

From Data Collection to Long-Term Preservation and Use: Arctic data as part of a global system

Peter L. Pulsifer

Research Scientist, National Snow and Ice Data Center, University of Colorado Boulder

Chair, IASC-SAON Arctic Data Committee

Co-Lead, **IARPC** Arctic Data Sub-Team

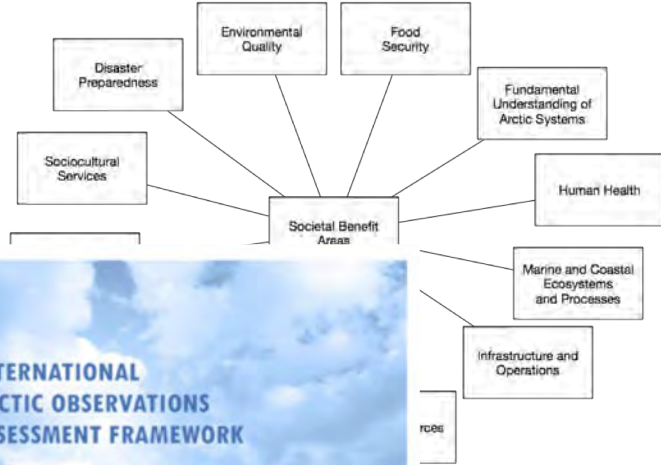
Co-Lead GEO Cold Regions Initiative

3 December 2018, UAK Research School, Longyearbyen, Svalbard, Norway

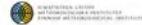
The World of Data

Polar Data in the Global Data System

Why Does Data Matter?



Arctic Observation Value Tree Analysis. IDA Science and Technology Policy Institute, Oslo, Norway. 73 pp.



Arctic Observation Value Tree Analysis
- assessment for physical and oceanic variables

Mikko Strahlendorff, FMI
thanks to IDA STPI and SAON input



<https://www.skepticalscience.com/>



<https://osf.io/>



CLIMATE CHANGE
THREATENING POLAR BEAR
POPULATION, GROUP SAYS



Recent History of Arctic and Polar Data



Polar Data Forum Series

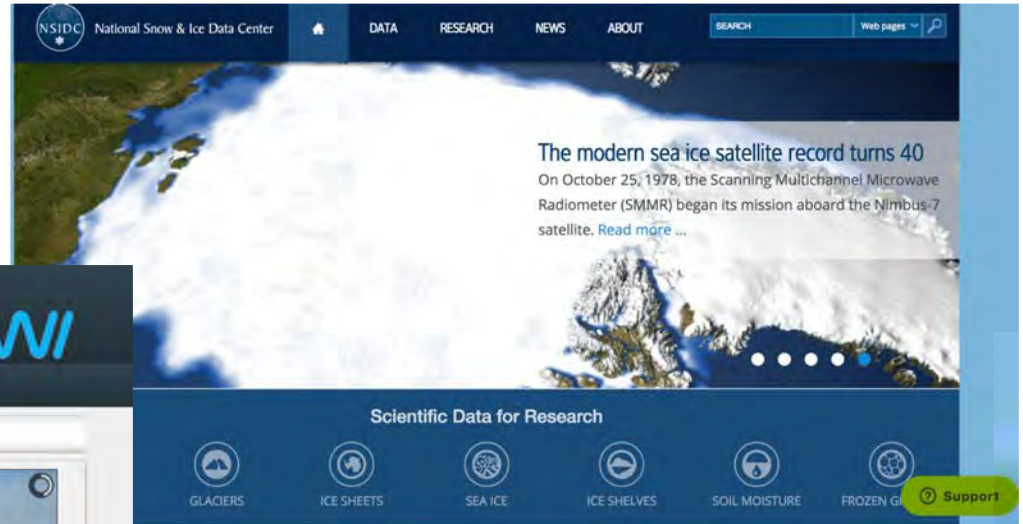
Domains of Data

The Many Flavours of Data

Physical Sciences



<https://gtnp.arcticportal.org/>



<https://nsidc.org/>

<https://geobon.org/>

Life Sciences

A screenshot of the GEO BON Dataset and Metric selection interface. It features a "Dataset" dropdown menu and a "Metric" dropdown menu. Below these, there is a "Plot Options" section with a "Country" dropdown menu and a "sum" button. A "Calculate" button is located at the bottom.

Q Back to search

Arctic Species Trend Index (ASTI)

Updated: 3 years ago

The Circumpolar Biodiversity Monitoring Program, a cornerstone programme of the Conservation of Arctic Flora and Fauna (CAFF), Arctic Council working Group is an international network of scientists, government agencies, Indigenous organizations and conservation groups working together to harmonize and integrate efforts to monitor the Arctic's living resources. CBMP experts are developing four coordinated and integrated Arctic Biodiversity Monitoring Plans to help guide circumpolar monitoring efforts. Results will be channeled into effective conservation, mitigation and adaptation policies supporting the Arctic. These plans represent the Arctic's major ecosystems (Marine, Freshwater, Coastal, Terrestrial). It is important that monitoring programs develop the most effective reporting strategies if they are to inform decision making. To facilitate effective and consistent reporting, the CBMP has chosen a suite of indices and indicators that provide a comprehensive picture of the state of Arctic biodiversity – from species to habitats to ecosystem processes to ecological services. These indices and indicators are developed in a hierarchical manner, allowing users to drill down into the data from the higher-order indices to more detailed indicators. These are being developed through an expert consultation process. The Arctic Species Trend Index (ASTI) is part of this suite of indicators and indices developed by CAFF's CBMP. It tracks trends in over 300 Arctic vertebrate species and comprises the Arctic component of the Living Planet Index. It is important to identify how wildlife and ecosystems are changing in order to develop

A screenshot of the GBIF Global Biodiversity Information Facility homepage. The header features the text "GBIF | Global Biodiversity Information Facility" and "Free and open access to biodiversity data". Below this is a navigation bar with links for "OCCURRENCES", "SPECIES", "DATASETS", "PUBLISHERS", and "RESOURCES". A search bar is located on the right. The main content area displays statistics: "Occurrence records: 1,038,516,773", "Datasets: 41,734", "Publishing institutions: 1,315", and "Species: Learn more about the number of species covered by data in GBIF.org". There are also images of a beetle, a flower, and a landscape.

<https://www.gbif.org/>

<https://www.caff.is/asti/>

<http://www.arctichorizons.org/>

Social Sciences

https://iseralaska.org/static/living_conditions/microdata.htm

**Survey of Living Conditions in the Arctic:
Inuit, Saami, and the Indigenous Peoples of Chukotka**



What is SLiCA?

The Survey of Living Conditions in the Arctic, or SLiCA, is an international joint effort of research and indigenous people to measure and understand living conditions in the Arctic. This website is intended to promote the use and understanding of SLiCA data.

FIND DATA - START SHARING DATA - MEMBERSHIP - SUMMER PROGRAM - TEACHING & LEARNING - DATA MANAGEMENT & CURATION

Search for studies, publications, variables, and webpages

Search

Log In/Create Account



ICPSR
Social science in the public interest

<https://www.icpsr.umich.edu/icpsrweb/>



HOME ABOUT BLOG WORKSHOPS LITERATURE LEADERSHIP LOGIN

ARCTIC HORIZONS

workshops join us

Search

Standards - Resources - Training - Community - Publications - About -

Document, Discover and Interoperate

The Data Documentation Initiative (DDI) is an international standard for describing the data produced by surveys and other observational methods in the social, behavioral, economic, and health sciences. DDI is a free standard that can document and manage different stages in the research data lifecycle, such as conceptualization, collection, processing, distribution, discovery, and archiving. Documenting data with DDI facilitates understanding, interpretation, and use — by people, software systems, and computer networks. Use DDI to Document, Discover, and Interoperate!

Specification Tools Learn Collaborate

<https://www.ddialliance.org/>

Indigenous Knowledge, CBM and Information Systems

- Growing group actively working to share Indigenous Knowledge, information and data
- Progress needed on bridging worldviews, concepts and semantics represented in information systems
- Indigenous Peoples must lead engagement and work with their knowledge – information sovereignty important

<http://www.inuitknowledge.ca/>



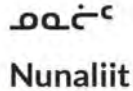
<http://www.arcticcbm.org/index.html>



<http://ititaq.ca/>



<http://eloka-arctic.org/>



Nunaliit Atlas Framework

<http://nunaliit.org/>

<http://www.inuitcircumpolar.com/community-based-monitoring.html>



<https://www.smartice.org/>



<https://arcticeider.com/siku>

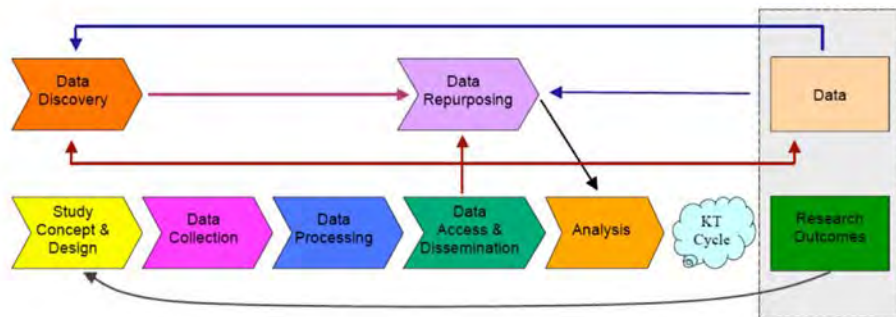


<http://trailmarksys.com/>

The Data Lifecycle

End to End Management and Use of Your Data

The Data Lifecycle



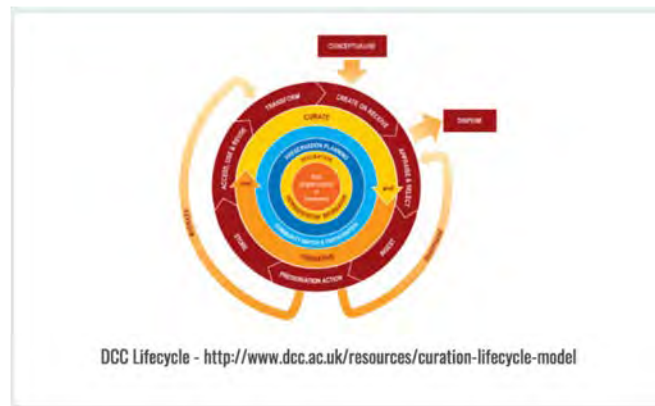
<http://datalib.library.ualberta.ca/~humphrey/lifecycle-science060308.doc>.



FGDC Lifecycle - <https://www.fgdc.gov/policyandplanning/a-16/stages-of-geospatial-data-lifecycle-a16.pdf>



DataONE Life Cycle - <https://www.dataone.org/data-life-cycle>



DCC Lifecycle - <http://www.dcc.ac.uk/resources/curation-lifecycle-model>

Research Data Canada

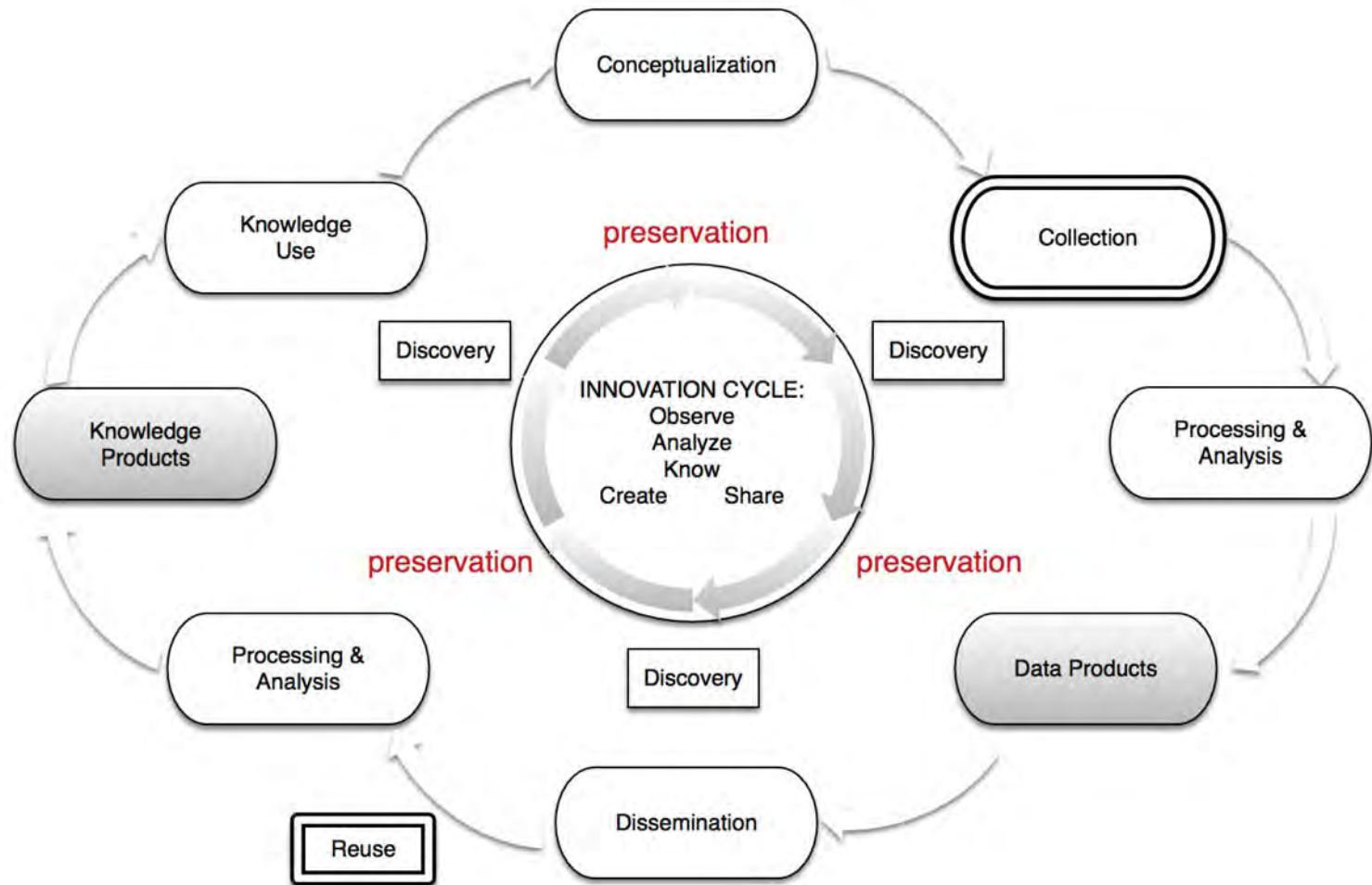
2. DISSEMINATION

1. PRODUCTION



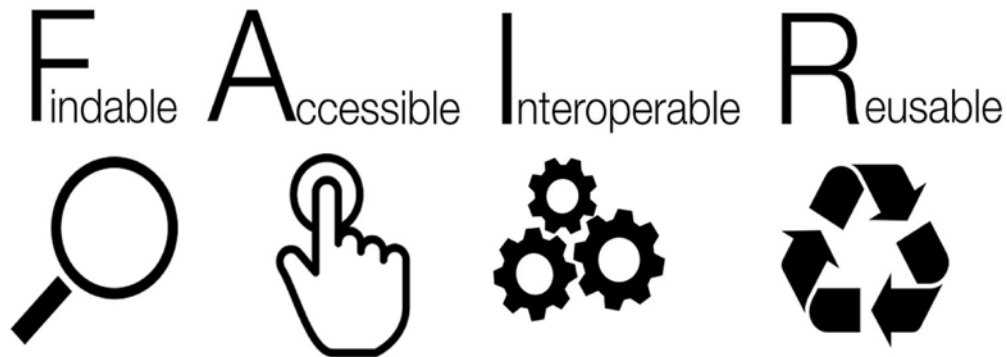
4. DISCOVERY & REPURPOSING

<http://dmtclearinghouse.esipfed.org/>

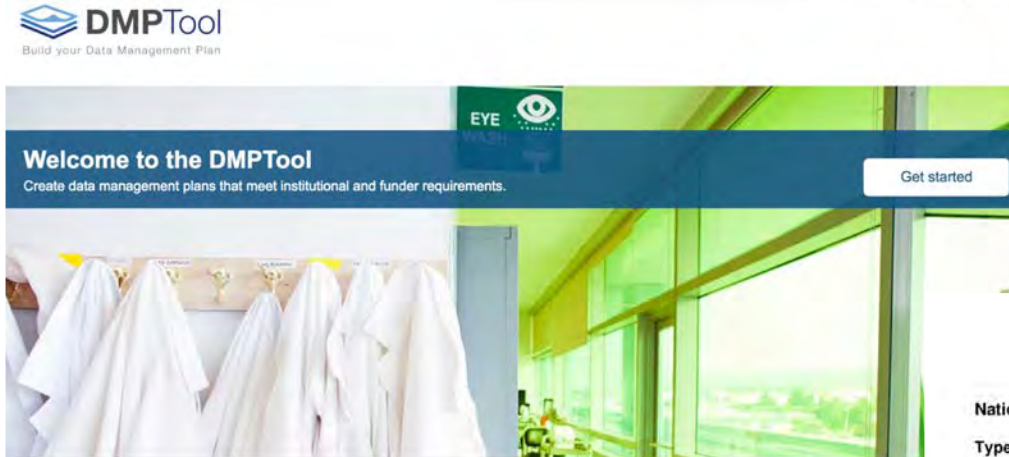


Grounding the Lifecycle Approach

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable
- “FAIR” principles



Data Management Planning



<https://dmptool.org/>

DMPTool
Build your Data Management Plan

Learn Sign in Language

Funder Requirements

Templates for data management plans are based on the specific requirements listed in funder policy documents. The DMPTool maintains these templates; however, researchers should always consult the program officers and policy documents directly for authoritative guidance. Sample plans are provided by a funder or another trusted party.

Search

Template	Download	Funder	Last Updated	Funder Links	Sample Plans (if available)
Alfred P. Sloan Foundation	DOCX PDF	Alfred P. Sloan Foundation	04-18-2018	Sloan Grant Proposal Guidelines	
Arctic Data Center: NSF Polar Programs (DRAFT)	DOCX PDF	National Science Foundation (NSF)	08-23-2018	NSF Arctic Data Center DMP Resources	
BCO-DMO NSF OCE: Biological and Chemical Oceanography	DOCX PDF	National Science Foundation (NSF)	04-24-2018	NSF OCE Sample and Data Policy NSF GEO Directorate Guidance	
Department of Defense (DOD)	DOCX PDF	Department of Defense (DOD)	09-15-2018	DOD Public Access Plan Data Archiving Plans for NIJ Funding Applications DOE Policy for Digital Research Data Management	
Preprint of Science (PLOS) Preprint	DOCX	Preprint of Science (PLOS)	As of 2018		

https://dmptool.org/public_templates

National Science Foundation (NSF): NSF-GEN: Generic

Types of data produced

The types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project.

Guidance:

- [NSF Proposal & Award Policies & Procedures Guide \(PAPPG\)](#)
- [NSF plans for data management and sharing of the products of research \(PAPPG\)](#)
- [NSF Dissemination and Sharing of Research Results](#)
- [NSF Frequently Asked Questions \(FAQs\) for Public Access](#)

Data and metadata standards

The standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies).

Guidance:

Data Collection

The Many Forms of Data Creation

Observations Become Data



Data Sharing

Making Your Data Available to the World for the Greater Good and Credit

Different Ways of Sharing



<https://multimedia.journalism.berkeley.edu/tutorials/ftp/>

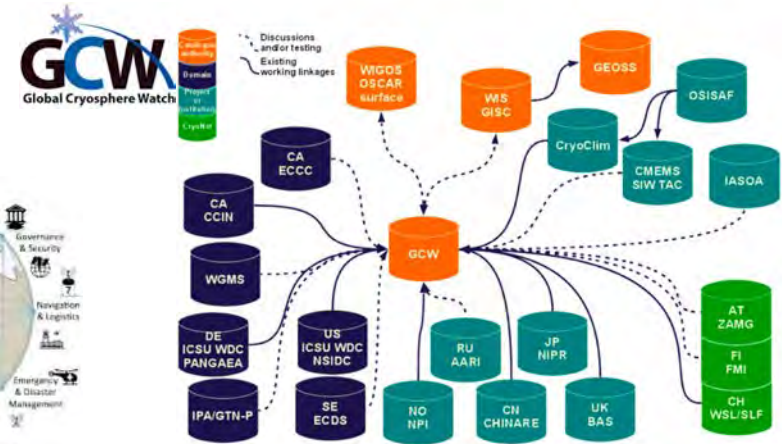
<https://www.pmel.noaa.gov/arctic-zone/data.html>

Data Interoperability

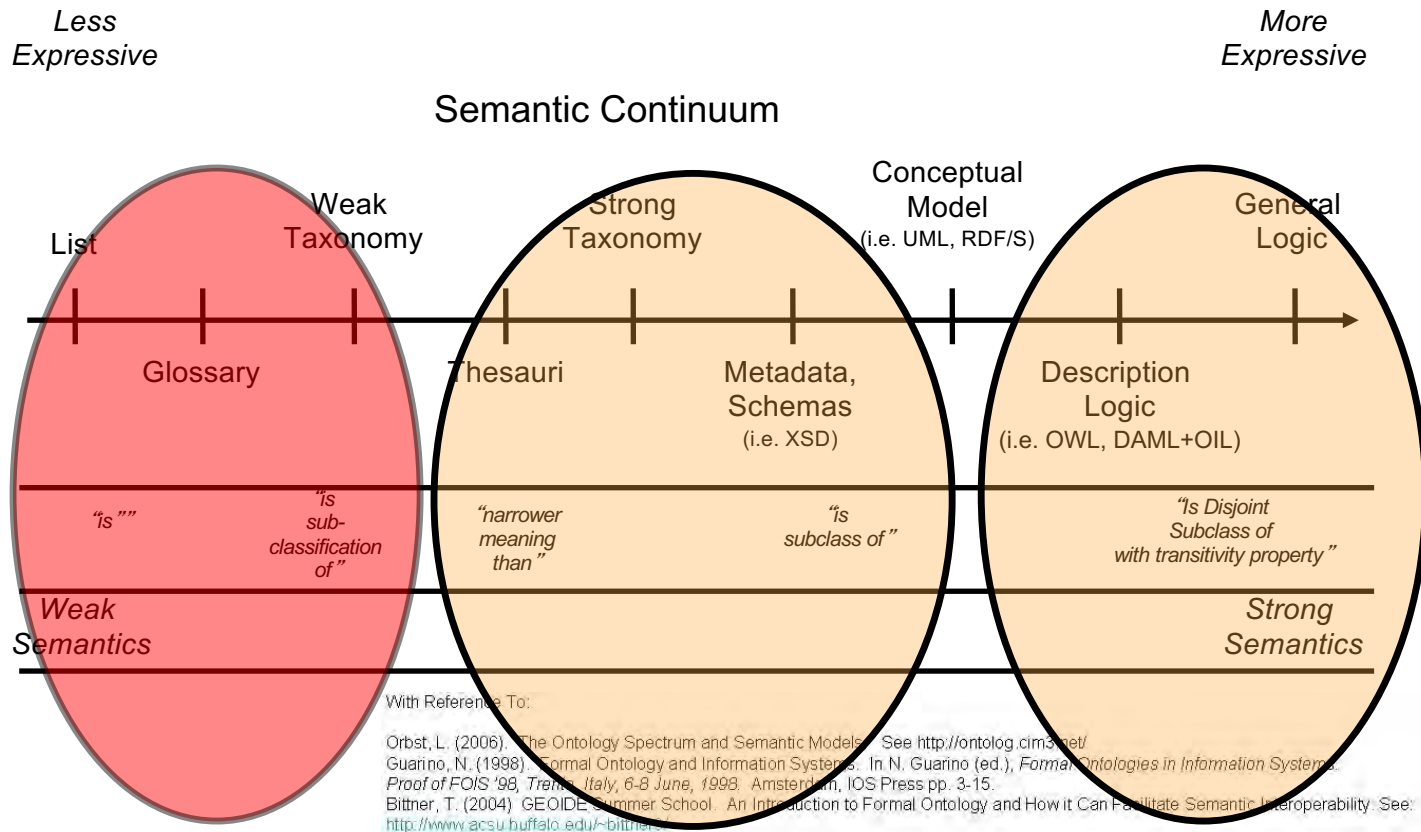
- “Live” data sharing between and among systems
- Standards and specifications
 - Discovery standards
 - Data standards
- “Services” (**Data as a Service**) use standards to make the data widely available “on demand”



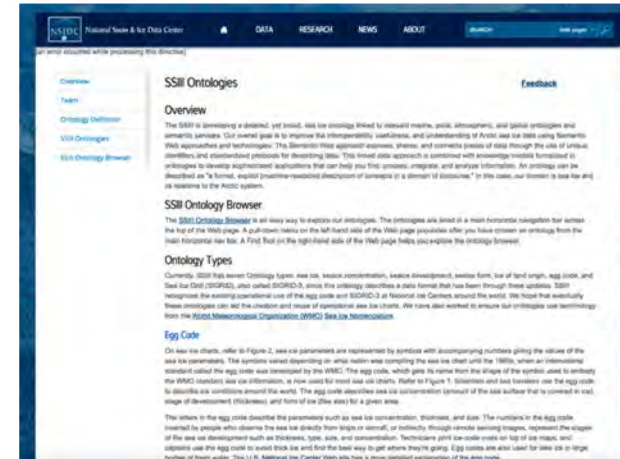
OAI-PMH



What is the meaning of this! ... data semantics



<https://github.com/rduerr/ssiii>



Data Discovery

Finding Data and Making Your Data Findable and Citable

Data Discovery Tools



ANTARCTIC MASTER DIRECTORY
A Global Change Master Directory Portal

HOME DATA SEARCH DATA SERVICES AUTHORIZING TOOLS NADC PORTALS SCAR PROJECTS ASTROPHYSICS

Find Data Sets by Topic:

- Agriculture**
agricultural/aquatic sciences,
agricultural chemicals...
- Oceans**
aquatic sciences,
bathymetry/seafloor
topography...
- Atmosphere**
aerosols, air quality...
- Paleoclimate**
ice core records, land records...
- Biological Classification**
animals/invertebrates...
- Solid Earth**
earth oceans/liquids...

Data Set Text Search
Go
Search tips



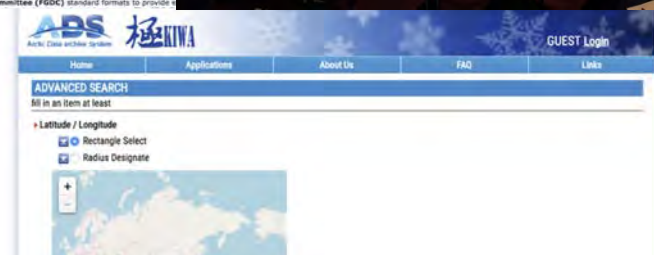
POLAR DATA CATALOGUE

Access to the Polar Data Catalogue

The Catalogue is a database of metadata and data that describes, indexes, and provides access to diverse data sets generated by Antarctic researchers. The metadata records follow ISO 15926 and Federal Geographic Data Committee (FGDC) standard formats to provide a common data structure. The records cover a wide range of disciplines from natural sciences and are available to the public and researchers alike and allows searching data using a map interface.

on the FGDC Search map below to start searching for datasets or sign in to the FGDC Catalogue.

to Search is also available for users with limited internet speed.

ADS Arctic Data System

Home Applications About Us FAQ Links

ADVANCED SEARCH
Fill in an item at least

Latitude / Longitude
☐ Rectangle Select
☐ Radius Designate



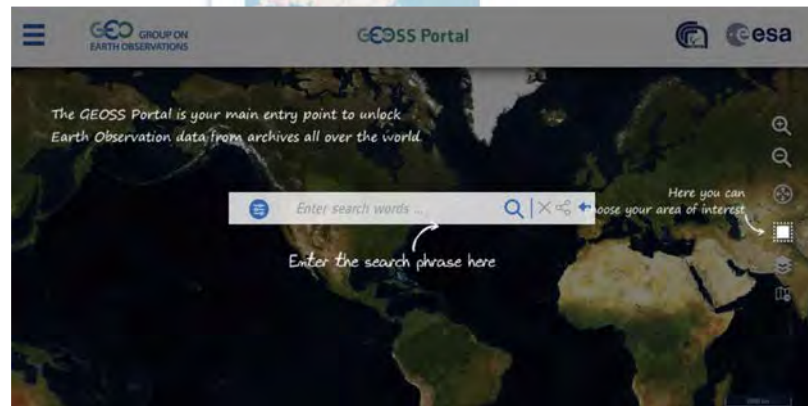
Arctic Data Explorer

Enter term(s) [] Find Data Now Reset

N: 90 S: 45 E: 180 W: -180 From yyyy-mm to yyyy-mm

Diverse Arctic Research Data
Search multiple repositories in one convenient place. Search results will direct you towards free data downloads.

Repository Name	Datasets
NSF Arctic Data Center (NSF-ADC)	10987
National Science and Ice Data Center (NSIDC)	539
UCAR NCAR - Earth Observing Laboratories (UCAR NCAR EOL)	2024
UCAR NCAR Research Data Archive (UCAR NCAR RDA)	172
NOAA National Oceanographic Data Center (NOAA NODC)	6152
Stenostrup Meteorological Institute (Steno)	201
NASA Earth Observing System (EOS) Climate House (ECHOS) (NASA ECHOS)	14390
International Council for the Exploration of the Sea (ICES)	470
U.S. Geological Survey ScienceBase (USGS ScienceBase)	809
Biological and Chemical Oceanography Data Management Office (BCO DMO)	1512
Polar Data Catalogue (PDC)	2350
ICAR The Digital Archaeological Record (IDAR)	81
Rolling Desk in Repertory (RDR)	913
NOAA National Centers for Environmental Information World Data Service for Paleoclimatology (NOAA WDS Paleoclim)	2700
Global Meteorological Network for Permafrost (GNET-P)	1484



GEOSS Portal

The GEOSS Portal is your main entry point to unlock Earth Observation data from archives all over the world.

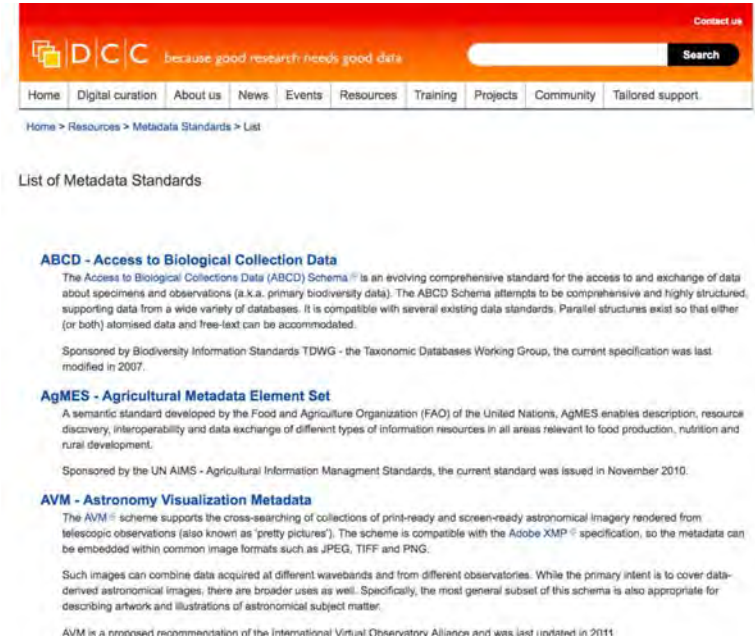
Enter search words [] [X] [] [] []

Here you can choose your area of interest

Enter the search phrase here

Metadata Standards

- Digital Object Identifier (publication, data)
- Dublin Core (general)
- DIF (Science)
- ISO 19115 (Geo)
- FGDC (Geo)
- Data Documentation Initiative (Soc. Sci)

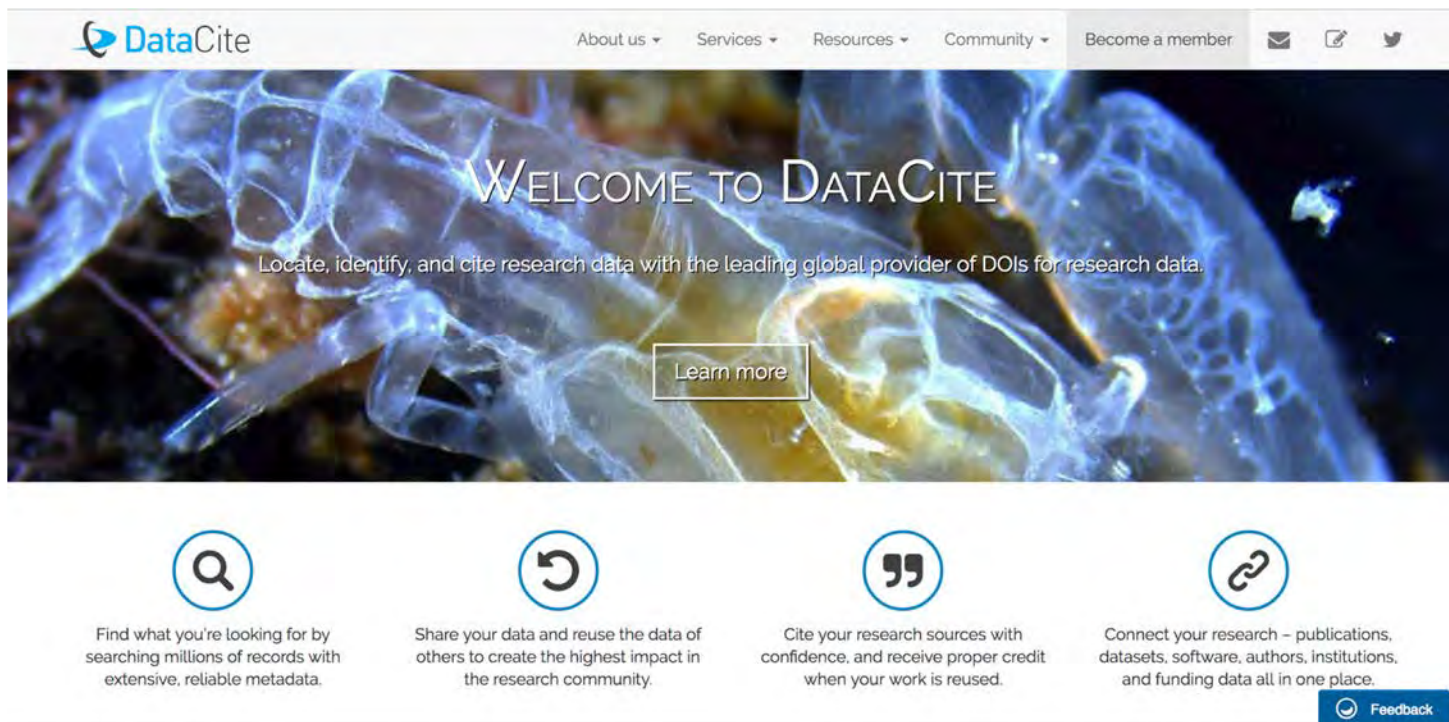


<http://www.dcc.ac.uk/resources/metadata-standards/list>

<http://bit.ly/MDStandards>

Data Publication

<https://www.datacite.org/>



The screenshot shows the DataCite website homepage. At the top is a navigation bar with the DataCite logo on the left and links for 'About us', 'Services', 'Resources', 'Community', and 'Become a member' on the right. Below the navigation bar is a large hero section with a background image of a blue, translucent, biological structure. The text 'WELCOME TO DATACITE' is centered in the hero section, followed by the tagline 'Locate, identify, and cite research data with the leading global provider of DOIs for research data.' and a 'Learn more' button. Below the hero section are four columns, each with a circular icon and a description of a service: 1. A magnifying glass icon for finding research data. 2. A circular arrow icon for sharing and reusing data. 3. A quote icon for citing research sources. 4. A link icon for connecting research publications, datasets, software, authors, institutions, and funding data. At the bottom right is a 'Feedback' button.

DataCite

About us ▾ Services ▾ Resources ▾ Community ▾ Become a member

WELCOME TO DATACITE

Locate, identify, and cite research data with the leading global provider of DOIs for research data.

[Learn more](#)

Find what you're looking for by searching millions of records with extensive, reliable metadata.

Share your data and reuse the data of others to create the highest impact in the research community.

Cite your research sources with confidence, and receive proper credit when your work is reused.

Connect your research – publications, datasets, software, authors, institutions, and funding data all in one place.

[Feedback](#)

Schema.org + Google Data Search



Welcome to Schema.org

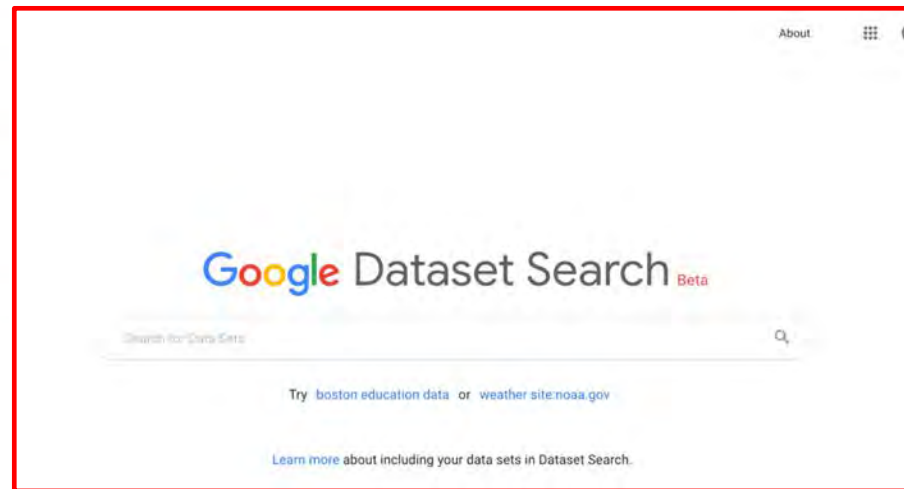
Schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond.

Schema.org vocabulary can be used with many different encodings, including RDFa, Microdata and JSON-LD. These vocabularies cover entities, relationships between entities and actions, and can easily be extended through a well-documented extension model. Over 10 million sites use Schema.org to markup their web pages and email messages. Many applications from Google, Microsoft, Pinterest, Yandex and others already use these vocabularies to power rich, extensible experiences.

Founded by Google, Microsoft, Yahoo and Yandex, Schema.org vocabularies are developed by an open [community](#) process, using the public-schemaorg@w3.org mailing list and through [GitHub](#).

A shared vocabulary makes it easier for webmasters and developers to decide on a schema and get the maximum benefit for their efforts. It is in this spirit that the founders, together with the larger community have come together – to provide a shared collection of schemas.

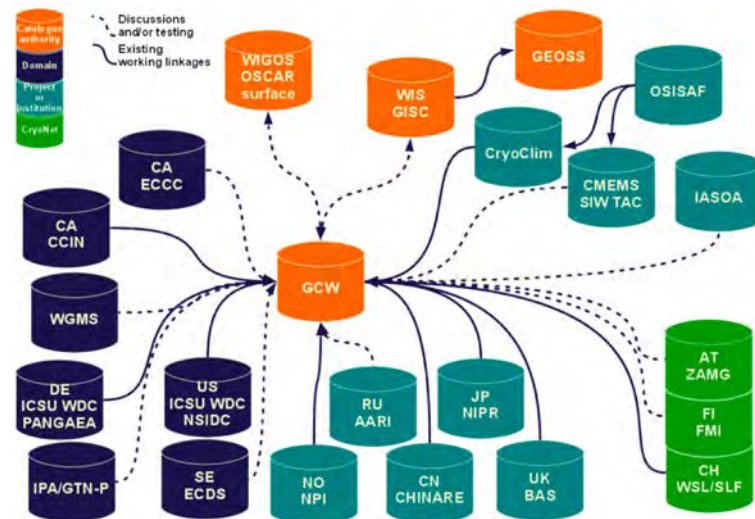
We invite you to [get started!](#)



<https://toolbox.google.com/datasetsearch>

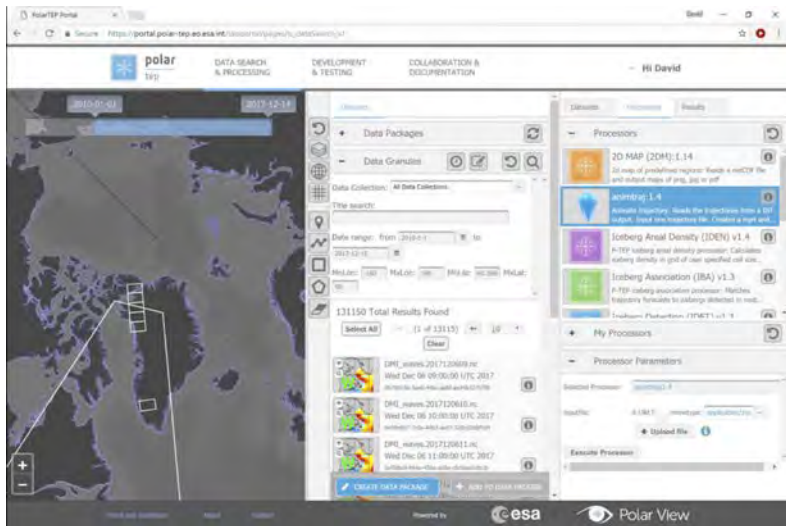
Data Reuse

Create Once, Use Many Times



Cloud Platforms/Virtual Research Environments

<https://portal.polar-tep.eo.esa.int>



Polar Thematic Exploration Platform

<https://researchworkspace.com>



<https://earthengine.google.com>



Google Earth Engine

Training Activities and Resources



The screenshot shows the homepage of the Data Management Training (DMT) Clearinghouse. At the top is a dark blue navigation bar with the ESIP logo and the text "Data Management Training". Navigation links include Home, Browse, Search, Submit, Help, and About. A "Log in" link is on the right. The main content area has a heading "Welcome to the DMT Clearinghouse" followed by a paragraph describing the clearinghouse as a registry for online learning resources. It mentions collaboration with the U.S. Geological Survey's Community for Data Integration, the Earth Sciences Information Partnership (ESIP), and DataONE. A contact email "clearinghouseEd@esipfed.org" is provided. A "Read More" button is below. To the right is a circular diagram labeled "DCC Lifecycle" with arrows indicating a cycle of "Discover", "Describe", "Distribute", and "Delete". Below the diagram is the URL "http://www.dcc.ac.uk/resources/curation-lifecycle-model". At the bottom are three teal boxes: "Search" with a text input and a "Search" button, "Browse" with a "Browse" button, and "Submit" with a "Submit" button.

<http://dmtclearinghouse.esipfed.org/>

Arctic Data Community

Working Together for Science and Society

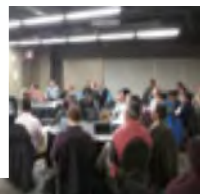
Data Sharing: Cooperation from Local to Global

- Significant progress made since International Polar Year
- Framework for cooperation exists
- Recent collaboration being leveraged to establish concrete “architecture”
- All perspectives and actors must be included



Arctic Data Committee

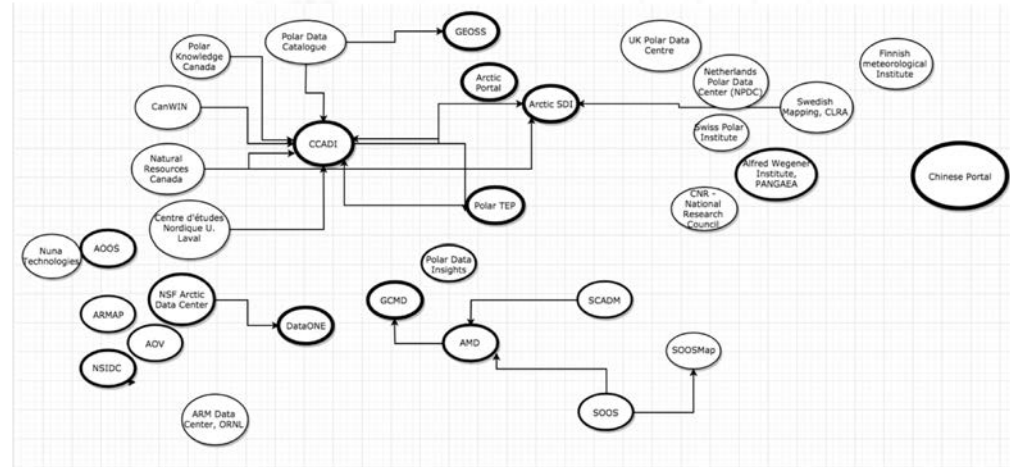
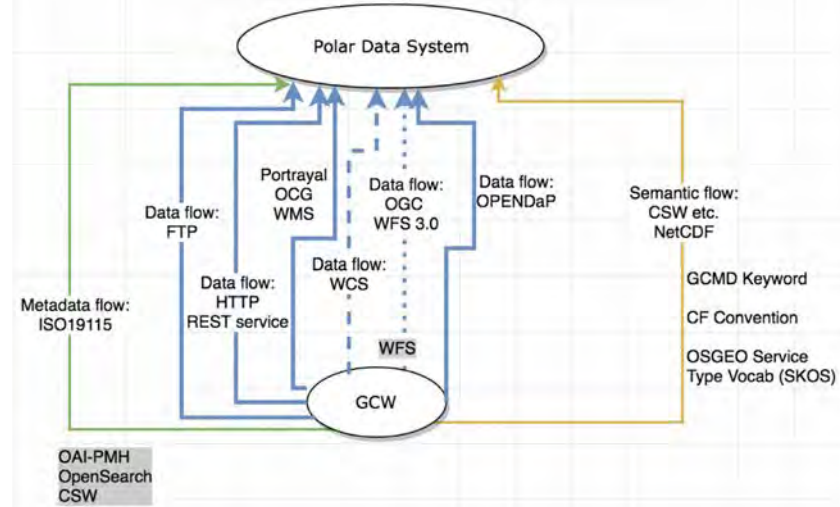
- Formed Nov '14
- IASC-SAON partnership
- National and voluntary members + Indigenous (2017)
- Promote and enable:
 - **Understanding the system**
 - **Effective data policy**
 - **Infrastructure**
 - **Ethically open access**
 - **Attribution**
 - **Standards and interoperability**



<http://arcticcdc.org>

Polar Data and Systems Architecture Workshop

28 – 30 November 2018, Geneva



Conclusions

- Think beyond your research project to broader sharing and data reuse – your data matters to the world!
- Think win-win
- Take advantage of the activities of the data community to help support your data management activities (e.g. identifying training materials, standards, repositories etc.)

Thank you!