Freeing Data through the Polar Information Commons

CODATA 2010, Spier, South Africa, 25-27 Oct. 2010

Taco de Bruin, NIOZ Royal Netherlands Institute for Sea Research
Robert S. Chen, CIESIN Columbia University and CODATA
Mark R. Parsons, NSIDC, University of Colorado
David J. Carlson, formerly IPY IPO, now at University of Colorado
Kathleen Cass, CODATA
Kim Finney, Australian Antarctic Division
Kaitlin Thaney, formerly Science Commons, now at MacMillan, UK
**John Wilbanks**, Science Commons
Kim Jochum, University of Alaska

Data are the common wealth of humanity — *Adama Samassekou*
Convener of the UN World Summit on the Information Society

The International Polar Year 2007-2008 (IPY)

- IPY was an intensive burst of internationally coordinated, **interdisciplinary**, scientific research and observations, focused on the Polar Regions (Arctic Ocean including adjacent seas and coastal regions, Southern Ocean, Antarctic Continent)

- The IPY Data Policy calls for IPY data to be made available "fully, freely, openly, and on the shortest feasible timescale"

- The IPY Framework document: "In fifty years time the data resulting from IPY 2007-2008 may be seen as the most important single outcome of the programme"

- IPY has resulted in an unprecedented amount of data of the Polar Regions
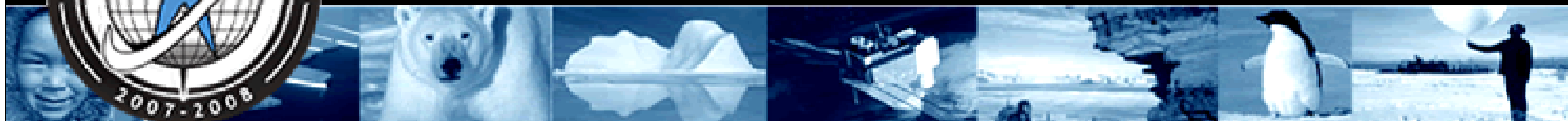
The International Polar Year 2007-2008 (IPY)

- However, access to and long-term preservation of many key IPY datasets remains problematic, especially with respect to Arctic data

- Several disciplines, involved in IPY, have no tradition of data sharing

- Several disciplines have no discipline-based data centres

- As a result, scientists may not know how or where to submit data, in order to make the data available
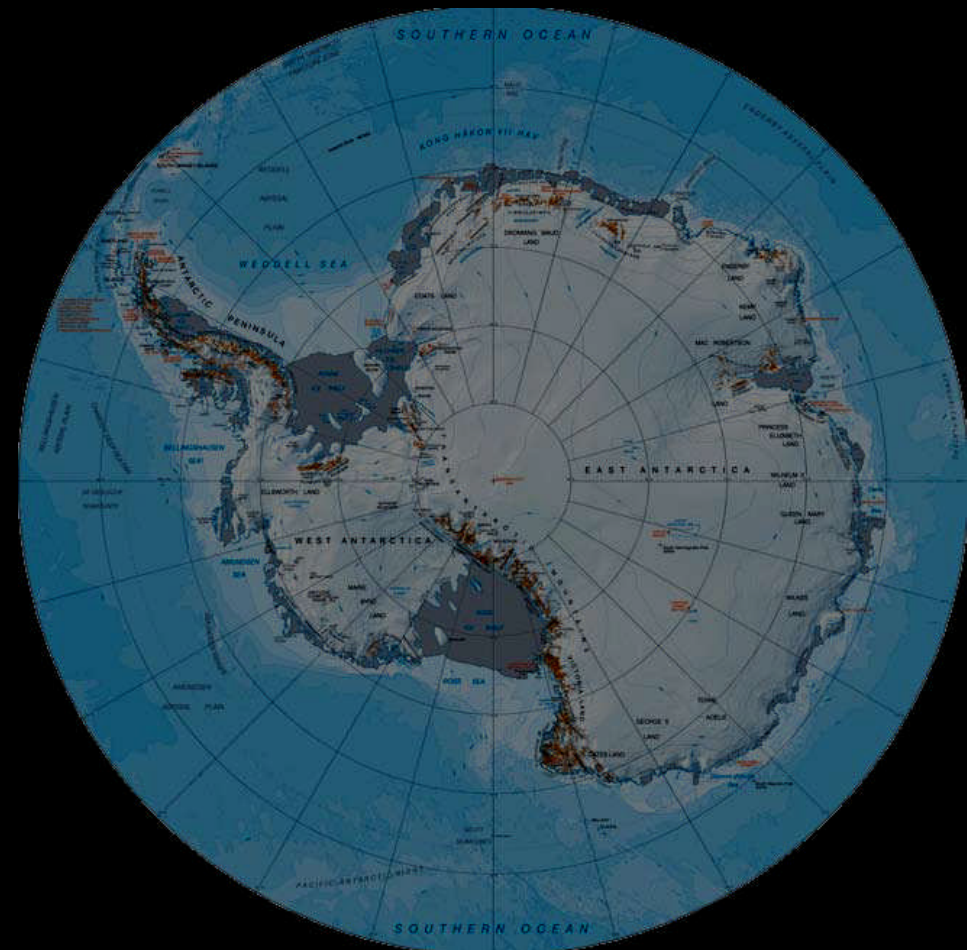
The Polar Information Commons (PIC)

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 indigenous peoples

The Polar Information Commons Vision

Data and information about the Polar regions are public goods that should be shared ethically and with minimal constraint

The PIC is a shared virtual commons that parallels the 'legal' commons in the Antarctic and the 'shared interest' commons in the Arctic

The PIC both provides an institutional framework and a technical infrastructure for sharing and preservation of polar data in the short and long term
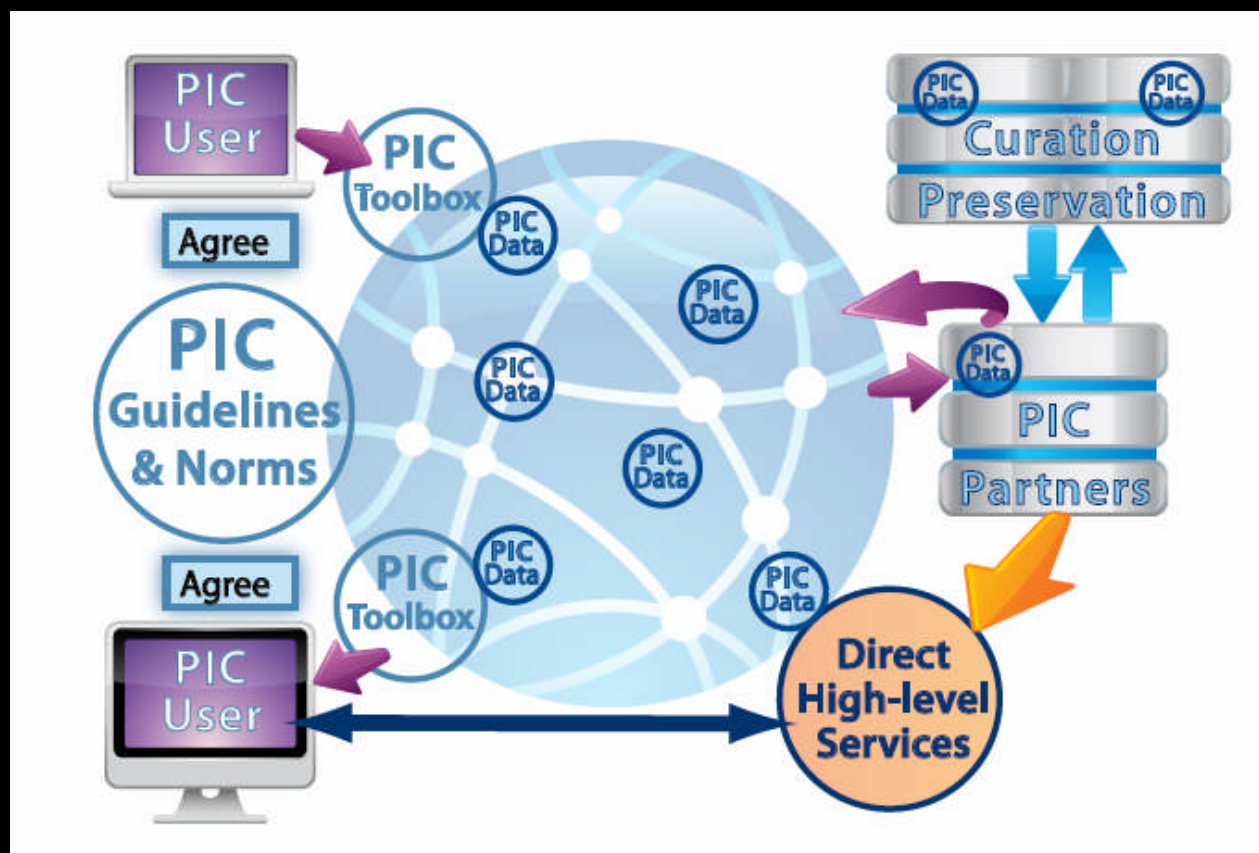
The PIC builds an interdisciplinary community of providers and users of scientific, polar data

How does the PIC work?

Submit data to the PIC cloud

Polar data sources expose their data to the world through the PIC badge and open protocols

Data centers monitor new PIC data and assess and acquire important data for formal archiving, curation, and access

Australian Research Computing Infrastructure (ARCS) is initially hosting the PIC Cloud

How does the PIC work?

Stages and roles

Data providers

- Badge data as belonging to the PIC
- Expect users to comply with PIC norms and behaviour
- Are ensured of long term preservation of their data

PIC Cloud

- Initial host for data, including 'orphan' data
- Data are exposed to enable preservation and use of the data

Data Centres

- Ensure long term preservation of data, by adopting all data
- Add value by providing various high-level services

Data users

- Comply with PIC norms and expected behaviour
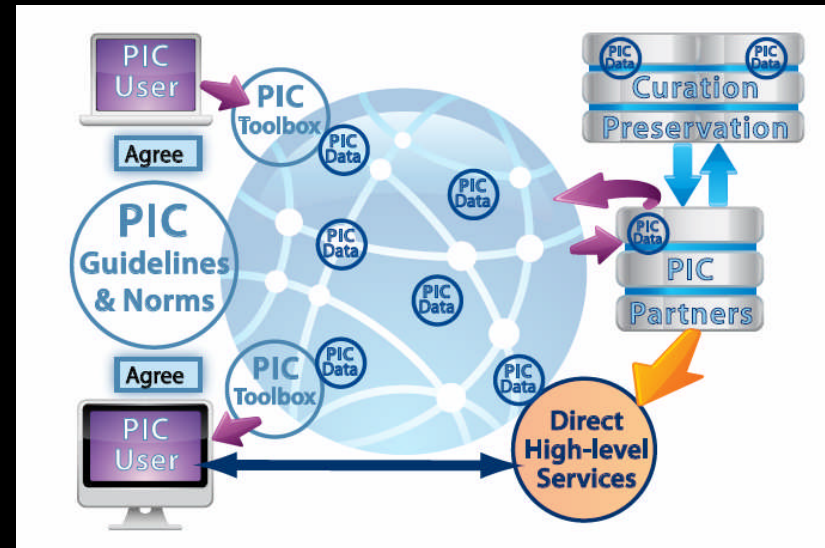
How does the PIC work?

Badging data

- Labels data as belonging to the PIC
    - Norms and expected behavior, both by users and providers
    - Makes rights of users and providers explicit
    - Puts contents as close as (legally) possible to public domain

- Makes (automated) searching for the data possible

- Badging tools incorporated in discovery metadata writing tools

- Logos in metadata link to norms and expected behavior:
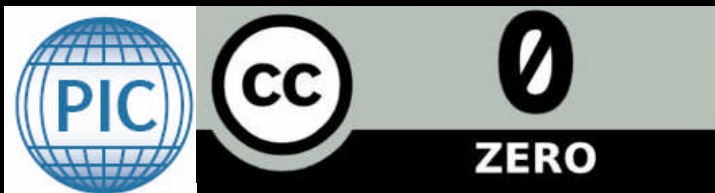
Proposed Norms for PIC Users

- Formal scientific publication citation is desired and PIC users acknowledge authorship and co-authorship of materials that is used from PIC

- PIC users agree that they will also give appropriate recognition to the role of the PIC as a digital community resource

- PIC users agree that they will make reasonable and timely efforts to 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

- PIC users acknowledge that they themselves are responsible for determining whether the PIC materials they use are of sufficient quality and appropriateness for their objectives. However the PIC badge is not a certification of quality

- PIC users agree that in all cases they will 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

Proposed Norms for PIC Data Providers

- PIC 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

- PIC contributors agree to label their contributions digitally with the "PIC badge", which specifies rights of access and links back to this statement of norms, and agree to make these contributions accessible and searchable online

- PIC contributors agree to provide at least the minimum information about their contributed materials requested by the PIC

- PIC contributors agree that, if requested, they 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.

- PIC contributors agree that they will 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.

How does the PIC work?

Data centres or 'data adopters'

- Long term preservation

- World Data System (WDS) requirements for certified data centres

- Continue to keep data accessible under same access rights

- Notify provider that data set is adopted

Who is leading the PIC?

- Committee on Data for Science & Technology (CODATA)
- IPY Data Management Subcommittee (also a CODATA TG)
- International Arctic Science Council (IASC)
- Scientific Committee on Antarctic Research (SCAR)
- International Polar Year International Program Office (IPY IPO)
- World Meteorological Organization (WMO)
- World Data System (WDS) Transition Team
- International Union of Geodesy and Geophysics (IUGG)
- Royal Netherlands Academy of Sciences
- Science Commons
- Association of Polar Early Career Scientists (APECS)
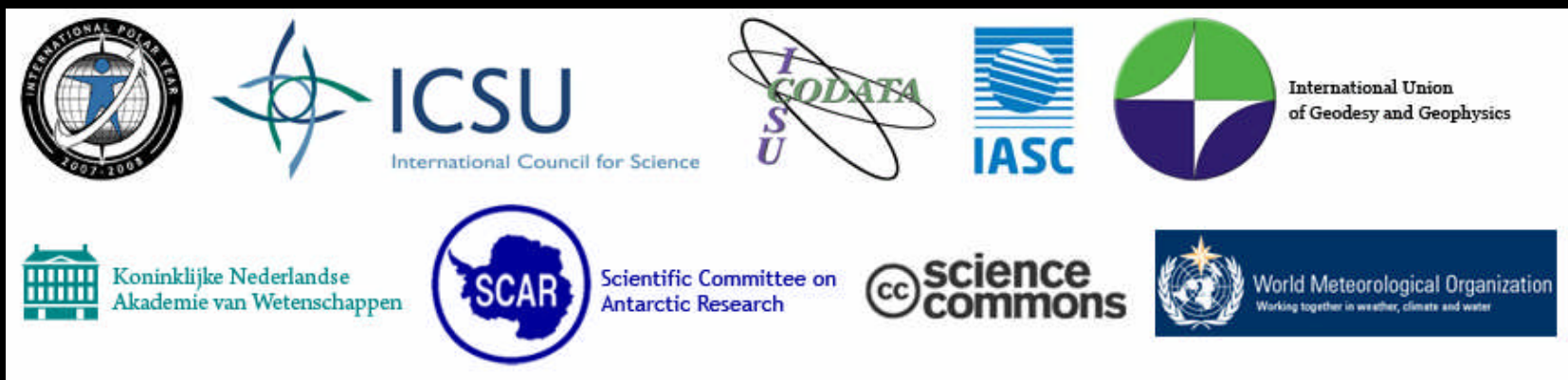
The PIC is operational

Launched at IPY Conference, Oslo, June 2010

The PIC is operational

Data are the common wealth of humanity — *Adama Samassekou*
Convener of the UN World Summit on the Information Society

PIC Home

Overview of PIC Cloud

How to use the Cloud

Submit Dataset

Retrieve Dataset

Contact Us

Dataset Upload Form

Please complete the following information and upload your files

First Name: [   ]

Last Name: [   ]

Phone: [   ]

Email: [   ]

Title: [Dr   ▼]

Country: [Abkhazia   ▼]

Data Set Title: [   ]

Data Set Abstract:

The PIC is operational

The PIC is operational

Data are the common wealth of humanity — *Adama Samassekou*
Convener of the UN World Summit on the Information Society

PIC Home

Overview of PIC Cloud

How to use the Cloud

Submit Dataset

Retrieve Dataset

Contact Us

Dataset Search and Retrieval Form

1. Retrieve all submissions made by a person "x":

First Name:

Last Name:

Search

2. Retrieve all data submitted by people from country "x".

Country: Please select a country

Search

3. Retrieve all data containing the title words "x".

Data Set Title:

Search

The PIC is operational

Data are the common wealth of humanity — *Adama Samassekou*
Convener of the UN World Summit on the Information Society

PIC Home

Overview of PIC Cloud

How to use the Cloud

Submit Dataset

Retrieve Dataset

Contact Us

Dataset Search Results

To view the abstracts for the datasets you have retrieved click on the "abstract" hyperlink. To download dataset packages of interest click on the dataset "download" hyperlink.

Datasets meeting your criteria include:

8. **Temperature-density, temperature-conductivity and conductivity-density relationships for marine-derived saline lake waters**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/BIOSPHERE/AQUATIC ECOSYSTEMS/LAKES/SALINE LAKES/LRjbWOUrWxNMWXtZBSHB/IIZhJUZtpY1275674233053.zip
Dataset Creator: John Gibson   Email: john.gibson(at)utas.edu.au
Click to see Abstract

1. **UV radiation at Davis station summer 1998**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/ATMOSPHERE/ATMOSPHERIC RADIATION/ULTRAVIOLET RADIATION/knASCTUwdCFTkygXbOqT/oQPOGyaJLF1276160986920.zip
Dataset Creator: Andrew Davidson   Email: andrew.davidson(at)aad.gov.au
Click to see Abstract

15. **: Ice sheet topography and surface characteristics in eastern Wilkes Land, East Antarctica**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/CRYOSPHERE/GLACIERS_ICE SHEETS/fPUSnjGxmnZCXerYxLPI/VMPMZtrYcU1275630822771.zip
Dataset Creator: Ian Goodwin   Email: ian.goodwin(at)newcastle.edu.au
Click to see Abstract

5. **Impact of changes in UV and visible radiation on the reflective properties of plant photosynthetic surfaces**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/BIOLOGICAL CLASSIFICATION/PLANTS/dpAvyFxCMLrsJoUNUbWO/GmIVNcJrWD1275676237036.zip
Dataset Creator: Sharon Robinson   Email: sharonr(at)uow.edu.au
Click to see Abstract

2. **Macquarie Island Cetacean Sightings and Strandings**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/BIOLOGICAL CLASSIFICATION/ANIMALS_VERTEBRATES/FJcETUFykFBMcBoYKiUN/MkytPVHeNI1275622204517.zip
Dataset Creator: Geoff Copson   Email: geoffco(at)dpiwe.tas.gov.au
Click to see Abstract

17. **data for Thala Valley Clean Up Project 2003V4**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/HUMAN DIMENSIONS/ENVIRONMENTAL IMPACTS/CONTAMINANTS/iRgNhauofmNriPGyFnKJ/gutPTuAbas1276088958660.zip
Dataset Creator: Chris Paterson   Email: chris.paterson(at)aad.gov.au
Click to see Abstract

3. **Ducks and Mallards of Macquarie Island**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/BIOLOGICAL CLASSIFICATION/ANIMALS_VERTEBRATES/jbQadpzpLBvrOYhoqCfW/yzxsFQAEHi1275622007568.zip
Dataset Creator: Knowles Kerry   Email: liz.kerry(at)keypoint.com.au
Click to see Abstract

18. **Modification of the type of dietary fat at an Antarctic station: impact on cardiovascular risk factors**   Download: https://df.arcs.org.au/ARCS/projects/PICCLOUD/data/EARTH SCIENCE/HUMAN

Can you help?

- Inputs on PIC vision, design, community norms (see PIC web site)

- Contributions to key PIC elements

- Contributions of important polar data and information resources

- Outreach to the broader community

Thanks!

http://www.polarcommons.org