Creating a Polar Information Commons

Knowledge is the common wealth of humanity
- Adama Samassekou
Convener of the UN World Summit on the Information Society

Bob Chen, CODATA
Mark Parsons, NSIDC
David Carlson, IPY Programme Office
Kathleen Cass, CODATA
Taco de Bruin, NIOZ
Kim Finney, AADC
Kim Jochum, APECS
Kaitlin Thaney, Science Commons
Sponsors

• International Council for Science
• World Meteorological Organization
• International Arctic Science Committee
• Scientific Committee on Antarctic Research
• International Polar Year International Program Office
• World Data System
• International Union of Geodesy and Geophysics
• Royal Netherlands Academy of Sciences
• Creative Commons/Science Commons
• National Science Foundation
Larsen B Iceshelf breakup, 31 January to 7 March 2002

If we knew what it was we were doing, it would not be called research, would it?
- Albert Einstein

**Arctic Ocean**

September Sea Ice Extent:
Observations (red) versus
Models and Model Mean (± s.d.)
(averaged model data and s.d. in black)
The problem of coastal erosion

Courtesy IARC, Dave Sanches
• earlier break up / later freeze up (2-3 weeks each)
• increased weather variability / traditional forecasts no longer work
• sea ice thinner; poorly formed (poor strength/integrity)
• seals molt too late in summer and some skins are in poor condition; impact on making clothing
• seasonal calendar off; some names no longer apply
• snow is harder due to shifting wind patterns; impact on making igloos
• glaciers and aniuvat (perennial snow patches) melting more rapidly
50,000 participants from 63 nations
IPY Data Policy

http://www.ipy.org/Subcommittees/final_ipy_data_policy.pdf

Special Cases:
• Human subjects
• Intellectual property of traditional knowledge
• Where data release may cause harm

“...the IPY Joint Committee requires that IPY data, including operational data delivered in real time, are made available fully, freely and on the shortest feasible timescale.”
What are the Data

- National Science Board 2005:
  - Reference collections
  - Community or Resource collections
  - Research collections

Fetterer and Knowles. 2004. Sea Ice Index. nsidc.org/data/seaice_index/


Survey on Data Withholding in Academic Genetics

Respondent was denied data from colleagues with published results 47%

Respondent denied data related to their publication 12%

From Campbell, et al. 2002, JAMA
“A biologist would rather share their toothbrush than share their data”
—Carole Goble

ZaCky ➔ http://www.flickr.com/photos/zacky8/
Reasons for not sharing data

From Campbell, et al. 2002, JAMA
Identity—The “Real Polar Man”
Realizing that the Antarctic Treaty (1959) not only established a physical commons, but also an ‘information commons’:

- Article III-1c: “...to the greatest extent feasible and practicable... scientific observations and results from Antarctica shall be exchanged and made freely available”
The Polar Information Commons (PIC)

Realizing that, though the situation in the Arctic is more complex, there is a clear need for information sharing based on shared interests:

- Monitoring and prediction of Arctic climate changes and associated impacts on the global environment
- Opening of sea routes
- Marine pollution
- Fisheries and wildlife management
- Energy extraction
- Sustainable development of northern communities
The scientific community is also increasingly recognizing the importance of open data access

- ICSU Principle of Universality of Science
- Statute 5: “This principle embodies freedom of movement, association, expression and communication for scientists, as well as equitable access to data, information and research materials.”

“Research cannot flourish if data are not preserved and made accessible.” – Editorial, Nature, 2009

http://tinyurl.com/dataspecial
Why?

- Rapid, unprecedented polar change
- Wise management of resources
- Improved decision support
- Effective international cooperation on resource and geopolitical issues
- Deeper understanding and better prediction of change
- Scattered, sparse, and sporadic data
Our Vision

• Data and information are **public goods** that should be shared ethically and with minimal constraint.

• The PIC would be a **shared virtual resource** that parallels the “legal” commons in the Antarctic and the “shared interest” commons in the Arctic

• The PIC would provide both an **institutional framework** and a **technical infrastructure** to encourage sharing and preservation of polar data in the short and long term

• The PIC would harness the **collective expertise, capabilities, and resources** of the broad community of stakeholders with interest in the polar regions
Our vision:
Data are open, linked, useful, and safe.
Positive deviance says that if you want to create change, you must scale it down to the lowest level of granularity and look for people within the social system who are already manifesting the desired future state. Take only the arrows that are already pointing toward the way you want to go, and ignore the others. Identify and differentiate those people who are headed in the right direction. Give them visibility and resources. Bring them together. Aggregate them. *Barbara Waugh*
1. Intellectual foundation for the protocol

The motivation behind this memorandum is interoperability of scientific data.

The volume of scientific data, and the interconnectedness of the systems under study, makes integration of data a necessity. For example, life scientists must integrate data from across biology and chemistry to comprehend disease and discover cures, and climate change scientists must integrate data from wildly diverse disciplines to understand our current state and predict the impact of new policies.

The technical challenge of such integration is significant, although emerging technologies appear to be helping. But the forest of terms and conditions around data make integration difficult to legally perform in many cases. One approach might be to develop and recommend a single license: any data with this license can be integrated with any other data under this license.
calls for data providers to **waive** all rights necessary for **data extraction** and **re-use**

requires provider place **no additional** obligations (like share-alike) to limit downstream use

request behavior (like attribution) through **norms** and **terms of use**
CC Zero waiver + SC norms

waive rights → public domain

+ attribution / citation through community norms, not a contract
PIC Norms for Contributors

• acknowledge that their submitted materials are already in the public domain, or that they have clear rights to make these materials openly accessible through the PIC.

• label their contributions digitally with the “PIC badge,” which specifies how rights have been waived and links back to this statement of norms, and agree to make these contributions accessible and searchable online.

• provide at least the minimum information about their contributed materials requested by the PIC submission interface as accurately and completely as possible.

• will make reasonable efforts to provide additional information about their contributed materials, e.g., to help document the quality of their submitted materials and to ensure their long-term usability.

• make reasonable efforts to provide appropriate notification to the PIC community (e.g., through PIC interfaces) of any significant errors in their contributed materials or descriptions, if any are discovered after submission.
PIC Norms for Users

- **acknowledge the authorship** or source of all materials that they use from the PIC, including use of an appropriate citation when data are used in a formal scientific publication. See example citation guidelines at [http://ipydis.org/data/citations.html](http://ipydis.org/data/citations.html)

- give **appropriate recognition to the PIC** as a digital community resource in their publications and value-added products and services.

- notify the relevant **PIC contributors** (or the PIC community more generally) about their use of specific digital materials from the PIC, and about any suspected significant errors, limitations, or other problems that they may have discovered in the course of their use of those materials.

- acknowledge that they are **primarily responsible** for determining whether the PIC materials they use are of sufficient quality and appropriateness for their objectives.

- **contribute back** to the PIC any value-added data, information, or other digital content derived entirely or largely from PIC materials, with appropriate citation of and documentation about PIC and non-PIC inputs.
Welcome to the Polar Information Commons (PIC):
Establishing the Framework for the Long-term Stewardship of Polar Data and Information

The polar regions are changing rapidly with dramatic global effect. Wise management of resources, improved management of resources, improved decision support, and effective international cooperation on resource and geopolitical issues requires deeper understanding and better prediction of these changes. Understanding and knowledge are built on data and information, yet polar information is scattered, scarce and sporadic. We are scientists and greater than us are themselves. We envision a Polar Information Commons (PIC) as a shared virtual resource mirroring the geographic common interests of humanity. The PIC would serve as an open, virtual repository for vital scientific data and information, and would provide a shared, community-based cyber-infrastructure fostering innovation, improving scientific, and encouraging participation in research, education, planning, and management in the polar regions.

The PIC builds on the legacy of the International Polar Year and we seek active participation and ideas from national governments, international organizations, and the scientific and data management communities at large to build this common resource.
Thank You
parsonsm@nsidc.org