestic geographic names data, the official vehicle for geographic names use by all departments of the Federal Government, and the source for applying geographic names to Federal electronic and printed products.

6.2.5 Harris Corporation

Data types available through Harris which should be considered for integration include access to Japan Meteorological Agency (JMA) Himawari-8 and NOAA GOES-R sensor data through Harris WxConnect ground station/s, access to real-time satellite AIS data steam, and access to current commercial or government imagery data through an OGC service interface. In addition, it is possible to provide data analytics based on the real time data described above. There are prototype analytics to process the real time data to provide derived information. These currently include COTS imagery analytics and test cases for vessel track intersection, vessel deviation from reported course and speed.

6.2.6 National Geospatial-Intelligence Agency

The National Geospatial-Intelligence Agency (NGA) has a variety of datasets currently available to the public in their NGA GEOINT Services Arctic Support page found at here.

Users can discover and access the following datasets and information in the Arctic:

- **Digital Elevation Data** - a Digital Elevation Model (DEM) is a digital representation of ground topography or terrain that can be represented as a raster or as a triangulated irregular network. DEMs of 5-meter and 2-meter resolution are currently available.
- **Reference Graphics** - (1) Arctic Maritime Reference Graphic depicting ice extent, energy resource potential, search and rescue areas, an elevation view of the region, as well as maritime boundaries and IHO World-Wide Navigational Warning Service (WWNWS) and (2) The Arctic: A GEOINT Perspective that provides information about the changing landscape of the Arctic and a depiction of shipping routes, energy resources and the change of ice extent from 1980-2100.
- **Nautical Charts** - a representation of a portion of marine navigable waters and adjacent coastal area used by vessels sailing in the Arctic. Twenty-four charts are provided for the Arctic region.
- **Sailing Directions** publications that describe coastal weather, currents, ice, dangers, coastal features, port descriptions, as well as a graphic key of navigation products are available. Four publications are provided for the Arctic region.
- **World Port Index** a publication that gives the location, characteristics, known facilities, and available services of most ports and shipping facilities and oil terminals throughout the world.

Furthermore, the following are also available as NGA Data Services in several standard protocols to include the OGC developed Web Map Service (WMS):

- Arctic Sea Routes
- Natural Resources and Industry
- NAVTEX Sites

- Arctic Search and Rescue Zones
- Exclusive Economic Zones
- Sea Ice Index
- Energy Resources Potential
- Airfields
- Arctic Maritime Boundaries
- Natural Earth Features
- Arctic Basic DEMs
- Arctic World Port Index

NGA also provides Maritime Safety Information which contains a wealth of Safety of Navigation information including the following in KMZ and other formats:

- WWNWS Broadcast Warning Messages specifically HYDROARC Navigational Warnings, which contain information about persons in distress, or objects and events that pose an immediate hazard to navigation within the five Arctic Navigation Warning Areas (NAVAREA).
- Anti-shipping Activity Messages (ASAM) providing the locations and descriptions of specific hostile acts against ships and mariners.
- Mobile Offshore Drilling Units (MODU) facilities designed or modified to engage in drilling and exploration activities.

In addition to what is currently available, NGA is working to make available the following for the Arctic:

- The first ever publically available, high resolution, satellite-based elevation map of Alaska by this spring and the entire Arctic by summer 2017.
- Human Geography data including languages.
- Videos focused on Arctic boundaries.
- Improved gravity measurement and modeling.

The data available is provided in a variety of open formats and standards to support the ultimate users needs in the Arctic. Any of these datasets could potentially be included in an SDI for the Arctic given the key drivers and communities of interest that develop in the region.

6.2.7 US National Oceanic and Atmospheric Administration Office of Coast Survey

NOAA's Electronic Navigational Charts (ENCő) are vector data sets that represent NOAA's newest and most powerful electronic charting product. NOAA ENCs conform with the International Hydrographic Office (IHO) S-57 international exchange format, comply with the IHO ENC Product Specification, and are provided with incremental updates that supply Notice to Mariners corrections and other critical changes. In addition to downloading single ENCs in S-57 format, ENCs can be accessed via:

- ENC Direct to GIS which allows users to display, query, and download all available base editions of NOAA ENCő data in a variety of GIS/CAD formats, using Internet mapping service technology.
- NOAA ENCő Online which optimizes the viewing of the entire ENC suite, using the display rules defined by the IHO S-52 specifications for chart content and display aspects of ECDIS.

Both ENC Direct and ENC Online provide OCG WMS capabilities.

NOAA also provides the ability to access up-to-date nautical charts via the Office of Coast Survey's new Chart Tile Service prototype. The NOAA Chart Tile Service provides standardized nautical chart tilesets for the public, eliminating the need for application developers to regularly undergo the cumbersome process of transforming NOAA BSB files into tilesets. It provides geo-referenced charts compatible with the Web Map Tile Specifications (WMTS) and Tile Map Service Specification (TMS). All tilesets are published on a weekly basis.

NOAA identified the following 10 datasets, which should be included in an Marine-SDI. To provide an idea of what datasets are important nationally, the following list was derived from regional input from the US Northeast and Mid-Atlantic regions, as well as a national Coastal Zone Management survey conducted by NOAA.

1. **Marine Boundaries**: jurisdictional boundaries, territorial sea, submerged lands, agency jurisdictions. Maritime Limits and Boundaries for the United

States are measured from the official U.S. baseline, recognized as the low-water line along the coast as marked on the NOAA nautical charts in accordance with the articles of the Law of the Sea. The Office of Coast Survey depicts on its nautical charts the territorial sea (12 nm), contiguous zone (24 nm), and exclusive economic zone (200 nm, plus maritime boundaries with adjacent/opposite countries). Digital data is available in a variety of formats, including a dynaminc OpenGISő Web Map Service (WMS) which can be accessed online.

2. **Bathymetry**: nearshore lidar, medium and high resolution. NOAA's National Center's for Environmental Information (NCEI) archives data and products that illustrate the depths and shapes of underwater terrain with underwater maps, digital elevation models, multibeam data, and more. These include:

    - Bathymetric Data Map Interface
    - Digital Elevation Models
    - ETOPO Gridded Relief Models
    - International Hydrographic Organization Data Center for Digital Bathymetry
    - Ocean and Coastal Mapping Data for Northeast U.S. Sandy Recovery

    Many of these datasets, including those conducted by NOAA's National Ocean Service (NOS) are publically available via REST services through NOAA's GeoPlatform. Coastal topographic lidar datasets, including topobathy lidar is available from NOAA's Office of Coastal Management's DigitalCoast.

3. **Shoreline**: NOAA's National Geodetic Survey provides the Continually Updated Shoreline Product (CUSP) and the National Shoreline through the Shoreline Data Explorer. CUSP is also available as an OGC WMS.

4. **Fishing Data**: Content of the dataset: Vessel Trip Report and Vessel Monitoring System databases in collaboration with commercial fishing stakeholders to illustrate the distribution timing, and intensity of commercial fishing activity, with attributes for gear type and harvest species.

5. **Marine Wildlife** (whales, birds, turtles, bats) data: distribution, probability of occurrence, and migration paths. Datasets on marine mammals and turtles are available from NOAA's Office of Coastal Management's

DigitalCoast. Additionally, NOAA's National Marine Fisheries Service (NMFS) produces GIS data depicting areas designated as critical habitat for species listed as threatened or endangered under the U.S. Endangered Species Act.

6. **Benthic Habitat**: corals, seagrasses, submerged aquatic vegetation (SAV). Benthic habitat datasets are available from NOAA's Office of Coastal Management's DigitalCoast. Additionally, NOAA's National Marine Fisheries Service (NMFS) produces a GIS Data Inventory on Essential Fish Habitats (EFH).

7. **Restricted Areas**: Information pertaining to uses of the marine environment by the Department of Defense for training, classified or restricted areas, with unexploded ordnances, etc. Many restricted areas are depicted on NOAA's navigational charts. NOAA's ENC Direct to GIS provides the ability to access NOAA ENC data in a variety of GIS/CAD formats. ENC Direct provides WMS capabilities for the six different ENC scale bands: WMSoverview, WMSgeneral, WMScoastal, WMSapproach, WMSharbour, WMSberthing

8. **Vessel Traffic Data**: Automatic Identification System (AIS) shipping data, maintained channels. AIS data records for U.S. coastal waters are available for download thorough MarineCadastre.gov for calendar years 2009, 2010, 2011, 2012, 2013, and 2014. Records are filtered to one minute and formatted in zipped, monthly file geodatabases by Universal Transverse Mercator (UTM) zone. Vessel Density plots are also available as a map service through NOAA's GeoPlatform.

   Information about shipping channels can be found on NOAA navigational charts. Both Coastal Maintained Channels and Shipping Lanes and Regulations are available as national-scale, standardized data sets based on several Federal Geographic Data Committee (FGDC) and Open Geospatial Consortium standards from NOAA's ENC Direct to GIS: Coastal Maintained Channels, Shipping Lanes and Regulations.

9. **Sand, Gravel, and Cobble Mining Locations**: Offshore energy and mineral extraction leases, and planning areas; Resource areas where materials are actively mined from the seafloor; Areas which have been identified as potential "donor" sites for materials.

MarineCadastre.gov is a joint BOEM and NOAA initiative providing authoritative data to meet the needs of the offshore energy and marine planning communities. Access to ocean use and planning areas can be found here.

10. **Human Uses**: recreational fishing, boating, diving, surfing, swimming, nature viewing, ecotourism, aquaculture, etc. Could also contain submarine cables for telecommunications and energy, including modern fiber optic.

    NOAA's National Marine Protected Areas (MPA) Center has developed and applied a participatory GIS process to gather ocean use data through workshops that engage local and regional ocean experts. It is intended to be a flexible and mobile approach that can be refined and adapted for any region or domain and to address multi-scaled management decisions.

6.2.8 Arctic Hub

The Arctic Hub is a collaboration space for people interested in Arctic observation. The ArcticHub cyberinfrastructure provides an impressive Arctic and related data management directories and initiatives overview.

6.2.9 Polar Community

A number of technological trends are providing a flood of new data concerning the polar regions. Of particular relevance are space-based technologies such as earth observation, satellite telecommunications, global navigation satellite systems (GNSS), and ship-borne automatic identification systems (AIS), and a wide variety of in-situ observational networks. Each has a role to play in monitoring the vast and harsh polar regions and each is undergoing significant improvements in capabilities. However, ensuring the interoperability of these diverse data streams requires the development and implementation of appropriate data standards.

The International Arctic Science Committee (IASC) Sustaining Arctic Observing Networks (SAON) has already initiated efforts to provide a map' that will document the participants and their interactions. The roadmap for this project includes establishing a linked open data end-point that will allow people to query the database using (Geo)SPARQL. In concert with activities under the EU-PolarNet project and other research initiatives (e.g. the Polar Data Catalogue, NSF-funded BCube Informatics Project), these efforts are establishing a clear picture of the data and computing services available within the community.

6.2.10 Natural Resources Canada, Public Safety Geoscience Program

The Public Safety Geoscience Program develops a huge amount of geospatial data and place-based knowledge about the Arctic. This includes, but is not limited to: seismic hazard model, space weather forecasts, research into historic tsunamis and floods, marine sample data, videos of submarine features taken by remotely operated vehicles, photos taken by UAVs, marine seismic data, journal articles about regional and local geohazards, slope stability research on the offshore continental slope, marine research station data in the EDS, and maps documenting the location and extent of geohazards and other geological features. Much of this is not included in the CGDI and could be included in both the CGDI and the Arctic Spatial Data Infrastructure.

6.2.11 University Carleton, Geomatics and Cartographic Research Centre

The Geomatics and Cartographic Research Centre (GCRC) is an official Research Centre in the Department of Geography and Environmental Studies, Carleton University, Ottawa, Canada. GCRC offers a number of valuable resources called Atlases, among them:

- SIKU (Inuktitut for sea ice) Atlas developed in partnership with Inuit communities in Nunavut
- Kitikmeot Place Name Atlas
- Pan-Inuit Trails Atlas

6.2.12 Yukon Government

GeoYukon allows users to search, display and download Yukon government's most authoritative and recent spatial data. GeoYukon allows to view, query and download the spatial data and its related metadata.

Chapter 7

Standards

Standards are one of the key pillars of any SDI. They are key for the quality and development of interoperable geographic information and geospatial software during the entire life cycle of any data set. Standards define how data is created, archived, used, and discovered at and exchanged between components. They address different levels of interoperability such as syntax, semantics, services, profiles, or cultural and organizational interoperability. There are excellent publications available discussing the value of standards and role of standards in geospatial information management (OGC/ISO TC211/IHO, 2014) or the usage of standards in SDIs (United Nations, 2013). We therefore concentrate here on experiences made by the SDI developers and users community and refer to external literature for further details on the various standards. A good starting point to learn more about important standards is the Website of the Open Geospatial Consortium.

An approach often used by cookbooks is to classify standards in the context of SDI following three categories as introduced by GPC Group:

- **Data Content Standards** For understanding the contents of different data themes by providing a data model of spatial features, attributes, relationships, and a data dictionary.
- **Data Management Standards** For handling spatial data involving actions such as discovery of data through metadata, spatial referencing of data, collection of data from the field, submission of data by contractors to stakeholders, and tiling of image-based maps.
- **Data Portrayal Standards** For visual portrayal of spatial data using cartographic feature symbology.

This approach is often used by the various cookbooks that exist for the development and operation of an SDI (New Zealand Geospatial Office, 2011; United Nations, 2013). Here, we follow a different approach and discuss standards depending on their functionality domain, i.e. data format & access standards, metadata and catalogs, geodata integration, and orthogonal standards.

7.1 Data Format & Access Standards

> Serving data online by leveraging the latest advancements is critical to help polar scientists better conduct research. Post all data center holdings, especially the polar gridded/raster data, via web services, such as OGC web services. *Report on Workshop on Cyberinfrastructure for Polar Science, 2013* (Pundsack et al., 2013)

The polar community has reported that research and analysis of catalogues such as the Global Change Master Directory has shown that polar data community has adopted OGC standards. For example, the Atlas of the Cryosphere hosted at NSIDC, the Arctic Sensor Web of the Arctic Institute of North America, and the Polar Data Catalogue use OGC standards to make data and maps available for inclusion in external sites and applications. Additionally, the ESA thematic exploitation platforms (including Polar TEP) have been instructed to use OGC standards when available and develop best practices for implementation of the standards. The use of OGC standards will include resource catalogues, processing service execution, processing service packaging, and processing containers. It is anticipated that the resulting TEP best practices definitions will be contributed to OGC in the future. The major challenge in developing increased usage of the OGC approach will be in community building, adequate support (e.g. cookbooks, easily deployed stacks), and a clear value proposition. The focus should be on mature OGC standards that are core OGC services: WMS, WMTS, WCS, WFS, WPS, SOS and CSW.

On the encodings and information model side, the SWE standards SensorML and O&M ISO 19156 with WaterML for the exchange of hydrological time-series as well as GeoSciML should be used if more specialized data such as observations and measurements needs to be exchanged. The pilot may provide an excellent laboratory environment to experiment with new, less mature standards in conjunction with established technology.

Further on, the following standards shall be further explored as part of the pilot:

- KML: KML is a file format used to display geographic data in an Earth browser such as Google Earth.
- GeoPackage: GeoPackage is an open, standards-based, platform-independent, portable, self-describing, compact format for transferring geospatial information. Since a GeoPackage is a database, it supports direct use, meaning that its data can be accessed and updated in a "native" storage format without intermediate format translations. GeoPackages are interoperable across all enterprise and personal computing environments, and are particularly useful on mobile devices like cell phones and tablets in communications environments with limited connectivity and bandwidth.
- AIXM: Aeronautical Information Exchange Model (AIXM)
- S-57: IHO Special Publication 57 (S-57) includes a description of the data format, product specification for the production of ENC data, and an updating profile

7.2 Metadata and Catalogs

Many catalogs and registries make use of OGC Services and their corresponding ISO TC211 documents. For example, the International Hydrographic Organization (IHO) S-100 provides the data framework for the development of the next generation of ENC products, as well as other related digital products required by the hydrographic, maritime and GIS communities. S-100 is based on the ISO 19100 series of geographic standards. This means that S-100 based data is compatible with data created according to the relevant ISO standards. An

S-100 online registry based on the ISO 19135 standard (Procedures for Registration of Geographical Information Items), has been established for the registration, management and maintenance of the various dictionaries of items recognized under the S-100 framework. The registry contains the following principal subordinate registers:

- Feature Concept Dictionary (FCD) register
- Portrayal register
- Metadata register
- Product Specifications register
- Data Producer Code register

It is recommended that metadata follow the ISO 19115 (Geographic Information - Metadata) and corresponding ISO 19139 (Geographic Information - Metadata XML schema implementation), or their respective profiles, CSDGM (FGDC Content Standard for Geospatial Metadata), the Dublin Core, or INSPIRE guidelines and implementation rules. In addition, the emerging DCAT standard may be analyzed in more detail for its applicability in SDIs.

One issue that has been reported is today's focus of spatial data infrastructure metadata standards, which are suitable for business-to-business integration, but not suitable for consumption by ordinary citizens (e.g. elementary school students). There is a need to develop standards which make maps and spatial data suitable for re-use by citizens of limited experience and resources. Communities such as e.g. the W3C Maps for the HTML community have the objective to develop the concepts, software and community associated to the needs of developing a standard for maps suitable for adoption by browsers, and thereby for citizens who produce and consume HTML. As such, an SDI for Arctic would serve as an excellent community-driven initiative which could help stimulate development of the standards and software of Map Markup Language and the <web-map>/<MAP> element; an idea that could be experimented with as part of this pilot.

7.3 Geodata Integration

Combining multiple sources of geospatial information - a necessary key step in the geospatial knowledge generation cycle or geospatial data integration on-demand is still a challenge if it comes to high volumes of data or extremely high update frequencies. A solution can only be achieved through the conversion of traditional data archives into standardized data architectures that support parallel processing in distributed and/or high performance computing environments as well as complex stream processing. A common framework is required that will link very large multi-resolution and multi-domain datasets together and to enable the next generation of analytic processes to be applied. A solution must be capable of handling multiple data streams rather than being explicitly linked to a sensor or data type.

Success has been achieved using a framework called a discrete global grid system (DGGS). A DGGS is a form of Earth reference that, unlike its established counterpart the coordinate reference system that represents the Earth as a continual lattice of points, represents the Earth with a tessellation of nested cells [6]. Generally, a DGGS will exhaustively partition the globe in closely packed hierarchical tessellations, each cell representing a homogenous value, with a unique identifier or indexing that allows for linear ordering, parent-child operations, and nearest neighbour algebraic operations. Further experiments with DGGS and service support should be part of this pilot to gain new insights into large data volumes processing and integration.

7.4 Orthogonal Standards

There are a number of orthogonal standards that may play a role in SDI setup and operation and should be explored further as part of this pilot, for example ISO 27001 (Information Security) and ISO 20000 (IT Service Management).

Chapter 8

Applications

The feedback on current applications used in the Arctic, type of services that should be available as part of an SDI for Arctic, or type of applications shall be developed that make use of Arctic SDI or become integral part of it was surprisingly low. It was partly discussed in section 6.1 and is further elaborated here. Nevertheless, the lack of more concrete examples is a hint on possible gaps between data providers, SDI component operators, and SDI consumers that require further investigation. It is exactly the goal of this pilot to demonstrate the value of Arctic SDI to stakeholders, and therefore overcoming the catch-22 situation that potential stakeholders are not aware of the capabilities of the Arctic SDI and therefore not using it; and the provider side not being able to better adapt to users' needs, as they are not formulated and expressed.

The list of current applications includes on the abstract level items such as:

- geospatial portal development
- incident and event management and decision support applications
- mobile platform integration and cross platform mobile app development
- data warehouse management
- Internet mapping
- metadata management
- gateway creation
- comprehensive data connectivity

More detailed applications include

- water license management and water quality and quantity flow analysis, visualization and modeling
- wildlife management including range planning and habitat monitoring
- landscape change detection
- Permafrost monitoring
- Infrastructure monitoring and planning
- Environmental planning (air quality)
- Land use monitoring and planning
- Decision Support Systems
- Species at risk identification and monitoring
- Climate change research

The list key applications that shall be made available as part of an Arctic SDI include:

- Tool to monitor climate change
- Tool to measure, monitor and track permafrost conditions
- DEM/Landsat viewer/analysis app
- Dashboard for disaster monitoring/response

- Metadata harvester/broker
- SDI Portal for non-technical users
- SDI portal for scientists
- Story map series on Polar Region threats
- Glacier tracker
- Oil spill tracker/forecaster

The wide gamut of potential services shall support the following aspects:

- Ease of discovery and use
- Local community emphasis
- Possibility of using in non or low-bandwidth environments
- Use of geospatial standards
- Service level agreements to address business and contractual agreements
- The Arctic SDI should strive to increase the use and accessibility of scientific data stored in the USG Distributed Active Archive Centers (DAACs). We propose the development of a discovery mechanism for finding this data that leverages semantic mediation and similar search mechanisms.
- Scientific applications are important for understanding and developing the Arctic, but some additional SDI applications could include more general applications like tourism and eco-tourism; search and rescue; economic and resource development; and social and demographic applications.

A very interesting suggestion that was being made addresses an ontology and linked open data application. Though this is rather an experimental field, it should allow better understanding of future trends for data sharing and collaborative research. The scenario is described in full detail in Appendix E.

For demonstration purposes, there are some examples that could serve as a starting point for further discussion. It shall be discussed in phase two what type of apps could be powered by an SDI for the Arctic.

- The Columbia glacier app allows users to check the retreat of the glacier, compare 2009 and 2013 images, generate elevation profiles, and produce stats on glacier retreat during this period.

- Esri produced a Story Map to highlight the use of elevation data and its benefits in the Arctic region.
- The Polar Knowledge iPhone app allows to get information from across the spectrum of polar knowledge from Canada and the circumpolar world. It allows to connect with Arctic and Antarctic experts, to find out who is doing polar research and where, to locate polar science infrastructure and monitoring sites; and to learn about northern communities.
- USG agencies have deployed services that could be leveraged to serve the Arctic SDI Community, such as the NOAA Bathymetric Survey Client.
- There are a number of examples developed by the Russian Arctic and Antarctic Research Institute for the use of GIS and the deployment of Web services, such as the national protected areas map service (including the Arctic). The Institute also publishes weekly Arctic ice analysis maps.
- The World Wildlife Fund (WWF) Norway developed a map service that combines data about the environment and human activity in the Arctic.

Chapter 9

Use Cases & Scenarios

The goal of this pilot is to demonstrate the value of an SDI for the Arctic. It has been agreed that this could best be done by the implementation and description of a number of use cases and scenarios that make use of a number of data sets discovered and served by the Arctic SDI and visualized by Arctic SDI client components. The following overview is intended as a starting point for the Arctic Spatial Data Pilot Phase-2.

9.1 Linking Indigenous and Scientific Knowledge

One of the open challenges is the effective integration of indigenous and scientific (observation driven) knowledge. The pilot should, working in partnership with indigenous peoples, communities and their representative organizations, establish effective methods for linking indigenous knowledge with scientific and operational knowledge. This includes projects such as the SIKU

Atlas (Inuktitut for sea ice) developed in partnership with Inuit communities in Nunavut, the Kitikmeot Place Name Atlas, and the Pan-Inuit Trails Atlas. These are just a few examples of projects where Indigenous communities drove the development of the technology involved to ensure that their needs were met. The Nunaliit Atlas Development Framework established by GCRC is another example of this inclusive development approach (Hayes, P. L. Pulsifer, and Fiset, 2014). The full scenario may involve Inuit and western scientists and decision makers using knowledge and data provided in an integrated fashion from both sources (indigenous and scientific).

A key strategic priority of the Geographical Names Board of Canada is the support of indigenous naming, as well as building indigenous capacity for geographical naming in Canada's North. Place names are extremely important to Indigenous Peoples as a reflection of their culture and heritage, and also as a tool for language preservation. GNBC naming authorities may indicate the specific language of place names in the national names database. The language may be defined from a standardized ISO list of 74 languages used by Canada's Indigenous Peoples, as well as English and French.

9.2 Geohazards & Weather

Offshore geohazard research is another thematic domain that would serve well for pilot phase-2 demonstration scenarios. Focus is on support for safe development of infrastructure (Beaufort Sea and Baffin Bay) by improving the understanding of offshore geology and slope stability issues due to permafrost degradation. Marine research cruises collect seismic data, core samples, video, multibeam bathymetry, and water samples. These data are interpreted, and the findings published in journal articles. Each of the constituent pieces could be part of the Arctic Spatial Data Infrastructure.

Geohazard risk assessments: Combining knowledge about geohazards (in this case earthquakes and/or floods) and the built environment and demographics, a risk assessment examines the potential losses and consequences that could be generated, and allows for cost benefit analyses of mitigation. This could be of particular interest in an area of frequent hazards or potential development. Research is currently taking place in southern Canada, but could be extended to areas of interest in the North.

The Space Weather component of the Public Safety Geoscience Program has a lot of geospatial data and knowledge, and this could be included in the Arctic Spatial Data Infrastructure. Collaborators in this research also involve domestic utilities and may provide use cases of data integration use SDI between publicly collected data and privately held data, depending on agreement by parties external to the program.

9.3 Marine Use Cases

Although the primary use of hydrographic data is the safety of navigation, it can serve many other purposes when included in an SDI:

- Habitat mapping and heritage assessment
- Conservation assessment and designation
- Site selection (e.g. renewable energy and oil and gas extraction)
- Route optimisation
- Vessel location and disposal monitoring
- Homeland security and defence
- Aggregates extraction
- Fisheries regulation
- Coastal protection and shoreline management
- Licensing and consent evaluation
- Emergency planning and management
- Survey planning and execution

Any emergency scenario will likely include a marine aspect, as this article emphasizes. The receding ice in the Arctic has been leading to increased maritime traffic and resource exploration in areas that are not well surveyed and remote. In addition, many areas of the Arctic can only be reached by air or water. This increases the risk for incidents such as vessel groundings, oil spills, danger to wildlife, or even human catastrophe (disease). Each of these scenarios will require marine data in order to make informed decisions.

A non-hazard scenario might address the combination of available space based imagery and Automatic Identification System (AIS) messages that would allow tracking of sea ice and commercial vessels operating within the Arctic SDI area of interest. Analysis of imagery data to track ice would allow projection of future ice movement into shipping lanes indicated by regular AIS routes. In this use case, an agency associated with an Arctic SDI member state would request monitoring of shipping lanes in their EEZ for ice intrusion and potential intersection with specified vessels through the generation of an Area of Interest (AOI). To do this, the agency would request access to available government licensed or commercially available imagery data for a specified period. This data would be made available to the Arctic ADI either through a service interface to an Arctic SDI provided data server. Imagery analytics would determine to number and positions of sea ice objects in the defined AOI. Subsequent imagery collections would be processed to determine the movement, and fragmentation of ice objects, as well as the entry of new ice objects into the AOI. Projections of travel path for ice objects may be able to be predicted. At the same time, ship track data can be requested from ExactAIS WFS services. A boundary condition for nearest approach to sea ice, or projected ice oath, could be established. Warning to specific vessels could then be generated and transmitted.

Another use of AIS data is for intrusion detection and identification into environmentally sensitive areas. An agency associated with the Arctic SDI would request monitoring of vessel traffic intruding on an area defined as environmentally sensitive. The agency would define the area of interest for that region. The Arctic SDI would request AID point reports that intersect the define area of interest. The ship identification from the AIS MMSI information and location of the intrusion would be provided as a reporting service to the requesting agency. It would also be possible to maintain a database of all intrusions. This would allow a generation of historic patterns of intrusion for required for action. As an extension of this use case, available commercial or government licensed satellite imagery could be conducted to confirm the identity of the MSSI information.

9.4 Terrestrial Use Cases

It is recommended that a small number of cases crosses or overlaps multiple jurisdictions. Here, wildlife applications (particularly related to Caribou or other important terrestrial mammals) such as habitat management may be of particular interest (an application focusing on aspects of the Boreal Caribou Recover Strategy for example). Applications related to emergency response and multi-agency response could also demonstrate value. Finally, trans-boundary management issues such as water sheds and cumulative impact of human induced and naturally occurring ecosystem changes provide good cases for demonstration. Those use cases could be extended with additional elements such as the monitoring of the status of feeding areas for migratory birds, or the changes on Arctic biodiversity including the northward movement of more southern species, shrubbing and greening of the land, etc.

9.5 Climate Change

Climate change is one of the most prominent scientific fields of research in the arctic. A typical use case would include aspects such as the monitoring of sea level rise, including evaluation of areas that are suffering the worst impacts and the estimation of damage to infrastructure; the tracking of glacier movement including the estimation of potential impact to shipping lanes or coastal infrastructure and the projection of future conditions; or the monitoring of the status and condition of the permafrost layer, including the evaluation of impact to existing infrastructure with projections of future conditions.

9.6 Other Scenario Aspects

There are a number of aspects that are independent on the concrete scenario. Instead, they are applicable to almost all scenarios. One very important aspect in this context is the low to no Internet bandwidth in some areas. This aspect that was mentioned several times should be addressed in at least one pilot implementation scenario.

Another aspect addresses typical issues caused by cross-boundary events, e.g. a downed aircraft near an international border such as between Alaska and the Yukon. This requires bringing together a wide variety of disparate data and cross border interoperability.

Though the focus on the first phase of the Arctic Spatial Data Pilot has been on the North American Arctic, pan-arctic applications could be addressed in the second phase. One useful application that could be enabled through an SDI for the Arctic are pan-Arctic ice charts. Currently, if ice charts are used in scientific studies, they need to be gathered from different agencies depending on the study area. The various ice charts cover different areas and can have widely different data formats, file formats, and accuracies. Harmonizing the various agencies that produce ice charts and delivering them through one tool would be a very welcome development.

Consideration should be given on how the United Nations Sustainable Development Goals apply in the Arctic region and what role the Arctic SDI can play in meeting the Targets of the SDGs. This could be part of an OGC Pilot even in the suggested Pilot scenarios.

FIGURE 9.1: United Nations Sustainable Development Goals

9.7 Pilot Development

This pilot is not trying to develop a different, new operational professional-grade Arctic SDI, instead it supports the *Arctic SDI* by demonstrating its value to stakeholders. Nevertheless, the following parts shall be considered during the development of the second phase of the pilot, even though they usually apply to the development of a professional-grade SDI. These parts would be concurrent with putting up an open data site portal with the idea that that portal will evolve based on the following:

Part 1 Assessment and Planning

1. Project Initiation- host key stakeholders to discuss key issues that have direct bearing on the form and structure of an SDI, select an Advisory and Technical Committee to support all phases of the project.
2. Conduct Stakeholder Survey-survey organization interested in Arctic issues related to an SDI to understand their needs and requirements.
3. Prepare Inventory and Assessment of Existing Arctic Systems, Portals, Applications, Data Sets and Databases. This involves assessing key existing GIS systems and imagery in the Arctic region. Inventory should cover formats, datums, metadata, standards used, etc.
4. Finalize requirements related to an SDI with an Arctic focus with Advisory and Technical Committee agreement.
5. Prepare an SDI Portal and Implementation Strategy. Implementation Strategy should include 5-year sustainability plan covering hosting and data sharing agreements with a security plan.
6. Implement the next version of an SDI Open Data Portal based on the Implementation Strategy. This could have hosted datasets and networked Arctic community resources and sample apps for use.

Part 2 Program Implementation

1. Outreach and adoption
2. Continue to populate data sets via network of networks and hosted data
3. Improve Arctic SDI based on feedback and technology evolution
4. Online training for end users and providers
5. OGC Pilot to improve interoperability arrangements between providers and users testing that OGC services are discoverable in heterogeneous global community of providers.

Appendix A

Appendix: Stakeholders: Individual points of contact

List of individuals that can help establishing contact to important organizations and local communities.

- Peter L. Pulsifer, Chair (U.S. Representative), Arctic Data Committee; Co-Chair U.S. IARPC Arctic Data Coordination Team; and National Snow and Ice Data Center, CIRES, University of Colorado, email: peter.pulsifer @colorado.edu
- David Arthurs, Managing Director, Polar View Earth Observation, email: david.arthurs@polarview.org
- Allison Gaylord, Alaska Data Integration Working Group, Nuna Technologies
- Anton Van de Putte, Chair, Standing Committee on Antarctic Data Management
- Beatrix Schlarb-Ridley, Work Package Lead, Infrastructures, Facilities, and Data, EU-PolarNet
- Craig Tweedie, Principal Investigator, Arctic Research Mapping Application, Arctic Observing Viewer, Barrow Area Information Database
- D.R. Fraser Taylor, Director, Geomatics and Cartographic Research Centre, Carleton University
- Halldór Jóhannsson, Executive Director, Arctic Portal
- Jennifer Sokol, Manager, Partnerships and Engagement, Polar Knowledge Canada

- Joe Casas, NASA, Co-Chair Interagency Arctic Research Policy Committee (IARPC) Arctic Data Coordination Team
- John Falkingham, International Ice Charting Working Group
- Jonathan Pundsack, Principal Investigator, National Science Foundation Antarctic and Arctic Data Consortium (A2DC)
- Julie Friddell, Associate Director, Canadian Cryospheric Information Network
- Lawrence Hislop, Director, Climate and Cryosphere (CliC)
- Maribeth Murray, Principal Investigator, Canadian Consortium for Arctic Data Interoperability
- Masaki Kanao, Associate Professor, Polar Data Center, National Institute of Polar Research (NIPR)
- Ola Gråbak, Applications Engineer, European Space Agency
- Phillippa Bricher, Data Officer, Southern Ocean Observing System
- Tom Barry, Executive Secretary, Conservation of Arctic Flora and Fauna (CAFF), Akureyri, Iceland

Appendix B

Appendix: Governance Lessons Learned

Institutional dimension
Success of SDI initiates are based on the trust & goodwill established in a community over years and based on personal relationships
Barriers between the geospatial and IT communities hamper efforts to build SDI
SDI governance arrangements and existing government (including Whole of Government) and domain governance arrangements need to be harmonised
Key agencies typically have multiple roles in SDI including lead agency, custodial, secretariat of governance body, coordinator, and operator of the infrastructure. These roles must be clearly understood and separated.
Governance operates through representational processes and the effectiveness of the process is dependent on the quality of the representation
For effective governance, agencies acting as community representative must subordinate the interests of their own organisation to those of constituency being represented
Current SDI approaches and governance models reflect a hierarchical structured government world view, which does not adequately accommodate network and market-oriented realities of the geospatial industry
SDI implementation requires leaders in individual agencies to champion initiative
With institutional changes (organisations, people and roles), business drivers, motivation, priorities and power balance changes. This results in a loss of momentum and a need to re-build partnerships and trust
Governance mechanism should encourage participation of and give voice to smaller agencies
Government to government business is hampered by weak contract arrangements and weak enforcement mechanisms with contract performance based on trust

FIGURE B.1: Governance lessons learned, from (Paul Box and Rajabifard, 2009), part 1

Business dimensions
There is a need to understand the business drivers for SDI
There is a need to ensure alignment of business outcomes of individual agencies with collective interests
There is a need to convincingly answer the question "What's in it for me?" for potential participants in an SDI initiative.
Selling SDI to participants using the benefit of increased efficiency alone is problematic, as this may be perceived as a threat – i.e. it represents loss of budget/staff/status/power
Justification for obtaining agency funding to participate in SDI is problematic as the benefits are realized by users outside of the organization (which maybe beyond the business goals of the funding organization).
Reticence of agencies to participate in SDI can be overcome by demonstrated positive benefits and results
Data
Complexity, lack of standards and guidance related to data licencing is cited as a key barrier to the sharing and publication of geospatial data
Liability concerns related to incorrect data or misuse of published data are cited as reason for non-publication of data
There is a need to adopt a transactional view towards data access, pricing and licencing. The current approach of accessing data sets based around physical data storage reality is out-dated
There is an accountability gap with regard to data custodianship. Custodial responsibility is mapped to an organizations. However in reality, geospatial data is typically managed in business units and with corporate IT governance focuses on corporate data.
Services and technology
Technical limitation (bandwidth, speed) are cited as barriers for low uptake of geospatial web services by GIS users
Entrenched business practices are cited as barriers for low uptake of geospatial web services by GIS users
Service quality limitation (e.g. availability) are cited as barriers for low uptake of geospatial web services by GIS users
to enable SDI stakeholders them to participate in the SDI adequate technical support and mutual learning opportunities are required

FIGURE B.2: Governance lessons learned, from (Paul Box and Rajabifard, 2009), part 2

Appendix C

Appendix: Polar Data Portals

The following summarizes a selection of data portals and initiatives that are relevant to polar information.

The Sustaining Arctic Observing Networks (SAON): The Sustaining Arctic Observing Networks (SAON) process was initiated by the Arctic Council (AC) and has been underway since early 2007. Its purpose is to support and strengthen the development of multinational engagement for sustained and coordinated pan-Arctic observing and data sharing systems that serve societal needs, particularly related to environmental, social, economic and cultural issues. SAON promotes the vision of well-defined observing networks that enable users to have access to free, open and high quality data that will realize pan-Arctic and global value-added services and provide societal benefits. Its goal is to enhance Arctic-wide observing activities by facilitating partnerships and synergies among existing observing and data networks (building blocks), and promoting sharing and synthesis of data and information. SAON also is committed to facilitating the inclusion of Arctic indigenous people in observing activities, in particular by promoting community-based monitoring (CBM) efforts.

Global Earth Observation System of Systems (GEOSS): The Group on Earth Observations (GEO) is an intergovernmental organization working to improve the availability, access to and use of Earth observations by building a Global Earth Observation System of Systems (GEOSS), which provides decision-support tools to a wide variety of users. As with the Internet, GEOSS will be a global and flexible network of content providers allowing decision makers to access an extraordinary range of information at their desk. The GEOSS Portal is the main entry point to Earth Observation data from all over the world. The GEOSS information services for cold regions coordinates joint, global efforts to provide Earth observations and information services to decision-makers over a vast cold regions area, including the North Pole, South Pole, Himalaya-Third Pole and mountain cold regions.

The Arctic Portal: The Arctic Portal is a comprehensive gateway to Arctic information and data on the Internet, increasing information sharing and co-operation among Arctic stakeholders and granting exposure to Arctic related information and data. The Arctic Portal is operated in consultation and co-operation with members of the Arctic Council and its Working Groups, Permanent Participants, Observers and other Stakeholders. The Arctic Portal is a network of information and data sharing and serves as host to many web sites in a circumpolar context, supporting co-operation and outreach in science, education, and policy making.

The Arctic Portal is managed as non-profitable organization, located in Akureyri, Iceland, under an international board of directors.

The WMO Global Cryosphere Watch (GCW): GCW is an international mechanism for supporting all key cryospheric in-situ and remote sensing observations. To meet the needs of WMO members and partners in delivering services to users, the media, public, decision and policy makers, GCW provides authoritative, clear, and useable data, information, and analyses on the past, current and future state of the cryosphere.

Polar Data Catalogue: The PDC is a database of metadata and data files that describes, indexes, and provides access to diverse data sets generated by Arctic and Antarctic researchers. Geographic focus is on Canada, but the PDC holds international collections, too, such as hundreds of metadata records of CAFF's Circumpolar Biodiversity monitoring programme (CBMP). The records follow ISO 19115 and Federal Geographic Data Committee (FGDC) standard formats to provide metadata exchange with other data centres. Interoperability via OGC WMS, OGC WFS, OAI-PMH, and CSW (GeoNetwork) are in place for sharing metadata and data. The metadata records cover a wide range of disciplines from natural sciences and policy, to health and social sciences. Datasets are available for free public download, with new files being added on a regular basis as we work with researchers to prepare and submit their datasets.

Polar Knowledge Canada: Polar Knowledge Canada (POLAR) is on the cutting edge of Arctic issues and strengthens Canada's position internationally as a leader in polar science and technology. POLAR also promotes the development and distribution of knowledge of other circumpolar regions, including

Antarctica. It will provide a world-class hub for science and technology research in Cambridge Bay, Nunavut called the Canadian High Arctic Research Station. As part of Canada's Northern Strategy, POLAR improves economic opportunities, environmental stewardship and quality of life for Northerners and other Canadians.

Arctic Data Centre: Arctic Data Centre is a WMO Information System Data Collection and Production Centre building on the IPY legacy.

Arctic Data Explorer: The Arctic Data Explorer (ADE) is a cross-domain data discovery tool for searching distributed repositories. The current search space includes the holdings of the ACADIS Gateway, NCAR's Earth Observing Lab, National Snow and Ice Data Center, Norwegian Meteorological Institute, the Polar Data Catalogue, and the US National Oceanographic Data Center, and others. The ADE features an ISO-based metadata store, an available OpenSearch (ESIP-compliant) endpoint for automated searching, and metadata brokering technologies that allow for ingest of feeds in many formats.

Arctic Observing Viewer (AOV): AOV is a web mapping application for Arctic Observing data collection sites. This prototype is now available for visualization, synthesis, strategic assessment, and decision support for U.S. SEARCH/AON and other initiatives. It provides the who, what, where, and when of data collection activities (sites with repeat measurements such as towers, boreholes, weather stations, etc).

Arctic Research Mapping Application (ARMAP): ARMAP is a suite of online, interactive maps and web services that support Arctic science. The application displays details and field locations for over 2300 research projects funded by the US NSF and seventeen other agencies. A variety of web data services are also available for use by other organizations.

The International Arctic System for Observing the Atmospheres (IASOA) mission is to advance coordinated research objectives from independent pan-Arctic atmospheric observatories through (1) strategically developing comprehensive observational capacity, (2) facilitating data access and usability through a single gateway, and (3) mobilizing contributions to synergistic science and socially-relevant services derived from IASOA assets and expertise. The IASOA data access portal provides (through ISO-19115-2 metadata) discovery and access-level information for 700 atmospheric datasets from the ten Arctic Observatories of IASOA.

Sustaining Arctic Observing Networks (SAON): SAON has a search facility for Arctic observational data and metadata harvested from a series of data management institutions.

AbiskoGIS: AbiskoGIS is a research station-based initiative from the Abisko Scientific Research Station, in sub-Artic Sweden, that contains a mix of project data and metadata of monitoring activities.

Arctic Data archive System (ADS): The purpose of the Arctic Data archive System (ADS) is to archive and distribute multiple observational (atmosphere, ocean, terrestrial, and ecology) and satellite and model simulation datasets, and promote utilization of these datasets. ADS is the central repository of archived data on Arctic research in Japan

Natural Environment Research Council Arctic Office (UK): The aim of the Office and its website is to coordinate UK scientific research in the Arctic. The Office does provide information in its own right through a web map service (map.arctic.ac.uk).

The Norwegian Polar Data Centre and Norwegian Polar Institute Maps and Services are infrastructure nodes in development at the Norwegian Polar Institute, primarily focused on managing and distributing data from the institute itself. The data centre holds scientific, environmental and topographic data from the Norwegian Arctic, and distributes the data through open web services. The data centre also holds the responsibility as a National Antarctic Data Centre (NADC) for Norway, and Antarctic metadata are harvested by the Antarctic Master Directory. The metadata services are being connected to other networks as well, including the Norwegian IPY data catalogue (DOKIPY).

Sea Ice Prediction Network (SIPN): Decline in the extent and thickness of Arctic sea ice is an active area of scientific effort and one with significant implications for ecosystems and communities in the Arctic and globally. Forecasting for seasonal timescales (i.e., the summer and into fall) is of particular interest to many stakeholders since many activities that take place in the Arctic are planned over the summer months, and many species are sensitive to the behavior of summer sea ice. However, seasonal forecasting is particularly challenging due to the variable nature of weather and ocean behavior over that timescale as well as current limits to data and modeling capabilities. SIPN builds and expands on the Sea Ice Outlook project. The Sea Ice Prediction Network (SIPN), launched in the fall of 2013, will develop a collaborative network of scientists and stakeholders to advance research on sea ice prediction and communicate sea ice knowledge and tools.

Svalbard Integrated Arctic Observing System (SIOS): SIOS is an international infrastructure project. There are 28 partners from Europe and Asia involved. The essential objective is to establish better-coordinated services for the international research community with respect to access, data and knowledge management, logistics and training.

ECDS - Environment Climate Data Sweden: A data center that is part of the Swedish national infrastructure where research (meta)data can be stored and explored. This service has the potential to be a hub for exploring Swedish Arctic research data in a wider Arctic network.

The Norwegian Institute for Air Research (NILU) organizes atmospheric contaminants data for AMAP. The data are accessible through their EBAS database.

The WMO Information system (WIS) is the single coordinated global infrastructure responsible for WMO telecommunications and data management functions. It is the pillar of the WMO strategy for managing and moving weather, climate and water information in the 21st century.

GAWSIS is related to, but more specific than, the WMO listing above. It is an over-arching, coherent metadata system for the six world data centers that support the WMO's Global Atmospheric Watch (GAW) program: WDCGG (Gases), WOUDC (Ozone/UV); WDCPC (Precipitation and Chemistry); WWRDC (Radiation); WDCA (Aerosols/AOD); and WDC-RSAT (Remote Sensing).

Arctic Biodiversity Data Service (ABDS) by CAFF (Conservation of Arctic Flora and Fauna) and Seabird Data portal. Arctic Biodiversity Data Service (ABDS) by CAFF (Conservation of Arctic Flora and Fauna) and its Seabird Data portal. ABDS allows for the combination of geo-referenced data at various spatial, temporal, and taxonomic scales. It contains a range of data focusing on biodiversity,from Arctic assessment and monitoring programmes. It contains a range of data types focusing on various aspects of biodiversity, including vector and satellite imagery. The ABDS also contains an Arctic Integrated Publishing Toolkit (IPT) via which it links with the Global Biodiversity Information Facility (GBIF) GBIF and serves as the Arctic node for UNESCOs Ocean Biogeographic Information System (OBIS). ABDS also partners with the Polar Data Catalogue to allow discovery of access to data

Exchange for Local Observations and Knowledge of the Arctic (ELOKA): ELOKA provides data management and user support to facilitate the collection, preservation, exchange, and use of local observations and knowledge of the Arctic. ELOKA partners with Indigenous communities around the circumpolar Arctic to establish ethical and culturally appropriate mechanism (technical, policy, partnerships) for sharing Indigenous knowledge and information in digital and other forms.

National Snow and Ice Data Center (NSIDC): The National Snow and Ice Data Center (NSIDC) supports research into our worlds frozen realms: the snow, ice, glaciers, frozen ground, and climate interactions that make up Earths cryosphere. NSIDC manages and distributes scientific data, creates tools for data access, supports data users, performs scientific research, and educates the public about the cryosphere.

World Glacier Monitoring Service Meta Data Browser: The WGMS collects standardized observations on changes in mass, volume, area and length of glaciers with time (glacier fluctuations), as well as statistical information on the distribution of perennial surface ice in space (glacier inventories). Such glacier fluctuation and inventory data are high priority key variables in climate system monitoring.

The ArcticROOS is a GOOS Regional Alliance for the Arctic: The secretariat is located at the Nansen Environmental and Remote Sensing Center in Norway. It has been established by a group of 14 member institutions from nine European countries working actively with ocean observation and modelling systems for the Arctic Ocean and adjacent seas. The aim of the ArcticROOS is to promote, develop and maintain operational monitoring and forecasting of ocean circulation, water masses, ocean surface conditions, sea ice and biological/chemical ocean constituents in the Arctic Ocean. The ArcticROOS website contains metadata and results of long-term observations from partners. The main focus is on cryospheric and oceanographic data.

The International Council for the Exploration of the Sea (ICES) is an intergovernmental organization whose main objective is to increase the scientific knowledge of the marine environment and its living resources and to use this knowledge to provide unbiased, non-political advice to competent authorities. The ICES Marine Data Center organises marine data on, among other things, contaminants, and the biological effects of these, biological communities, oceanography, and fisheries. ICES is the data center for these data for AMAP.

International Bathymetric Chart of the Arctic Ocean (IBCAO): The goal of IBCAO is to develop a digital data base that contains all available bathymetric data north of 64 degree North, for use by mapmakers, researchers, institutions, and others whose work requires a detailed and accurate knowledge of the depth and the shape of the Arctic seabed.

Norwegian Satellite Earth Observation Database for Marine and Polar Research (NORMAP): The overall goal of NORMAP is to create and maintain a data repository, including metadata of the high latitude and Arctic regions based on earth observation data from polar orbiting satellites to facilitate and stimulate high quality and original multidisciplinary earth system research, application and education in marine, polar and climate sciences.

Oden Mapping Dat: Bathymetric data from multi-beam sounding on icebreaker Oden cruises. The data have been extensively downloaded and used and can be regarded as an example of a local initiative by an individual research group leading to a success story in sharing of research data. The data is one of the sources to the IBCAO.

Arctic Observation Network Social Indicator Project: The Arctic Observation Network Social Indicator Project (AON-SIP) was supported by the National Science Foundation from 2007 to 2011. This website is maintained to offer access to data compiled by the AON-SIP by researchers and policy makers.

Frozen Ground Data Center: The International Permafrost Association (IPA) has developed a strategy for data and information management to meet the requirements of the cold regions science, engineering, and modeling communities. A central component of this strategy is the Global Geocryological Data (GGD) system, an internationally distributed system linking investigators and data centers around the world. The National Snow and Ice Data Center (NSIDC) in collaboration with the International Arctic Research Center (IARC) serves as a central node of the GGD.

Appendix D

Appendix: Polar Community

The following are brief introductions to the organizations that have contributed to this Request for Information response on behalf of the wider **polar data community**.

D.1 Arctic

The **International Arctic Science Committee (IASC)** has a broad mandate to encourage, facilitate and promote cooperation in all aspects of Arctic research in all countries engaged in Arctic research and in all areas of the Arctic region. IASC cuts across all sciences and helps to promote science development, provides scientific advice and policy level documents, aims to maintain freedom and ethical conduct in science, and engages in long-term science visioning and planning. [iasc.info]

The **Sustaining Arctic Observing Networks (SAON)** activities are complementary to IASCs, focusing on the specifics of establishing a long-term Arctic-wide observing activities that provide free, open, and timely access to high-quality data.[www.arcticobserving.org]

The **Arctic Data Committee (ADC)** is a merger of the former IASC Data Standing Committee and SAON Committee on Data and Information Services. The overarching purpose of the ADC is to promote and facilitate international collaboration towards the goal of free, ethically open, sustained, and timely access to Arctic data through useful, usable, and interoperable systems. [arcticdc.org]

The **Arctic Portal** is a comprehensive gateway to Arctic information and data on the Internet, increasing information sharing and co-operation among Arctic stakeholders and granting exposure to Arctic related information and data. The

Arctic Portal is a network of information and data sharing and serves as host to many web sites in a circumpolar context, supporting co-operation and outreach in science, education, and policy making. [arcticportal.org]

D.2 Antarctic

The **Scientific Committee on Antarctic Research (SCAR)** is an inter-disciplinary committee of the International Council for Science (ICSU). SCAR is charged with initiating, developing and coordinating high quality international scientific research in the Antarctic region (including the Southern Ocean), and on the role of the Antarctic region in the Earth system. The scientific business of SCAR is conducted by its Standing Scientific Groups which represent the scientific disciplines active in Antarctic research and report to SCAR. [www.scar.org]

The **Standing Committee on Antarctic Data Management (SC-ADM)** helps facilitate co-operation between scientists and nations with regard to scientific data. It advises on the development of the Antarctic Data Directory System and plays a major role in the International Polar Year data system (IPYDIS). Members of SC-ADM are usually managers of the National Antarctic Data Centres or a relevant national contact. [www.scar.org/data-products/scadm]

The **Southern Ocean Observing System (SOOS)** was launched in August 2011 with the mission to establish a multidisciplinary observing system to deliver the sustained observations of the Southern Ocean that are needed to address key challenges of scientific and societal relevance (e.g., climate change, sea-level rise, impacts of global change on marine ecosystems). The SOOS Data Management Sub-Committee is charged with encouraging data sharing and discovery for essential observations of dynamics and change in the Southern Ocean. [www.soos.aq]

D.3 International Cryosphere

Climate and Cryosphere (CliC) aims to improve understanding of the cryosphere and its interactions with the global climate system, and to enhance the ability to use parts of the cryosphere for detection of climate change. CliC was established as a core project of theWorld Climate Research Programme in 2003. [www.climate-cryosphere.org]

D.4 Canada

Polar Knowledge Canada (POLAR) is on the cutting edge of Arctic issues and strengthens Canada's position internationally as a leader in polar science and technology. POLAR also promotes the development and distribution of knowledge of other circumpolar regions, including Antarctica. It will provide a world-class hub for science and technology research in Cambridge Bay, Nunavut called the Canadian High Arctic Research Station. POLAR improves economic opportunities, environmental stewardship and quality of life for Northerners and other Canadians. [www.canada.ca/en/polar-knowledge.html]

Since 1995, the main objectives of the **Canadian Cryospheric Information Network (CCIN)** have been: to provide a data and information management infrastructure for the Canadian cryospheric research community; to enhance public awareness and access to cryospheric information and related data; and to facilitate the exchange of information between researchers, northern communities, decision makers, and the public. CCIN operates the Polar Data Catalogue, a database of metadata and data that describes, indexes, and provides access to diverse data sets generated by Arctic and Antarctic researchers. The metadata records follow ISO 19115 and Federal Geographic Data Committee (FGDC) standard formats to provide exchange with other data centres. [www.ccin.ca][www.polardata.ca]

The **Canadian Consortium for Arctic Data Interoperability (CCADI)** is composed of a group of Canadas foremost Arctic scholars and Arctic data managers that promote collaboration, nationally and internationally, in the development of an integrated national data management system. CCADI seeks to facilitate information discovery, establish metadata and data sharing standards, enable interoperability among existing data infrastructures, and provide access to the broadest possible audience of users.

The research of the **Geomatics and Cartographic Research Centre of Carleton University** focuses on the application of geographic information processing and management to the analysis of socio-economic issues of interest to society at a variety of scales from the local to the international and the presentation of the results in new, innovative cartographic forms. Cybercartography is a new multimedia, multisensory and interactive online cartography and its main products are cybercartographic atlases using location as a key organizing principle. [gcrc.carleton.ca]

D.5 United States

The **National Snow and Ice Data Center (NSIDC)** supports research into our worlds frozen realms: the snow, ice, glaciers, frozen ground, and climate interactions that make up Earths cryosphere. NSIDC manages and distributes scientific data, creates tools for data access, supports data users, performs scientific research, and educates the public about the cryosphere. [nsidc.org]

The **Interagency Arctic Research Policy Committee (IARPC)** is chartered as a subcommittee under the National Science and Technology Council (NSTC). It consists of principals from 16 agencies, departments, and offices across the Federal government charged with enhancing both the scientific monitoring of, and research on, local, regional, and global environmental issues in the Arctic. [www.iarpccollaborations.org/teams/Arctic-Data]

The **Alaska Data Integration Working Group (ADIwg)** was formed to examine and address the technical barriers to efficiently integrate and share data within and among participating organizations. ADIwg evolved from, and supports the common interests of, the North Slope Science Initiative Oversight Group (NSSI), Alaska Ocean Observing System (AOOS), the Arctic Research Mapping Application (ARMAP), the North Pacific Research Board (NPRB), the Alaska Climate Change Executive Roundtable (ACCER), and their member agencies. [www.adiwg.org]

The **NSF Antarctic and Arctic Data Consortium (a2dc)** is a collaboration of NSF funded research centers and support organizations that provide polar scientists with data and tools to complete their research objectives. From searching historical weather observations to submitting geologic samples, polar researchers utilize the a2dc to search and contribute to the wealth of polar scientific and geospatial data. [www.a2dc.org]

The **Arctic Research Mapping Application** is designed for funding agencies, logistics planners, research investigators, students, and others to explore information about science being conducted across the Arctic. Hundreds of project locations and ship tracks are shown on the interactive web map, with easy access to details on funding agency, funding program, scientific discipline, principal investigator, project title, and much more. [armap.org]

The **Arctic Observing Viewer** is a web mapping application in support of U.S. SEARCH, AON, and other Arctic Observing networks. A collaborative effort, it helps answer the questions such as: How can we know where to go if we don't know where we've been? What resources already exist? Is there overlap? Where are the gaps? [www.arcticobservingviewer.org]

The **Barrow Area Information Database (BAID)** is a resource for learning about the types of data collection activities in the Barrow area on the North Slope of Alaska. The BAID team collaborates with scientists and the local community to compile and share information via online web mapping applications. [barrowmapped.org]

D.6 Asia

The **National Institute of Polar Research (NIPR)** is Japans key institution for scientific research and observation in the polar regions. It maintains monitoring stations in Antarctica and the Arctic, conducts comprehensive polar research based on monitoring programs, and acts as a center for the cultivation of researchers. [www.nipr.ac.jp/English]

D.7 Europe

Approved by the European Commission in November 2014, the goal of **EU PolarNet** is to coordinate polar research in Europe and develop a comprehensive European polar research programme. Its purpose is to provide Europe with the capability to better understand the nature of environmental risks and allow policy-makers and governments better able to design measures to mitigate those risks. [www.eu-polarnet.eu]

The **European Space Agency (ESA)** is Europe's gateway to space. Its mission is to shape the development of Europe's space capability and ensure that investment in space continues to deliver benefits to the citizens of Europe and the world. [www.esa.int]

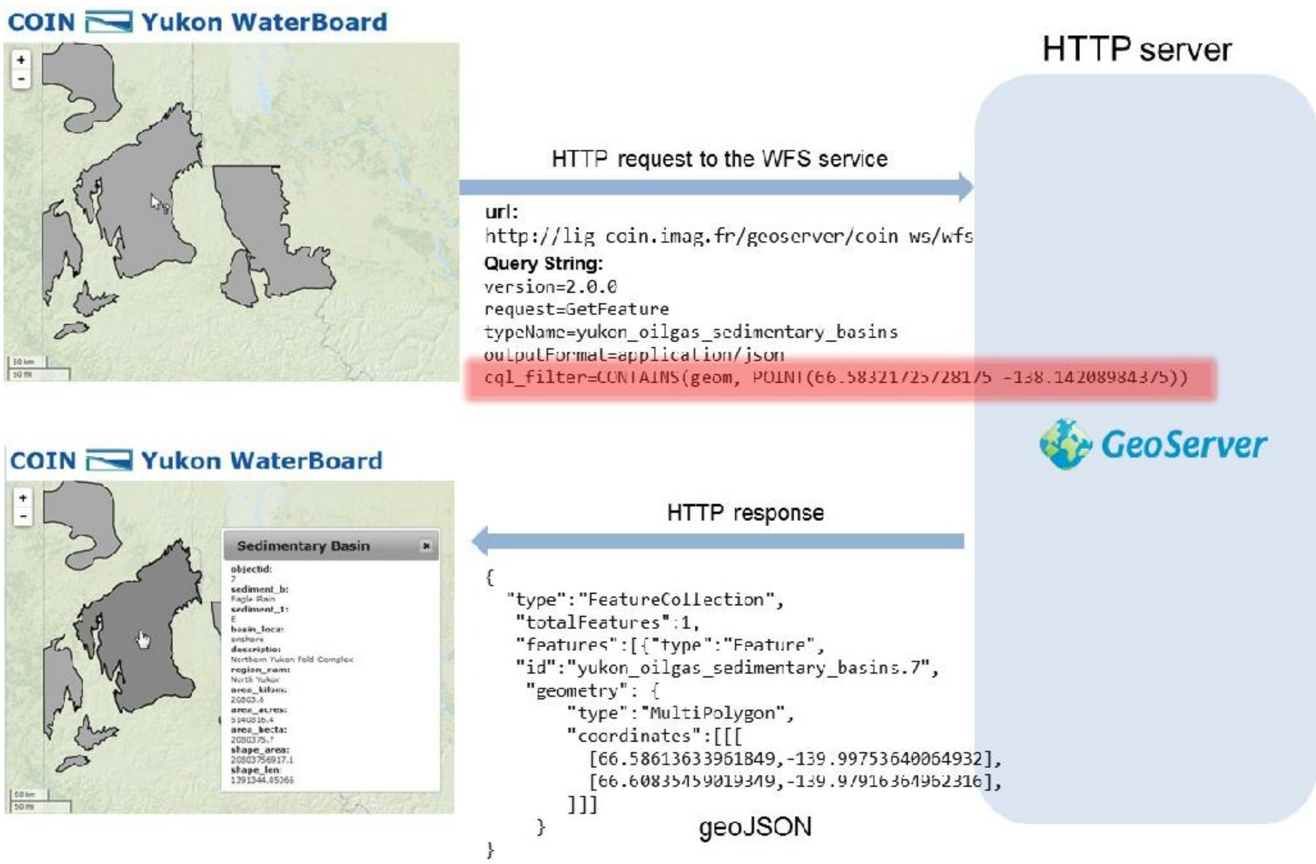
D.8 Operational

Polar View Earth Observation is a global organization providing satellite-based information and data services in the polar regions and the cryosphere that support safe and cost-effective marine operations, improved resource management, sustainable economic growth and risk protection across sectors and around the world. [polarview.org]

The **International Ice Charting Working Group (IICWG)** has provided a forum since 1999 for the worlds ice services to cooperate and coordinate on all matters concerning sea ice and icebergs. Its primary focus is on operational support to marine activities in ice-affected waters accompanied by a strong interest in scientific developments for continuous improvement. [nsidc.org/noaa/iicwg]

Appendix E

Appendix: A Semantic Web Driven Application Scenario



6 Application : COIN

COIN description [2014, 2015] : NRCan GeoConnections program.

6.1 Introduction

In the domain of the environment, the study of the evolution of observation data is a central task. Beyond the analysis of their own data, it is essential for researchers to cross them with other data. The evolution of the Internet greatly facilitates the possibilities of access and data exchange, as is evident in the geomatics implementation of spatial data infrastructures. However, the heterogeneity of models, data, metadata and formats and their change over time, remains a major difficulty in integrating different sources. Currently, after the tremendous growth of Web 2.0, we are witnessing an evolution of the world wide web to what the W3C refers to as web data: a model for simple, flexible and powerful data, the Resource Description Framework (RDF) (Cyganiak et al, 2014), which, based on web infrastructure, facilitates publication and exchange of data across the web; representation models, especially RDFS (Brickley and Guha 2004) and OWL (Hitzler et al., 2012), to define the form of ontologies knowledge to give a semantics to data semantics that can be exploited by computer programs. The initiative of open and linked data (Linked Open Data) (Bizer et al., 2009) follows this line, whose principles were set out by Tim Berners-Lee (Berners-Lee 2006) :

(1) use of URIs (Uniform Resource Identifiers) to name (identify) things,
(2) use of HTTP URIs to consult these addresses,
(3) when a URI is accessed, provide useful information using open standards (RDF, SPARQL, ...),
(4) include links to other URIs in order to discover more things.

These are the principles that we propose to apply to environmental data on water quality in northern Canada among others and which were used to study the impact of climate change in the Arctic and Subarctic regions (Lim al., 2008). In the present document we present the tabular data in RDF processing by associating a semantic model, enrichment by linking to other data, their access through a web interface mapping. COIN allowed us to understand the various steps necessary for such an approach and approach the different technologies necessary for its implementation. Our case study focused on the Arctic and

Subarctic ecosystems, biological and chemical data come either from water monitoring stations or sampling campaigns. The purpose of this data set is to help the researcher to answer the question: what is the status of biodiversity in the arctic and subarctic zones and establish a time line (Lim et al., 2005). Over the years a number of data have been collected, sometimes of different types and forms stored in various file formats. For easy operation, we propose a conceptual model (ontology) that will represent them in RDF format and then link them to other external data sources.

6.2 Ontology and linked open data application

To model data in the form of an RDF graph, an OWL ontology has been specifically defined for this application. This ontology includes a number of classes and properties for representing observational data. To represent spatial information (coordinates of sampling sites, geometry regions) we relied on GeoSPARQL standard proposed by OGC (OGC, 2012) (Battle Kolas, 2012). Figure 5 shows in the form of a UML diagram the various classes and relationships (owl: ObjectProperties) defined for our application.

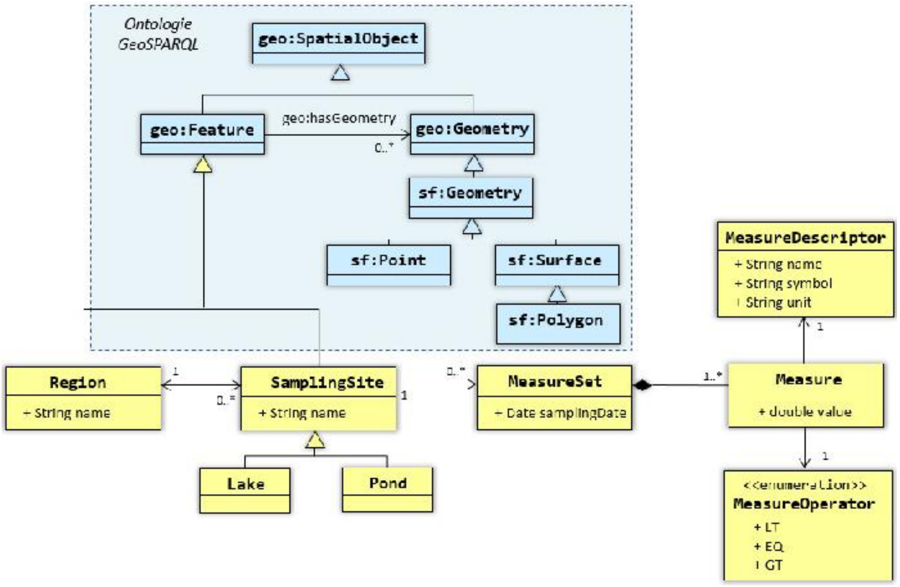


Figure 4: The ontology of the application and its links with GeoSPARQL ontology.

Classes Region and SamplingSite are defined as subclasses of class geo: Feature ontology GeoSPARQL. The spatial representation of these geographic entities is performed using geometry types GeoSPARQL based on OGC Simple Features model (OGC, 2011) and defined as subclasses of class geo: Geometry in our case: sf: Polygon for regions and sf: Point to the sampling sites.

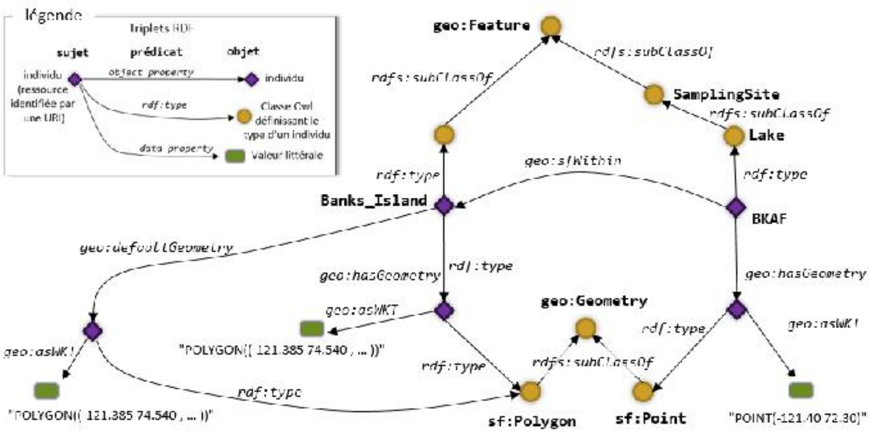


Figure 5: Extract from the RDF graph, based on data from the excel table, the BKAF site on Banks Island.

The spatial coordinates are defined for their part in the form of literals WKT (Well Known Text) associated with the object geometry by the property data (owl: DataProperty) geo: asWKT (Figure 4).

6.2.1 Domain Ontology

We have extended our application ontology with a general ontology from hydrology used by the Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI) (Couch et al., 2014).

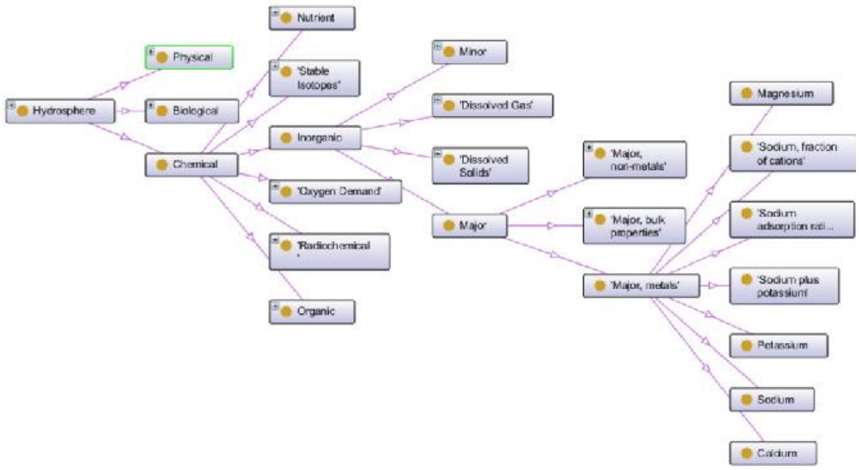


Figure 6: An extract of the ontology of the CUAHSI hydrosphere. Only are developed in this hierarchy, the classes corresponding to the major metal concepts.

This ontology defines a taxonomy that can structure hierarchically more than 4,000 words describing physical, chemical and biological measures related to water. It is used by the System Information CUAHSI (CUAHSI-HIS Hydrologic Information System) and consists of a set of servers and databases connected to client applications as web services to facilitate the discovery of time series data collected at a given point. We have taken this ontology defined in tabular form to translate it as an OWL class hierarchy (Figure 6).

The use of this ontology in our model is made by combining the type of corresponding measure in the terminology of CUAHSI with the descriptors measures identified during sampling.

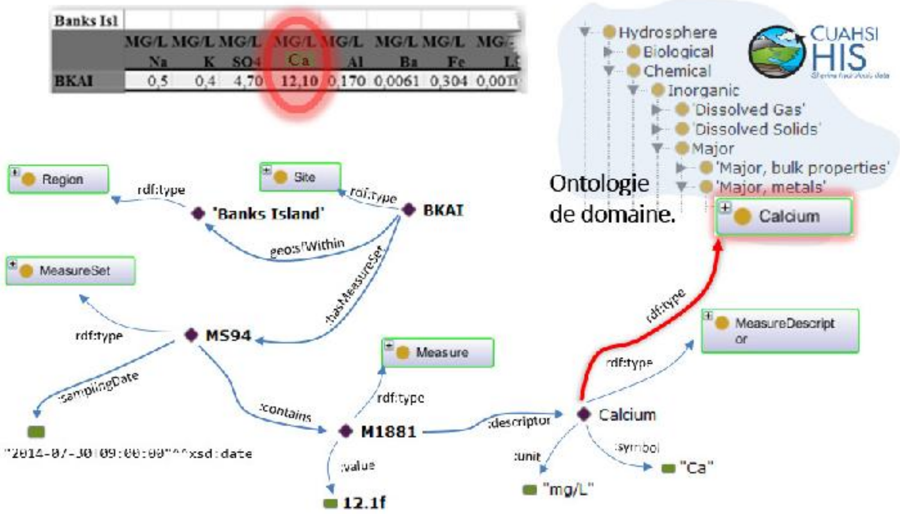


Figure 7: On the BKAI site at Banks Island a set of MS94 measures was taken. The M1881 12.10 measurement value is associated with a descriptor indicating that this is a measure of Calcium and its unit is mg / l. This descriptor is connected to the domain ontology CUAHSI by a relationship rdf: type.

6.2.2 Linking to external data sets

A second enrichment of the initial data that allows their representation in the form of an RDF graph is the ability to link with other external data sets published respecting the principles of open and linked data.

To demonstrate the potential of such enrichment we linked our data represented using our domain ontology with data from DBpedia (Lehmann et al., 2015). We also linked our study data on regions with data from GeoNames.

6.2.3 Liaison with Geonames

The gazetteer GeoNames, which contains more than 10 million names, offers a free geographic database, accessible on the Internet under a Creative Commons license. Besides the geographical coordinates of each location listed, GeoNames offers data such as altitude, population, administrative subdivision, the postal code and links to corresponding Wikipedia pages in multiple languages. In addition to an access via web services, or a raw copy "dump" of the database, GeoNames publishes its data in RDF. Thus, each geographic feature of GeoNames is represented as a web resource. This web resource is identified by a stable URI, which, by negotiating the content, gives access to either the HTML page showing the entity on a map, or to a RDF description of the entity based upon a vocabulary defined by a OWL ontology (Figure 8).

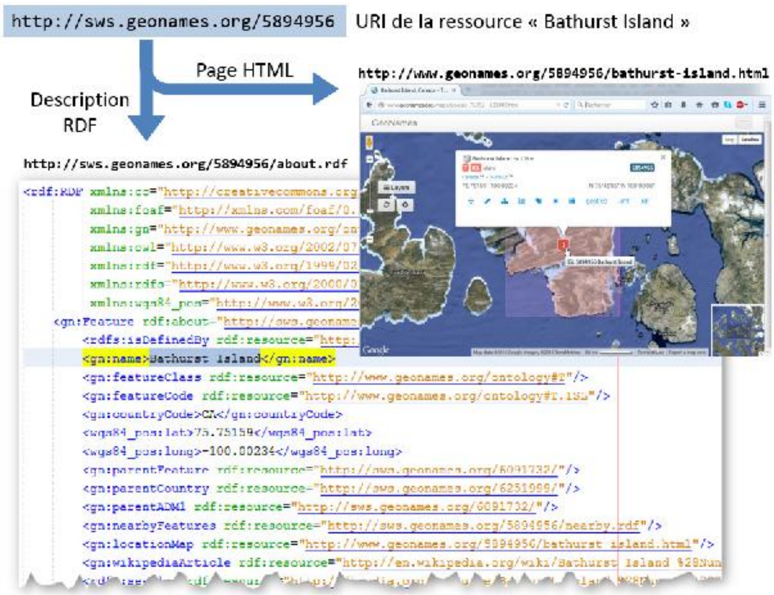


Figure 8: Content Negotiation with GeoNames. Depending on the value of the "*accept"* header of an HTTP request, the resource server sends either an HTML page or an RDF description of this resource.

For each of the regions in which the sites are located, we looked into GeoNames for an entry for the name used by filtering the results based on country and geolocation. This region of northern Canada is relatively poorly covered, sometimes we were driven to add new GeoNames entries (this was the case for example of Bathurst Island). An owl link sameAs is then used to associate a region representation in GeoNames (Figure 7).

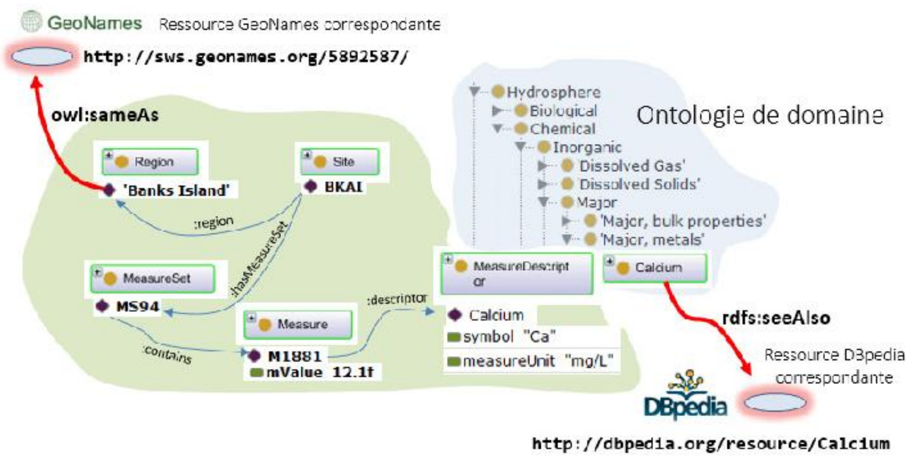


Figure 9: Data Binding to GeoNames and Dbpedia

6.2.4 *Liaison with Dbpedia*

We conducted liaison with DBpedia looking for a corresponding resource in DBpedia for each class of our domain ontology. To perform this alignment, we wrote a program comparing the label concepts from the ontology of CUASHI with labels DBpedia resources (more than 11 million of resources). Once the labels have been standardized (lowercase, removal of non-alphabetic characters and spaces) a search of identical labels is performed. And over 2,000 of the 4,000 defined concepts could be automatically found in Dbpedia. A rdfs link "*seeAlso*" was added between the class of our ontology and the corresponding DBpedia resource (Figure 7). For concepts for which this step fails, a comparison based on Levenshtein distance (Levenshtein 1966) was performed and a list of the best potential candidates was submitted, which then required a manual step to make a possible connection.

6.3 *Architecture and system implementation*

To explore and analyze the observational data thus represented, we propose a web mapping interface for data query through SPARQL and its spatial extension GeoSPARQL. For this raw data from excel files produced during sampling campaigns were converted to RDF and stored in a specially designed database for storage and data recovery RDF called "triplestore".

6.3.1 *RDF data storage*

Although published in 2012, few triplestores currently support the standard GeoSPARQL (Athanasiou et al, 2013.) Strabo (Kyzirakos et al, 2012.) USeekM and Parliament. Our choice fell on Parliament, which has a relatively good balance between ease of installation and use, a support GeoSPARQL and acceptable performance (although far from the performance offered by the spatial databases (Patroumpas et al., 2014)).

On the Parliament server, an RDF graph is created in which are loaded:
- Ontologies used: the domain ontology (prefix cuahsionto :), the application ontology (ccionto :) prefix;
- Observational data from Excel files and converted into RDF using vocabularies defined by previous ontology and GeoSPARQL vocabulary for their spatial dimension.

When charging, Parliament automatically performs a number of inferences: RDFS inference over a number of OWL inferences (equivalent classes or properties; inverse properties : symmetrical, transitive or functional). Once loaded, the data can be queried via GeoSPARQL requests transmitted (via http) to the access point of the Parliament server (Jetty server + Joseki).

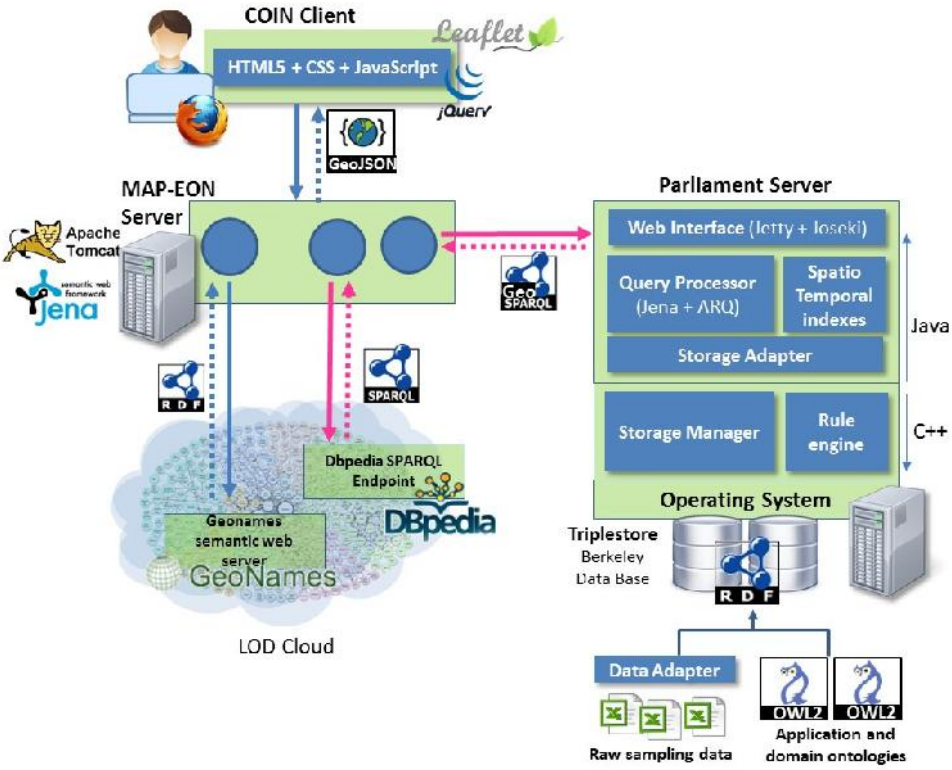


Figure 10: Architecture of COIN

6.3.2 The COIN web application

The COIN application is in the form of a web application. The user interface runs in a web browser on the client. Using HTML5, CSS3 and JavaScript technologies, it relies on Leaflet libraries for the map interface, and jQuery, jQueryUI to manage the interaction and AJAX communications with the server. The COIN Server has meanwhile been implemented using JEE web technologies. A number of servlets, running in an Apache Tomcat, are responsible for processing the requests from the client. As we will see later, these treatments result in some sending SPARQL queries to access points (the Parliament and DBpedia servers) or to retrieve RDF data (GeoNames server). To perform these RDF and SPARQL treatments we rely on the open source Java framework Apache Jena. The data returned by the server to the client use JSON and GeoJSON exchange formats (Buttler et al., 2008) to display geographic information on the map interface.

The interface of the COIN application allows the user to visualize the different sampling sites on an interactive map. The selection of a site on the map provides access to the various comments made on this site. In the dialog box that appears the user has three tabs (Figure 9) that allow the user:

- To filter the measures to display. Filtering is achieved through the hierarchy of concepts of the domain ontology. The RDFS inferences made in loading data help to automatically add a link ("*rdfs: subClassOf*") between each superclass of the Cuashi ontology to the descriptors measures. Only comments with a measurement concept as a descriptor are displayed, with a subclass descriptor of the selected concepts.
- To view all the comments for a given date.
- To view all the comments for a given measurement (time series).

With the links to DBpedia and GeoNames we presented in paragraph 2.3, the user can access additional information that the application will look for dynamically in the web data (Figure 11):

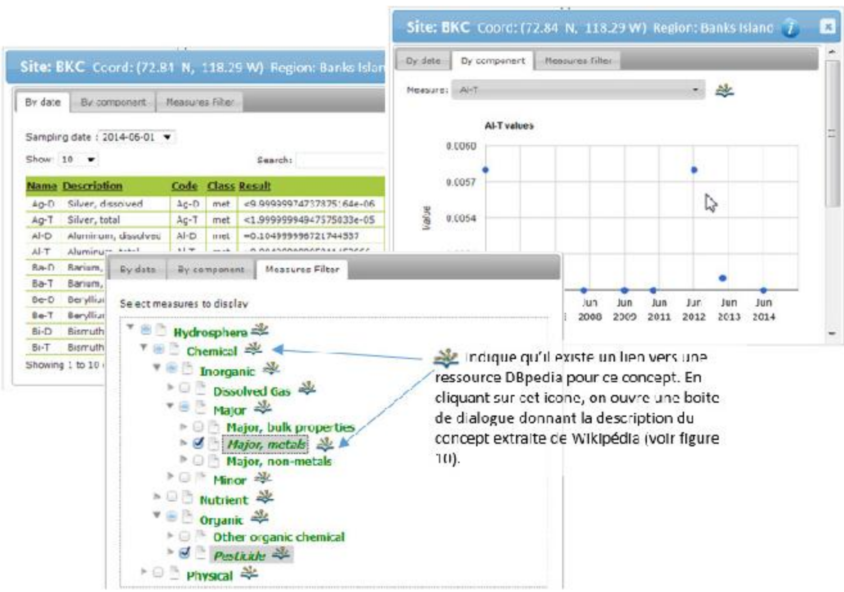


Figure 11: Filtering observations using the domain concepts of hierarchy. Are displayed as comments regarding heavy metals and pesticides.

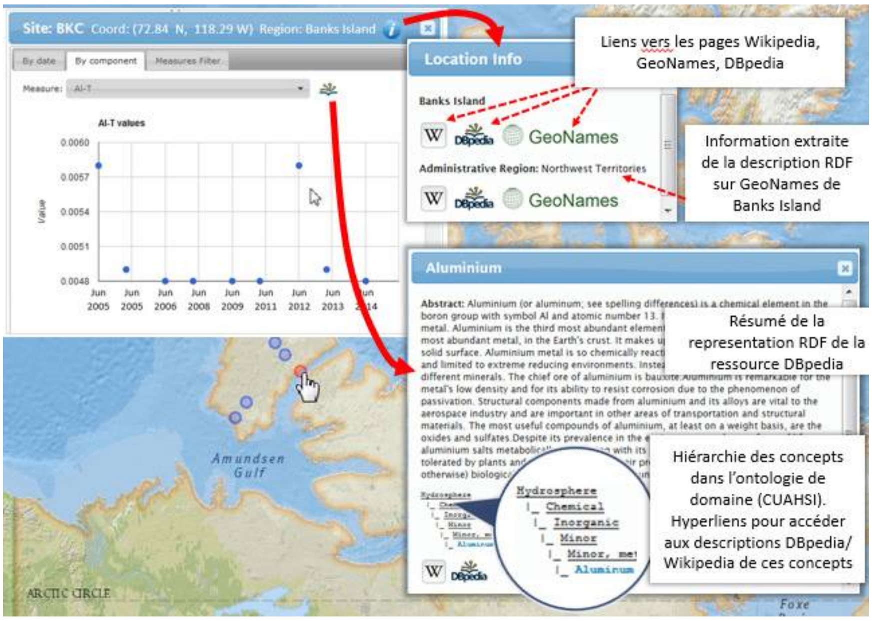


Figure 12: Search for additional information in the web of data using the links GeoNames a region (by Banks Island) or DBpedia a measure (in this case aluminum).

- The GeoNames link retrieves the administrative area of the upper level and Wikipedia pages describing aregion;
- The DBpedia link allows for a given observation in Wikipedia to provide a description of the item measured.

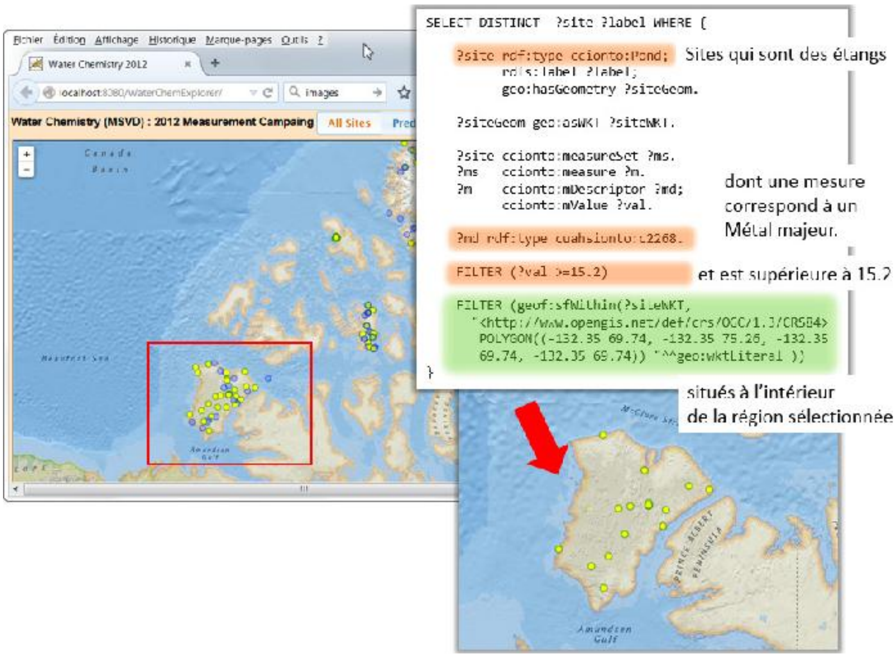


Figure 13: Request combining spatial filter and filter "semantic". Only are displayed ponds (yellow dots) in the selected region and having a heavy metal value greater than 15.2 mg / l.

Furthermore, the use of GeoSPARQL allows for queries combining both a spatial component and a semantic component. For example, the following query "Find all the sites located in an area for which there is a case for a heavy metal whose value is greater than 15.2 mg / l" can be expressed using a query GeoSPARQL (Figure 13):

- Using the RDF Schema type inferences to select only the "Pond" type sites (Pond) (RDF triplet pattern Site rdf: type ccionto: Pond?) And have an observation corresponding to a heavy metal (triplet pattern? md rdf: type cuahsionto: C2268 where md is a measure descriptor and cuahsionto: C2268 URI concept "Major Metal" in the domain ontology (ontology CUAHSI));

- GeoSPARQL uses quantitative spatial processing capabilities by selecting only sites whose geometry is inside selected area: triplets patterns geo Site: hasGeometry siteGeom?. and geo siteGeom: asWKT siteWKT?. possible to recover the geometry of a site in the variable WKT siteWKT; the spatial filter FILTER (geof: sfWithin (siteWKT, "<http://www.opengis.net/def/crs/OGC/1.3/CRS84> Polygon ((-132.35 69.74, 75.26 -132.35, -132.35? 69.74, 69.74 -132.35)) "^^ geo: wktLiteral)) allows to select only sites whose geometry is within the selected area.

6.4 Conclusion

COIN has the advantage of using semantic web techniques for integrating, publishing and visualization of heterogeneous observation data in their content, their format and temporality. The use of semantic queries using spatial-temporal operators also demonstrated the power of such a mechanism in the crossing of raw data and deduction of synthetic data. This experiment helped to update a number of shortcomings. We defined our own observations in our ontology application relying on GeoSPARQL for the spatial part. If it had the advantage of simplicity permitting us to "stick" to the raw data available to us, we lose in generality and interoperability. Ontologies exist for terrestrial observations such O&M and therefore work was conducted by the OGC for the standard definition Observations and Measurements (OGC, 2013). Using these ontologies should allow a better semantic analysis of what an observation and a measure are. The ontologies define particularly finely the requirements for data and associated parameters specifications. At the integration level, links to DBpedia and Geonames have shown the interest of linked data. It would be interesting to go a little further in search of other information available on the web that could enrich our application. Finally, from an architectural point of view, there is still no triplestore that completely implements the GeoSPARQL standard and provides entirely satisfactory performance. But the use of standards allows us to calmly consider migrating to new tools.

7 Scenarios and Use Cases

Presented in the previous paragraph: application.

8 Bibliography

Athanasiou S., Bezati L., Giannopoulos G., Patroumpas K., Skoutas D. (2013). Market and Reaserach Overview. *GeoKnow Deliverable 2.1.1.*

Battle R., Kolas D. (2012). GeoSPARQL: Enabling a Geospatial Semantic Web with Parliament and GeoSPARQL. *Semantic Web Journal,* vol. 3, n° 4, p. 355-370.

Berners-Lee T. (2006). *Linked Data – design issues,* publié en 2006, révisé en 2009, http://www.w3.org/DesignIssues/LinkedData.html.

Bizer C., Heath T., Berners-Lee T. (2009). Linked Data - The story so far. *International Journal on Semantic Web and Information Systems,* vol.5, n° 3, p. 1-22.

Brickley D. Guha R. V. (2004). *RDF vocabulary description language 1.0: RDF Schema. W3C Recommendation 10 February 2004.* http://www.w3.org/tr/rdf-schema/.

Butler H., Daly M., Doyle A., Gillies S., Schaub T., Schmidt C. (2008). *The GeoJSON Format Specification,* http://geojson.org/geojson-spec.html.

Couch A., Hooper R., Pollak J., Martin M., Seul M. (2014), Enabling Water Science at the CUAHSI Water Data Center, *7th Int'l Congress on Env. Modelling and Software,* San Diego, California, USA.

Cyganiak B., Wood D., Lanthaler M. (2014). *RDF 1.1 Concepts and Abstract Syntax. W3C Recommendation 25 February 2014.* http://www.w3.org/tr/rdf11-concepts/.

Harris S., Seaborne. A. (2012). *SPARQL 1.1 Query Language. W3C Proposed Recommendation 8 November 2012.* http://www.w3.org/tr/sparql11-query/.

Hitzler P., Krötzsch M., Parsia B., Patel-Schneider P.F., Rudolph. S. (2012). *OWL 2 Web Ontology Language Primer (Second Edition). W3C Recommendation 11 December 2012.* http://www.w3.org/tr/owl2-primer/

Madin J., Bowers S., Schildhauer M., Krivov S., Pennington D., Villa F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecological Informatics,* vol.2, p. 279-296.

Lehmann J., R. Isele, Jakob M., Jentzsch A., Kontokostas D., Mendes P. N., Hellmann S., Morsey M., van Kleef P., Auer S., Bizer C. (2015). DBpedia – A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia, *Semantic Web Journal*, vol.6, n° 2, p. 167-195.

Levenshtein V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady,* vol. 10, n° 8, p. 707-710.

Lim D., Douglas M., Smol J. (2005). Limnology of 46 lakes and ponds on Banks Island, N.W.T, Canadian Arctic Archipelago, *Hydrobiologia*, vol. 545, n° 1, p. 11-32.

Lim D., Smol J., Douglas M. (2008). Recent environmental changes on Banks Island (N.W.T., Canadian Arctic) quantified using fossil diatom assemblages. *Journal of Paleolimnology*, vol. 40, n° 1, p. 385-398.

Open Geospatial Consortium (2011). *OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture.* http://portal.opengeospatial.org/files/?artifact_id=25355

Open Geospatial Consortium (2012). *OGC GeoSPARQL - A Geographic Query Language for RDF Data*, https://portal.opengeospatial.org/files/?artifact_id=47664

Open Geospatial Consortium (2013). *OGC Abstract Specification, Geographic information - Observations and measurements.* https://portal.opengeospatial.org/files?artifact_id=41579

Patroumpas K., Giannopoulos G., Athanasiou S. (2014). Towards GeoSpatial semantic data management: strengths, weaknesses, and challenges ahead. *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems,* Dallas, Texas, USA.

Parundekar R., Knoblock C.-A., Ambite J.-L. (2010). Aligning Geospatial Ontologies on the Linked Data Web, *Proceedings of the GIScience Workshop on Linked Spatiotemporal Data, 2010 (LSTD-2010),* Zurich, Switzerland.

Kyzirakos K., Karpathiotakis M., Koubarakis M. (2012). Strabon: A Semantic Geospatial DBMS. *11th International Semantic Web Conference (ISWC 2012).*

Appendix F

Appendix: Indigenous Peoples

This work fully acknowledges the following paragraphs, partly provided by the Government of Canada. Though a (partly) Canadian perspective is provided here, we understand that Indigenous Peoples live in many areas of the Arctic.

F.1 Context

Indigenous Peoples is now the preferred collective term in Canada to include First Nations, Inuit and Métis (Environment and Climate Change Canada). Therefore the term is used throughout this report.

"Indigenous Peoples is a collective name for the original peoples of North America and their descendants. The Canadian constitution recognizes three groups of Indigenous peoples: Indians (commonly referred to as First Nations), Métis and Inuit. These are three distinct peoples with unique histories, languages, cultural practices and spiritual beliefs. More than 1.4 million people in Canada identify themselves as an Indigenous person, according to the 2011 National Household Survey.

Indigenous communities are located in urban, rural and remote locations across Canada. They include:

- First Nations or Indian Bands, located on lands called reserves in most cases
- Inuit communities located in Nunavut, NWT, Northern Quebec (Nunavik) and Labrador
- Métis communities located mainly in Alberta, British Columbia, Ontario, Manitoba and Saskatchewan

- Urban communities of Indigenous Peoples (including Métis, non-status Indians, Inuit and First Nation individuals) in cities or towns which are not part of reserves or traditional territories (for example, the British community in Winnipeg). (Indigenous Peoples and communities).

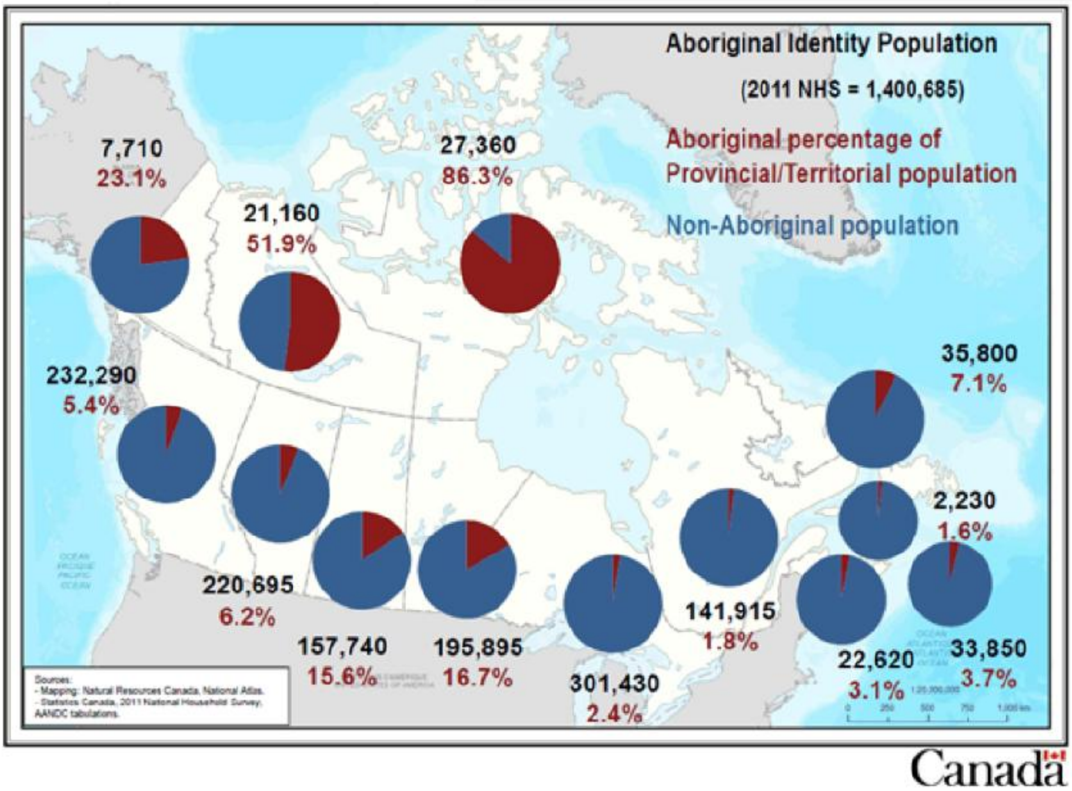


FIGURE F.1: Regional Indigenous Identity Population Proportions. This map shows the pie graphs for each province or territory that indicate the proportion of the regions population in 2011 that consists of Indigenous residents.

F.2 Arctic

[...] Northerners, including Indigenous Peoples who comprise 80 percent of the population in some regions of the Arctic (see F.1) have brought a number of issues to the worlds attention:

- the dangers and challenges posed by climate change;
- the need for sustainable economic development;

- and the importance of sharing experiences and knowledge with our circumpolar neighbours and the world. (Canada's International Gateway).

F.3 Consultation

The Government of Canada consults with Canadians on matters of interest and concern to them. Consulting is an important part of good governance, sound policy development and decision-making. Through consultation, the Crown seeks to strengthen relationships and partnerships with Indigenous peoples and thereby achieve reconciliation objectives. In addition to pursuing policy objectives, the federal government consults with Indigenous peoples for legal reasons. Canada has statutory, contractual and common law obligations to consult with Indigenous groups. The process leading to a decision on whether to consult includes a consideration of all of these factors and their interplay (Indigenous Consultation and Accommodation INAC). Links: Guidelines for Federal Officials to Fulfill the Duty to Consult.

F.4 Ownership, Control, Access and Possession (OCAP) Principles

As evidenced by the Indigenous Community Land and Resource Management review and the First Nations Regional and Inuit Longitudinal Health Survey (RHS), First Nations have taken considerable interest and active ownership of information concerning their communities, particularly in terms of health, culture and environment. In response to the RHS, the First Nations Centre developed the Ownership, Control, Access and Possession (OCAP) principles. The OCAP principles apply to all research, data or information initiatives that involve First Nations. The principles represent a comprehensive framework developed by First Nations to bring self-determination into the realm of research and information management. Of specific relevance to this project is the Data-Sharing Protocol between the First Nations and research partners. It establishes ownership of the data, including how and under what conditions the data may be shared. The protocol also sets out the principles and obligations that partners must adhere to when they collect, use, store and disclose individual or aggregate information. (Best practices for sharing sensitive environmental geospatial data). Link: http://fnigc.ca/ocap.html

F.5 Land Use and Occupancy Mapping

Indigenous peoples in Canada have been mapping aspects of their cultures for more than a generation. Indians, Inuit, Métis, non-status Indians and others have called their maps by different names at various times and places: land use and occupancy; land occupancy and use; traditional use; traditional land use and occupancy; current use; cultural sensitive areas; and so on. I use land use and occupancy mapping in a generic sense to include all the above. The term refers to the collection of interview data about traditional use of resources and occupancy of lands by First Nation persons, and the presentation of those data in map form. Think of it as the geography of oral tradition, or as the mapping of cultural and resource geography. (Tobias, T. 2000).

F.6 Indigenous Communities Strengthen Governance with Location-based Tools in the 21st Century

Today, Indigenous communities are taking advantage of the geomatics technologies that underpin popular online mapping applications and services. At the heart of geomatics lies location-based or geospatial information. By capitalizing on such information as maps and satellite images, Indigenous communities and organizations are better equipped to make informed decisions while managing their lands and resources.

The goal of land and resource management is to balance competing needs, such as the environment and the economy, with Indigenous peoples connection to the land, their cultural values and their responsibility to future generations. (Indigenous Communities and Geomatics).

F.7 Cultural Sensitivities and Considerations in Mapping

Cultural differences exist in the way Indigenous Peoples see, use, and understand maps and data sharing compared to southern perspectives. Representation of space may differ in Indigenous view in contrast with Western view (Chambers, et al, 2004). The following table provides one view on cultural sensitivities and considerations in mapping. It is provided here to exemplarily illustrate different perspectives.

Aboriginal Views	***Western Views***
Holistic view of landscape (derived from history of reliance on immediate environment)	*Economic view of landscape (economic layers, objects and entities)*
Borders permeable, fuzzy, or transparent	*Borders fixed*

FIGURE F.2: Cultural Sensitivities and Considerations in Mapping, source: Chambers, et al, 2004

Elders are traditionally the primary keepers of traditional knowledge, however it is likely the youth that will adopt mapping techniques and software need to consider how Elders knowledge can be passed and retained in mapping data without infringing on traditional rituals of information transfer, how will this affect the social structure (Chambers, et al, 2004).

Settlement areas (land claims, title lands) need to be considered for data collection and services.

F.8 Indigenous Community Land and Resource Management Geospatial Data Needs Assessment

The data needs assessment includes a review of ten Indigenous community land and resource management plans. This review contributed toward the development of a preliminary list of key geospatial data sets that are critical for land and resource management by Indigenous people. The list was validated and the data sets prioritized during workshops with Indigenous community managers and land planners. Indigenous Community Land and Resource Management: Geospatial Data Needs Assessment and Data Identification and Analysis.

F.9 Sensitive Information

For Indigenous Peoples, the land is not only a "home" but also the "food supply" where data sharing practices need to be in place to ensure sustainable land and resource management, cultural protection, and safety and security. Best practices have been developed for sharing sensitive environmental geospatial data (Best practices for sharing sensitive environmental geospatial data).

F.10 Indigenous Mapping Program

The Good Practices Guide: Success in Building and Keeping an Indigenous Mapping Program profiles practices that lead to success when implementing geomatics programs in Canada. The project team undertook a literature review including professional and scholarly research on factors for success when putting geomatics programs into operation, both in Canada and internationally. A list of potential success factors was developed and a survey questionnaire based upon these was written. Representatives of Indigenous organizations across Canada that were running or had operated local mapping programs were invited to fill out the questionnaire and participate in follow up interviews so they could share lessons they had learned while setting up and managing their programs. Their input was used to refine the list and complete a final list of good practices. Practices and advice are grouped under six headings: getting started; gaining leadership and community support; funding and finances; human resources and training approaches; technology, data, and data networks; and support networks. Under these, specific concrete points of advice on principles for success are provided. Additionally, examples from first hand experiences are shared in a case study format to highlight specific principles in action.

Bibliography

Atkinson, R and P Box (2008). *United Nations Spatial Data Infrastructure (UNSDI) Proposed Technical Governance Framework - Preliminary Report Version 1.1*. Tech. rep. UNSDI, pp. 1–52.

Box, Paul (2013). "The Governance of Spatial Data Infrastructure: A Registry Based Model". Ph. D. thesis. The University of Melbourne. URL: http://www.csdila.unimelb.edu.au/publication/theses/Paul%7B%5C_%7DBox%7B%5C_%7DThesis.pdf.

Box, Paul and Abbas Rajabifard (2009). "SDI governance: bridging the gap between people and geospatial resources". In: *Proceedings of GSDI 11 Conference, 15 - 19 June, 2009, Rotterdam, The Netherlands*. Vol. Rotterdam, pp. –—. URL: http://www.gsdi.orgwww.gsdi.org/gsdiconf/gsdi11/papers/pdf/151.pdf.

Commission on Global Governance (1995). *Our Global Neighborhood: The Report of the Commission on Global Governance*. Oxford: Oxford University Press.

Hayes, A., P. L. Pulsifer, and J. P. Fiset (2014). "The Nunaliit Cybercartographic Atlas Framework". In: *Developments in the Theory and Practice of Cybercartography: Applications and Indigenous Mapping*. Amsterdam: Elsevier.

ISAC (2016). *Arctic Observing Summit 2016 - Conference Statement*. URL: http://www.arcticobservingsummit.org/aos-2016-conference-statement-0 (visited on 04/04/2016).

New Zealand Geospatial Office (2011). *Spatial Data Infrastructure Cookbook*. v1.1. November. New Zealand Geospatial Office. ISBN: 9780473192181. URL: http://gsdiassociation.org/images/publications/cookbooks/SDI%7B%5C_%7DCookbook%7B%5C_%7DNew%7B%5C_%7DZealand%7B%5C_%7Dv1-1%7B%5C_%7D17Nov2011.pdf.

OGC/ISO TC211/IHO (2014). *A Guide to the Role of Standards in Geospatial Information Management*. Tech. rep. August. New York, NY, USA: OGC/ISO TC211/IHO, pp. 1–27.

Pulsifer, Peter (2015). *Indigenous Knowledge: Considerations for Arctic Research and Data Management. Report from "Sharing Knowledge: Traditions, Technologies, and Taking Control of our Future Workshop"*. Tech. rep. Boulder, Colorado: Exchange for Local Observations and Knowledge of the Arctic (ELOKA). URL: http://www.arcticobservingsummit.org/sites/arcticobservingsummit.org/files/Pulsifer-ELOKA--Extended%7B%5C_%7DSharing%7B%5C_%7DKnowledge%7B%5C_%7Dstatement.pdf.

Pundsack, J. et al. (2013). *Report on Workshop on Cyberinfrastructure for Polar Sciences*. Tech. rep. June. St. Paul, Minnesota: University of Minnesota Polar Geospatial Center, p. 17. URL: http://www.pgc.umn.edu/system/files/NSF%7B%5C_%7D2013%7B%5C_%7DCyberPolar%7B%5C_%7DWorkshop%7B%5C_%7DReport.pdf.

United Nations (2013). *Spatial Data Infrastructure (SDI) Manual for the Americas*. Tech. rep. New York, NY, USA: United Nations, p. 205. URL: http://unstats.un.org/unsd/geoinfo/RCC/docs/rcca10/E%7B%5C_%7DConf%7B%5C_%7D103%7B%5C_%7D14%7B%5C_%7DPCIDEA%7B%5C_%7DSDI%20Manual%7B%5C_%7DING%7B%5C_%7DFinal.pdf.