

# PROGRAMME and ABSTRACTS

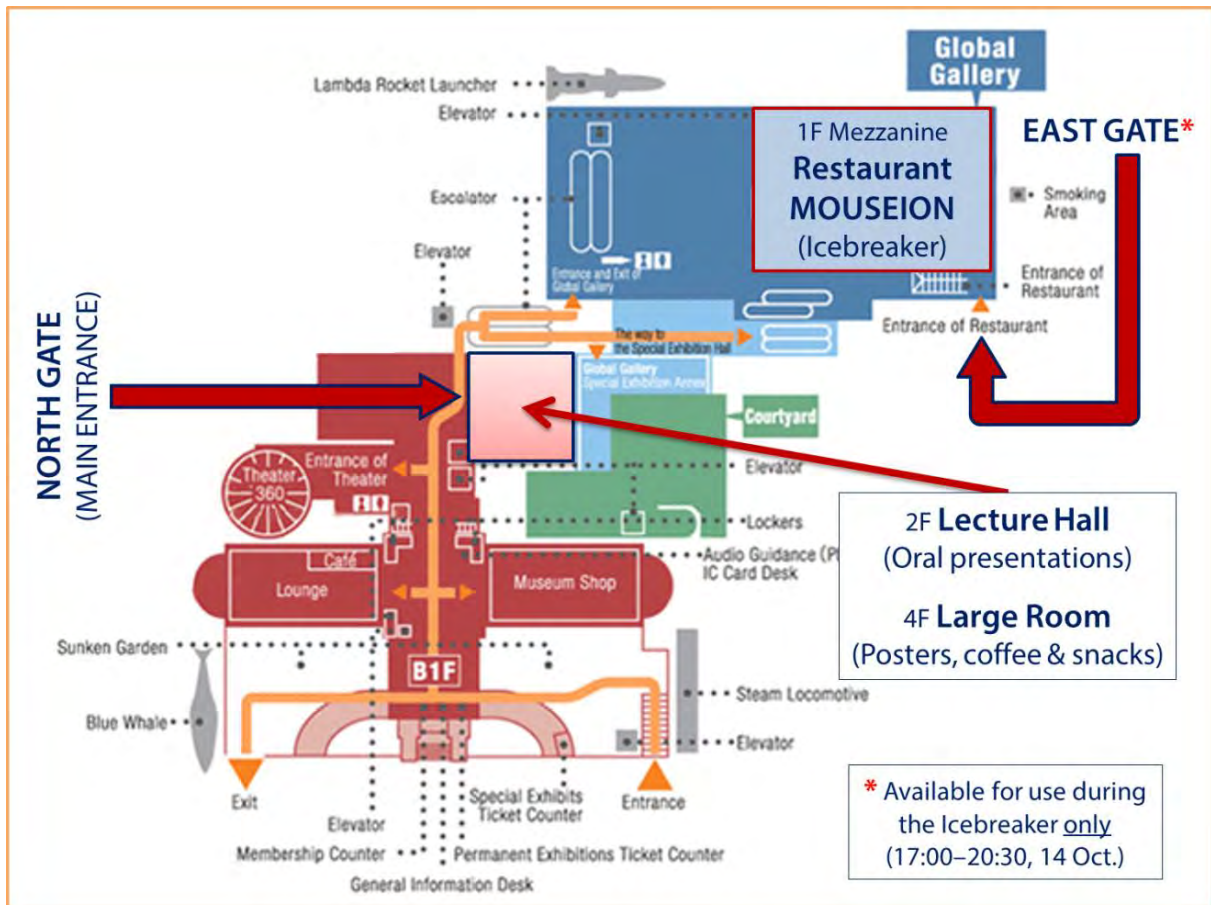
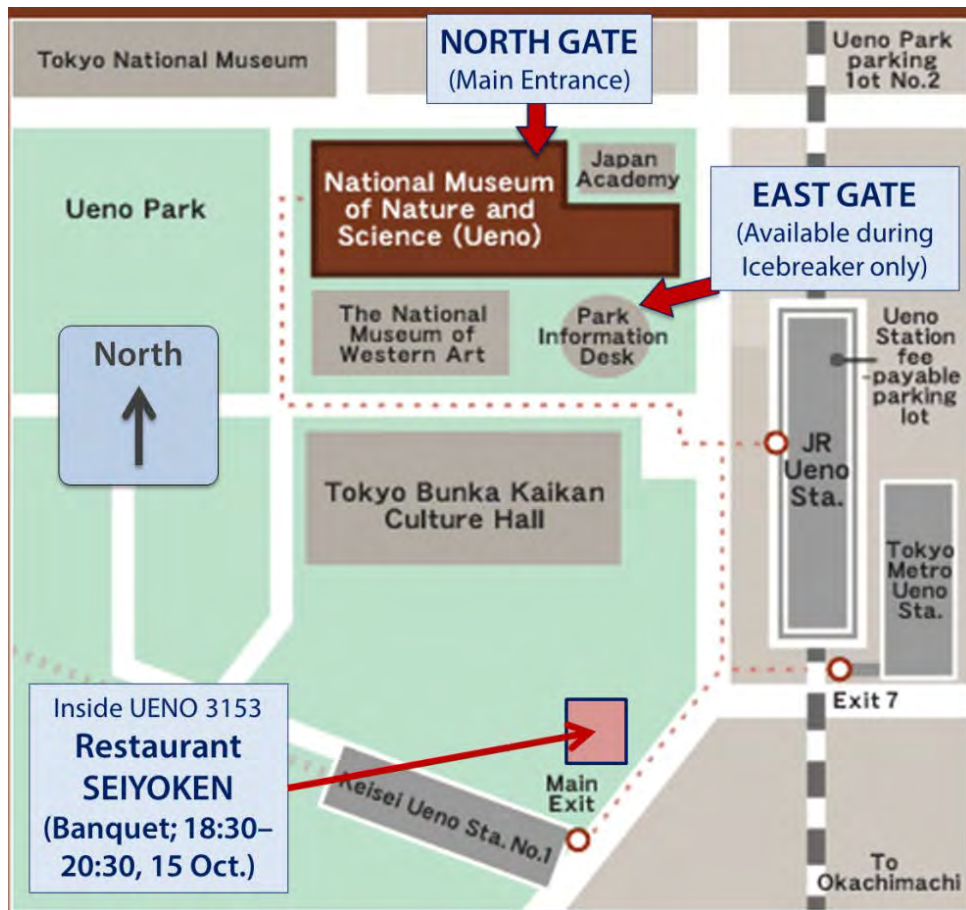
International Forum on

## Polar Data Activities in Global Data Systems



15–16 October 2013  
Tokyo Japan

<http://www.polar-data-forum.org/>



# Addenda

## Contents

Changes to the Forum Programme		A3
Full Programme (Revised)		A4
Poster Presentations (Revised)		A7
Antarctic Data Treasures: Recent Advances in Promoting and Curating the Vital Data Legacy of Antarctic Data	Taco de Bruin	A9



## Changes to the Forum Programme

---

(See the revised Programme and list of Poster Presentations for more information)

- Session 2: Best Practice – Data Administration (14:00–15:30 on Tuesday, 15 October)
  - Dongmin Jin was unable to attend the Forum. In his absence, the presentation ‘Establishing Korean Polar Data Policy and its Future Directions’ was given by Min Choul Lee of the Korea Polar Research Institute.
  - Rocío Zorilla was also unable to attend the Forum, and thus could not present ‘Design and Implementation of the Brazilian Information System on Antarctic Environmental Research’. Instead, Taco de Bruin presented ‘Antarctic Data Treasures: Recent Advances in Promoting and Curating the Vital Data Legacy of Antarctic Data’. The extended abstract for that presentation is given on Page A?.
  
- Wednesday, 16 October

Owing to a typhoon striking the Tokyo area on the morning of 16 October, the start of the second day of the Forum was delayed until 11:00. With the closing time of the Forum venue—the National Museum of Nature and Science—fixed at 19:00, the Forum schedule could not be extended. The remainder of the programme was therefore condensed, accordingly. In particular,

  - The keynote presentations and panel discussions in Session 5: Data Centre/Service Provider Accreditation Best Practice – Data Administration and Session 6: The Future of Scientific Polar Data Management were reduced from 30 min to 20 min in length.
  - The afternoon coffee break was limited to only 10 min, and the associated third-and-final Poster Session cancelled.

In addition, Ryosuke Shibasaki was unavailable to Chair Session 4-B: Best Practice – Data Sharing & Observing Networks and was replaced by Vasily Kopylov.
  
- Poster Presentations

Both Michael Nisilevich and Louise Newman were not able to participate in the Forum. As a result, P6 ‘Web-based Technologies in Data/Information Management for Polar Data in Russian World Data Centres’ was not displayed, and Kim Finney replaced Louise Newman as the presenter of P15 ‘Data Management Plans for the Southern Ocean Observing System (SOOS)’.

## Full Programme (REVISED)

<b>Monday 14 October</b>	
17:00–18:00	<b>Pre-registration (badges only)</b>
18:00–20:00	<b>Icebreaker</b>
<b>Tuesday 15 October</b>	
09:00–10:00	<b>Registration</b>
10:00–10:45	<b>Opening Remarks</b>
	<ul style="list-style-type: none"><li>• <b>Kazuyuki Shiraishi</b>, Director-General, National Institute of Polar Research (5')</li><li>• <b>Mizuhiko Hosokawa</b>, Senior Executive Director, National Institute of Information and Communications Technology (5')</li><li>• <b>Yoshihiro Hayashi</b>, Director-General, National Museum of Nature and Science (5')</li><li>• <b>Jerónimo López-Martínez</b>, President, Scientific Committee on Antarctic Research (8')</li><li>• <b>David Hik</b>, President, International Arctic Science Committee (8')</li><li>• <b>Bernard Minster</b>, Chair, ICSU World Data System (8')</li></ul>
10:45–11:00	<b>Scope of the Forum &amp; Practical Information</b>
	<ul style="list-style-type: none"><li>• <b>Kim Finney</b>, Chair of the Scientific Organizing Committee (10')</li><li>• <b>Masaki Kanao</b>, Chair of the Local Organizing Committee (5')</li></ul>
11:00–12:30	<b>Session 1: Data – Lessons Learned from IPY</b> (Session chair: Taco de Bruin) <b>Keynote:</b> <i>Overarching Lessons from the International Polar Year on How to Create a Functional Data Infrastructure.</i> <b>Mark Parsons</b> (30')
	<ul style="list-style-type: none"><li>• <i>Arctic Communities, Social Science, and Data Management: Human dimensions of an Arctic Data Coordination Network.</i> <b>Peter Pulsifer</b> (15')</li><li>• <i>The Polar Data Catalogue: Best Practices for Sharing and Archiving Canada's Polar Data.</i> <b>Julie Friddell</b> (15')</li><li>• <i>Preservation and Publication of IPY data – A Collaborative Effort of PANGAEA, ICSTII, and ICSU-WDS.</i> <b>Michael Diepenbroek</b> (15')</li></ul>
12:30–14:00	<b>Lunch and Poster Session</b> (See list of posters for details)
14:00–15:30	<b>Session 2: Best Practice – Data Administration</b> (Session chair: Wim Hugo) <b>Keynote:</b> <i>Administering Antarctic Data – A Practical Use-case.</i> <b>Kim Finney</b> (30')
	<ul style="list-style-type: none"><li>• <i>Establishing Korean Polar Data Policy and its Future Directions.</i> <b>Min Choul Lee</b> (15')</li><li>• <i>Antarctic Data Treasures: Recent Advances in Promoting and Curating the Vital Data Legacy of Antarctic Data.</i> <b>Taco de Bruin</b> (15')</li><li>• <i>The Arctic Cooperative Data and Information System: Data Management Support for the NSF Arctic Research Program.</i> <b>James Moore</b> (15')</li></ul>

15:30–15:45 <b>Coffee Break</b>
15:45–17:45 <b>Session 3: Best Practice – Publication &amp; Citation</b> (Session co-chairs: Mark Parsons & Toshihiko Iyemori) <b>Keynote:</b> <i>Best Practice—Data Publication and Citation.</i> <b>Karen Visser</b> (30') <ul style="list-style-type: none"> <li>• <i>From Data to Publications: The Polar Information Spectrum.</i> <b>Shannon Vossepoel</b> (15')</li> <li>• <i>Inter-university Upper Atmosphere Global Observation Network (IUGONET) Metadata Database and Possible Application for Polar Data Activities.</i> <b>Akiyo Yatagai</b> (15')</li> <li>• <i>Biodiversity.aq: Online Tools for Antarctic Biodiversity Data Discovery and Publication.</i> <b>Anton Van de Putte</b> (15')</li> </ul> <b>Panel discussion</b> (30') <ul style="list-style-type: none"> <li>• <b>Michael Diepenbroek</b> (PANGAEA, MARUM Centre for Marine Environmental Sciences, Bremen University, ICSU World Data System)</li> <li>• <b>Shuichi Iwata</b> (Data Science Journal, ICSU-CODATA)</li> <li>• <b>Karen Visser</b> (Australian National Data Service)</li> </ul> 17:45–18:00 <b>Group Photograph</b> 18:30–20:30 <b>Forum Banquet</b>
<b>Wednesday 16 October</b>
11:00–12:15 <b>Session 4-A: Best Practice – Data Sharing &amp; Observing Networks</b> (Session chair: Akira Kadokura) <b>Keynote:</b> <i>Data Sharing and Observing Networks—How can we do better?</i> <b>Lesley Rickards</b> (20') <ul style="list-style-type: none"> <li>• <i>Antarctic Space Weather Data Managed by IPS Radio and Space Services of Australia.</i> <b>Kehe Wang</b> (15')</li> <li>• <i>TRANSMIT Prototype: Cross-institutional Network Approach from Geophysical Database to User Application for GNSS Science and Industry.</i> <b>Hiroatsu Sato</b> (15')</li> <li>• <i>Hydrometeorological Database (HMDB) for Practical Research in Ecology.</i> <b>Alexander Novakovskiy</b> (15')</li> </ul> 12:15–13:30 <b>Lunch and Poster Session</b> (See list of posters for details)
13:30–15:00 <b>Session 4-B: Best Practice – Data Sharing &amp; Observing Networks</b> (Session chair: Vasily Kopylov) <b>Keynote:</b> <i>Current Data Practises in Polar Institutions and Networks: A Case Study with the HIACMS Project.</i> <b>Paul Berkman</b> (20') <ul style="list-style-type: none"> <li>• <i>EXPEDITION: An Integrated Approach to Expose Expedition Information and Research Results.</i> <b>Roland Koppe</b> (15')</li> <li>• <i>Assembling an Arctic Ocean Boundary Monitoring Array.</i> <b>Takamasa Tsubouchi</b> (15')</li> <li>• <i>Building on the IPY: Discovering Interdisciplinary Data through Federated Search.</i> <b>Siri Jodha Singh Khalsa</b> (15')</li> <li>• <i>GEO Cold Regions—The Interface To GEOSS For Polar And Mountainous Cold Region Observations.</i> <b>Yubao Qui</b> (15')</li> </ul> 15:00–15:10 <b>Coffee Break (Poster Removal)</b>

15:10–16:05 **Session 5: Data Centre/Service Provider Accreditation** (Session chair: Lesley Rickards)

**Keynote:** *Certification and Accreditation of Data Centres and Services*. **Michael Diepenbroek** (20')

- *ICPSR and the Data Seal of Approval: Accreditation Experiences and Opportunities*. **Jared Lyle** (15')

**Panel discussion** (20')

- **Michael Diepenbroek** (PANGAEA, MARUM Center for Marine Environmental Sciences, Bremen University. ICSU World Data System)
- **Toshihiko Iyemori** (World Data Centre for Geomagnetism)
- **Akira Kadokura** (National Institute of Polar Research)
- **Jared Lyle** (Interuniversity Consortium for Political and Social Research. Data Seal of Approval)

16:05–17:15 **Session 6: The Future of Scientific Polar Data Management** (Session co-chairs: Kim Finney & David Hik)

**Keynote:** *International Polar Data Management*. **Mustapha Mokrane** (20')

- *Towards an International Polar Data Coordination Network: An Arctic Perspective*. **Peter Pulsifer** (15')
- *Creating Web of Data for Science*. **Hideaki Takeda** (15')

**Panel discussion** (20')

- **Paul Berkman** (Bren School of Environmental Science and Management, University of California, Santa Barbara)
- **Taco de Bruin** (Royal Netherlands Institute for Sea Research)
- **Mustapha Mokrane** (ICSU World Data System)
- **Mark Parsons** (Research Data Alliance)
- **Hideaki Takeda** (National Institute of Informatics, Japan)

17:15–17:45 **Session 7: Polar Forum Communique**

**Facilitator:** **Rorie Edmunds** (ICSU World Data System)

17:45–18:30 **ICSU World Data System Town Hall Meeting**



## Poster Presentations (REVISED)

<b>P1</b>	<i>Hornsund GLACIO-TOPOCLIM Database—Polish Polar Station IPY Legacy</i>	<b>Bartłomiej Luks et al.</b>
<b>P2</b>	<i>Harmonizing Polar Biodiversity Data for Wider Access and Integration: A Collaboration between the Spanish Polar Data Center and GBIF-Spain</i>	<b>Oscar Bermúdez et al.</b>
<b>P3</b>	<i>'Quantarctica': New Standalone GIS Package for Antarctic Research, Operation, and Education using Open-source Software</i>	<b>Kenichi Matsuoka et. al. (Presenting author: Stein Tronstad)</b>
<b>P4</b>	<i>Polar Data Tools at the Integrated Earth Data Applications (IEDA)</i>	<b>Frank Nitsche et al. (Presenting author: Robert Arko)</b>
<b>P5</b>	<i>A Global Environmental Database Project at the National Institute for Environmental Studies and its Contribution to DIAS/GRENE</i>	<b>Hideaki Nakajima et al.</b>
<b>P7</b>	<i>Polar Magnetic Data at WDC Kyoto—Services under International Collaborations</i>	<b>Toshihiko Iyemori et al.</b>
<b>P8</b>	<i>Cosmic-ray Neutron Data held by WDC for Cosmic Rays</i>	<b>Takashi Watanabe et al.</b>
<b>P9</b>	<i>Observations and Data Handling at International Center for Space Weather Science and Education, Kyushu University, Japan .</i>	<b>Hideaki Kawano et al.</b>
<b>P10</b>	<i>Data and Science Lessons Learned from Atmospheric/Environmental Observations in Japan and Alaska</i>	<b>Yasuhiro Murayama et al.</b>
<b>P11</b>	<i>Japanese Contribution to Super Dual Auroral Radar Network (SuperDARN)</i>	<b>Nozumu Nishitani et al.</b>
<b>P12</b>	<i>Continuous Broadband Seismic Observation on the Greenland Ice Sheet under Greenland Ice Sheet Monitoring Network</i>	<b>Seiji Tsuboi et al.</b>
<b>P13</b>	<i>Current Status of Science Data Archives for the Data Obtained by the Japanese Antarctic Research Expedition</i>	<b>Akira Kadokura et al.</b>
<b>P14</b>	<i>Metadata Management at the Polar Data Center (PDC) of the National Institute of Polar Research (NIPR), Japan</i>	<b>Masaki Kanao et al.</b>

<b>P15</b>	<i>Data Management Plans for the Southern Ocean Observing System (SOOS)</i>	<b>Louise Newman et al. (Presenting author: Kim Finney)</b>
<b>P16</b>	<i>NICT Science Cloud: Distributed Storage System and Parallel Data Processing Applicable for Polar Research Data</i>	<b>Ken Murata et al.</b>
<b>P17</b>	<i>Operation of Data Acquisition, Transfer, and Storage System for Worldwide Observation Networks</i>	<b>Tsutomu Nagatsuma et al.</b>
<b>P18</b>	<i>IUGONET Data Analysis Software (UDAS) for Upper Atmosphere Study</i>	<b>Yoshimasa Tanaka et al.</b>
<b>P19</b>	<i>Development of a Sensor Observation Service (SOS) Javascript Library</i>	<b>Alex Tate and Paul Breen (Presenting author: Alex Tate)</b>
<b>P20</b>	<i>Outline of Arctic Data Archive System (ADS)</i>	<b>Hironori Yabuki and Takeshi Sugimura (Presenting author: Hironori Yabuki)</b>
<b>P21</b>	<i>The ICSU World Data System: Trusted Data Services for Global Science</i>	<b>Rorie Edmunds</b>

# Antarctic Data Treasures: Recent Advances in Promoting and Curating the Vital Data Legacy of Antarctic Data

**Taco F. de Bruin**<sup>1\*</sup>

<sup>1\*</sup> NIOZ Royal Netherlands Institute for Sea Research, Den Burg, The Netherlands  
Email: Taco.de.Bruin@nioz.nl

**Summary.** The work of the Scientific Committee on Antarctic Research's (SCAR's) Standing Committee on Antarctic Data Management (SC-ADM) centres on implementation of the SCAR Data and Information Strategy (DIMS) to develop the Antarctic Data Management System. This presentation will give a short introduction to SC-ADM, before focussing on DIMS Implementation Plan activities and outcomes of the recent SC-ADM meeting. The presentation will then conclude by providing a future vision for polar data.

**Keywords.** Scientific Committee on Antarctic Research, Standing Committee on Antarctic Data Management, Antarctic Data Management System, Data and Information Strategy.

## 1. Introduction

Preservation of, and free access to, Antarctic data is key to the further development of Antarctic science. The Standing Committee on Antarctic Data Management (SC-ADM) of the Scientific Committee on Antarctic Research (SCAR) coordinates the development of the Antarctic Data Management System (ADMS), which aims to provide centralized access to a repository system of distributed sources of Antarctic data.

The focus of the work of SC-ADM is on the implementation of the SCAR Data and Information Strategy (DIMS) and further development of the ADMS as defined in the Strategy document.

## 2. SC-ADM Activities

The presentation will give a short introduction of SC-ADM—consisting of representatives of SCAR's

National Antarctic Data Centres or national points of contact—and of its main activities.

It will then focus specifically on the DIMS Implementation Plan activities and on the outcomes of the recent SC-ADM meeting, such as the development of a best practices guidance document for the DIF metadata format, improvement of Antarctic data management communications; data management for the Southern Ocean Observing System; and the links with the SCAR science community, the World Data System, the International Arctic Science Committee, and the CODATA-initiated Polar Information Commons.

## 3. Conclusions

Towards the end, the presentation will provide a vision of future sharing and preservation of Antarctic data and polar science data in general.



**Proceedings of the  
International Forum on  
'Polar Data Activities in  
Global Data Systems'**

**15–16 October 2013  
National Museum of Nature and Science  
Tokyo, Japan**

**Edited and Published by  
WDS International Programme Office**

In Memory of Dr. Evgeny Kharin  
Director of the World Data Centre for Solar–Terrestrial Physics  
(6 February 1933 – 11 September 2013)

# Contents

Polar Data Activities in Global Data Systems	Kim Finney	ix
<b>Forum Programme</b>		
Sessional Themes		x
Programme Overview		x
Full Programme		xi
Poster Presentations		xiv
<b>Side Meetings/Events</b>		
SC-ADM-17 (Closed meeting)		xvi
Icebreaker		xvi
Banquet (& Group Photograph)		xvi
WDS Town Hall Meeting		xvi
WDS-SC#9 (Closed meeting)		xvi
<b>Data – Lessons Learned from IPY</b>		
Overarching Lessons from the International Polar Year on How to Create a Functional Data Infrastructure	Mark Parsons	2
Arctic Communities, Social Science, and Data Management: Human dimensions of an Arctic Data Coordination Network	Peter Pulsifer	4
The Polar Data Catalogue: Best Practices for Sharing and Archiving Canada’s Polar Data.	Julie Friddell	6
Preservation and Publication of IPY data – A Collaborative Effort of PANGAEA, ICSTII, and ICSU-WDS	Michael Diepenbroek	9
<b>Best Practice – Data Administration</b>		
Administering Antarctic Data – A Practical Use-case	Kim Finney	12
Establishing Korean Polar Data Policy and its Future Directions	Dongmin Jin	15
Design and Implementation of the Brazilian Information System on Antarctic Environmental Research	Rocío Zorilla	17
The Arctic Cooperative Data and Information System: Data Management Support for the NSF Arctic Research Program	James Moore	19
<b>Best Practice – Publication &amp; Citation</b>		
Best Practice—Data Publication and Citation	Karen Visser	24
From Data to Publications: The Polar Information Spectrum	Shannon Vossepoel	26
Inter-university Upper Atmosphere Global Observation Network (IUGONET) Metadata Database and Possible Application for ‘Polar Data Activities’	Akiyo Yatagai	28
Biodiversity.aq: Online Tools for Antarctic Biodiversity Data Discovery and Publication	Anton Van de Putte	31

## **Best Practice – Data Sharing & Observing Networks (A)**

Data Sharing and Observing Networks—How can we do better?	Lesley Rickards	34
Antarctic Space Weather Data Managed by IPS Radio and Space Services of Australia	Kehe Wang	36
TRANSMIT Prototype: Cross-institutional Network Approach from Geophysical Database to User Application for GNSS Science and Industry	Hiroatsu Sato	38
Hydrometeorological Database (HMDB) for Practical Research in Ecology	Alexander Novakovskiy	40

## **Best Practice – Data Sharing & Observing Networks (B)**

Current Data Practises in Polar Institutions and Networks: A Case Study with the HIACMS Project	Paul Berkman	44
EXPEDITION: An Integrated Approach to Expose Expedition Information and Research Results	Roland Koppe	49
Assembling an Arctic Ocean Boundary Monitoring Array	Takamasa Tsubouchi	51
Building on the IPY: Discovering Interdisciplinary Data through Federated Search	Siri Jodha Singh Khalsa	53
GEO Cold Regions—The Interface To GEOSS For Polar And Mountainous Cold Region Observations	Yubao Qui	56

## **Data Centre/Service Provider Accreditation**

Certification and Accreditation of Data Centres and Services	Michael Diepenbroek	60
ICPSR and the Data Seal of Approval: Accreditation Experiences and Opportunities	Jared Lyle	62

## **The Future of Scientific Polar Data Management**

International Polar Data Management	Mustapha Mokrane	66
Towards an International Polar Data Coordination Network: An Arctic Perspective	Peter Pulsifer	69
Creating Web of Data for Science	Hideaki Takeda	71



## Poster Presentations

Hornsund GLACIO–TOPOCLIM Database—Polish Polar Station IPY Legacy	Bartlomiej Luks et al.	76
Harmonizing Polar Biodiversity Data for Wider Access and Integration: A Collaboration between the Spanish Polar Data Center and GBIF-Spain	Oscar Bermúdez et al.	78
‘Quantarctica’: New Standalone GIS Package for Antarctic Research, Operation, and Education using Open-source Software	Kenichi Matsuoka et al. (Presenting author: Stein Tronstad)	80
Polar Data Tools at Integrated Earth Data Applications (IEDA)	Frank Nitsche et al. (Presenting author: Robert Arko)	82
A Global Environmental Database Project at the National Institute for Environmental Studies and its Contribution to DIAS/GRENE	Hideaki Nakajima et al.	84
Web-based Technologies in Data/Information Management for Polar Data in Russian World Data Centres	Evgeny Kahrin et al. (Presenting author: Michael Nisilevich)	87
Polar Magnetic Data at WDC Kyoto—Services under International Collaborations	Toshihiko Iyemori et al.	89
Cosmic-ray Neutron Data held by WDC for Cosmic Rays	Takashi Watanabe et al.	91
Observations and Data Handling at International Center for Space Weather Science and Education, Kyushu University, Japan	Hideaki Kawano et al.	94
Data and Science Lessons Learned from Atmospheric/Environmental Observations in Japan and Alaska	Yasuhiro Murayama et al.	96
Japanese Contribution to Super Dual Auroral Radar Network (SuperDARN)	Nozumu Nishitani et al.	98
Continuous Broadband Seismic Observation on the Greenland Ice Sheet under Greenland Ice Sheet Monitoring Network	Seiji Tsuboi et al.	100
Current Status of Science Data Archives for the Data Obtained by the Japanese Antarctic Research Expedition	Akira Kadokura et al.	102
Metadata Management at the Polar Data Center (PDC) of the National Institute of Polar Research (NIPR), Japan	Masaki Kanao et al.	104
Data Management Plans for the Southern Ocean Observing System (SOOS)	Louise Newman et al.	106
NICT Science Cloud: Distributed Storage System and Parallel Data Processing Applicable for Polar Research Data	Ken Murata et al.	108
Operation of Data Acquisition, Transfer, and Storage System for Worldwide Observation Networks	Tsutomu Nagatsuma et al.	110
IUGONET Data Analysis Software (UDAS) for Upper Atmosphere Study	Yoshimasa Tanaka et al.	112
Development of a Sensor Observation Service (SOS) Javascript Library	Alex Tate and Paul Breen (Presenting author: Alex Tate)	114
Outline of Arctic Data Archive System (ADS)	Hironori Yabuki and Takeshi Sugimura (Presenting author: Hironori Yabuki)	116
The ICSU World Data System: Trusted Data Services for Global Science	Rorie Edmunds	118



## Polar Data Activities in Global Data Systems

---

From March 2007 to March 2009, scientists from around the globe participated in the International Polar Year (IPY). This was the fourth such event in history, with the first Polar Year being declared in 1882–1883. This most recent Polar Year generated research from greater than 60 countries in the Physical, Biological, and Social Science disciplines and as a result, significantly added to our understanding of the polar regions and their role in Earth systems. Just as importantly, these activities shone a light on how we currently conduct data management in support of large multidisciplinary, cross-jurisdictional science projects. Given such rapid recent advances in technology, the quantity and variety of data being generated by polar research are unprecedented. Subsequently, many of the local-in-scale, fragmented, and independent data systems that were capable of supporting the first to third IPY were found to be far less suitable for underpinning global research in the 21st Century.

IPY 2007–2008 has therefore catalyzed a number of global science coordinating entities into action in order to address identified weaknesses in current polar data management and data systems. In addition to unilateral actions, the International Council for Science bodies of SCAR (the Scientific Committee on Antarctic Research) and WDS (the World Data System) have joined with IASC (the International Arctic Science Committee) to convene an International Forum on ‘Polar Data Activities in Global Data Systems’. Hosted by the Japanese research institutions of NICT (National Institute of Information and Communication Technology) and NIPR (National Institute of Polar Research), this Forum brings together data professionals and research scientists to explore the data-centric lessons that have been learnt from undertaking the fourth IPY. Forum participants will share experiences of current best practice in fields such as data administration, global data publication and citation, international data repository accreditation, and large-scale data sharing and observing networks. These best practices derive from scientific experience in general; not just that confined to the polar community. At the Forum’s conclusion, participants will discuss what the future holds for scientific polar data management given continual technological advances, the emerging era of Big Data and the globalization of science.

On behalf of the Scientific Organizing Committee,

Kim Finney, Chair  
(01/10/2013)

# Forum Programme

---

## Sessional Themes

- **Data – Lessons learned from IPY:** IPY Data Practices & Data legacy, Polar Information Commons
- **Best Practice – Data Administration:** Metadata, Vocabularies, Ontologies, Data Management Planning, Data Policy, Repository Practices & Standards
- **Best Practice – Publication & Citation:** Data Publication, Data Citation, Scientific Reward and Recognition Systems
- **Best Practice – Data Sharing & Observing Networks:** Virtual Observatories, Information and Communications Technology Infrastructure Protocols and Architectures, Sustainability and Governance Models, Cloud Computing and Storage, Real-time data handling
- **Data Centre/Service Provider Accreditation:** Current Accreditation Schemes and their Benefits
- **The Future of Scientific Polar Data Management:** Any theme as long as the presentation is strategically forward looking

## Programme Overview

Date	Morning (Approx. 09:00–12:30)	Lunch (12:30–14:00)	Afternoon (Approx. 14:00–17:00)	Evening (17:00–)
Saturday Oct. 12				SC-ADM Icebreaker
Sunday Oct. 13	SC-ADM-17		SC-ADM-17	SC-ADM Dinner
Monday Oct. 14	SC-ADM-17		SC-ADM-17	Polar Data Forum Pre-registration & Icebreaker
Tuesday Oct. 15	Polar Data Forum Registration, Opening Remarks & Scope. Session 1: Data – Lessons Learned from IPY	Poster session	Session 2: Best Practice – Data Administration Session 3: Best Practice – Publication & Citation	Group Photograph & Forum Banquet
Wednesday Oct. 16	Sessions 4-A & 4-B: Best Practice – Data Sharing & Observing Networks	Poster session	Session 5: Data Centre/Service Provider Accreditation Session 6: Future of Scientific Polar Data Management Session 7: Polar Forum Communique	WDS-SC Town Hall Meeting
Thursday Oct. 17	WDS-SC#9 (NICT Office, Kojimachi)		WDS-SC#9 (NICT Office, Kojimachi)	WDS-SC Banquet
	Excursion for SC-ADM members: NIPR in Tachikawa City (optional)			
Friday Oct. 18	WDS-SC#9 (NICT Office, Kojimachi)		WDS-SC#9 (NICT Office, Kojimachi)	
	Excursion for SC-ADM members: NICT in Koganei City (optional)			

## Full Programme

<b>Monday 14 October</b>	
17:00–18:00	Pre-registration (badges only)
18:00–20:00	Icebreaker
<b>Tuesday 15 October</b>	
09:00–10:00	Registration
10:00–10:45	Opening Remarks <ul style="list-style-type: none"> <li>• <b>Kazuyuki Shiraishi</b>, Director-General, National Institute of Polar Research (5')</li> <li>• <b>Mizuhiko Hosokawa</b>, Senior Executive Director, National Institute of Information and Communications Technology (5')</li> <li>• <b>Yoshihiro Hayashi</b>, Director-General, National Museum of Nature and Science (5')</li> <li>• <b>Jerónimo López-Martínez</b>, President, Scientific Committee on Antarctic Research (8')</li> <li>• <b>David Hik</b>, President, International Arctic Science Committee (8')</li> <li>• <b>Bernard Minster</b>, Chair, ICSU World Data System (8')</li> </ul>
10:45–11:00	Scope of the Forum & Practical Information <ul style="list-style-type: none"> <li>• <b>Kim Finney</b>, Chair of the Scientific Organizing Committee (10')</li> <li>• <b>Masaki Kanao</b>, Chair of the Local Organizing Committee (5')</li> </ul>
11:00–12:30	<b>Session 1: Data – Lessons Learned from IPY</b> (Session chair: Taco de Bruin) <p><b>Keynote:</b> <i>Overarching Lessons from the International Polar Year on How to Create a Functional Data Infrastructure.</i> <b>Mark Parsons</b> (30')</p> <ul style="list-style-type: none"> <li>• <i>Arctic Communities, Social Science, and Data Management: Human dimensions of an Arctic Data Coordination Network.</i> <b>Peter Pulsifer</b> (15')</li> <li>• <i>The Polar Data Catalogue: Best Practices for Sharing and Archiving Canada's Polar Data.</i> <b>Julie Friddeil</b> (15')</li> <li>• <i>Preservation and Publication of IPY data – A Collaborative Effort of PANGAEA, ICSTII, and ICSU-WDS.</i> <b>Michael Diepenbroek</b> (15')</li> </ul>
12:30–14:00	<b>Lunch and Poster Session</b> (See list of posters for details)
14:00–15:30	<b>Session 2: Best Practice – Data Administration</b> (Session chair: Wim Hugo) <p><b>Keynote:</b> <i>Administering Antarctic Data – A Practical Use-case.</i> <b>Kim Finney</b> (30')</p> <ul style="list-style-type: none"> <li>• <i>Establishing Korean Polar Data Policy and its Future Directions.</i> <b>Dongmin Jin</b> (15')</li> <li>• <i>Design and Implementation of the Brazilian Information System on Antarctic Environmental Research.</i> <b>Rocío Zorilla</b> (15')</li> <li>• <i>The Arctic Cooperative Data and Information System: Data Management Support for the NSF Arctic Research Program.</i> <b>James Moore</b> (15')</li> </ul>

15:30–15:45 <b>Coffee Break</b>
15:45–17:45 <b>Session 3: Best Practice – Publication &amp; Citation</b> (Session co-chairs: Mark Parsons & Toshihiko Iyemori) <b>Keynote:</b> <i>Best Practice—Data Publication and Citation.</i> <b>Karen Visser</b> (30') <ul style="list-style-type: none"> <li>• <i>From Data to Publications: The Polar Information Spectrum.</i> <b>Shannon Vossepoel</b> (15')</li> <li>• <i>Inter-university Upper Atmosphere Global Observation Network (IUGONET) Metadata Database and Possible Application for 'Polar Data Activities'.</i> <b>Akiyo Yatagai</b> (15')</li> <li>• <i>Biodiversity.aq: Online Tools for Antarctic Biodiversity Data Discovery and Publication.</i> <b>Anton Van de Putte</b> (15')</li> </ul> <b>Panel discussion</b> (30') <ul style="list-style-type: none"> <li>• <b>Michael Diepenbroek</b> (PANGAEA, MARUM Centre for Marine Environmental Sciences, Bremen University. ICSU World Data System)</li> <li>• <b>Shuichi Iwata</b> (Data Science Journal, ICSU-CODATA)</li> <li>• <b>Karen Visser</b> (Australian National Data Service)</li> </ul> 17:45–18:00 <b>Group Photograph</b> 18:30–20:30 <b>Forum Banquet</b>
<b>Wednesday 16 October</b>
09:30–10:45 <b>Session 4-A: Best Practice – Data Sharing &amp; Observing Networks</b> (Session chair: Akira Kadokura) <b>Keynote:</b> <i>Data Sharing and Observing Networks—How can we do better?</i> <b>Lesley Rickards</b> (20') <ul style="list-style-type: none"> <li>• <i>Antarctic Space Weather Data Managed by IPS Radio and Space Services of Australia.</i> <b>Kehe Wang</b> (15')</li> <li>• <i>TRANSIMIT Prototype: Cross-institutional Network Approach from Geophysical Database to User Application for GNSS Science and Industry.</i> <b>Hiroatsu Sato</b> (15')</li> <li>• <i>Hydrometeorological Database (HMDB) for Practical Research in Ecology.</i> <b>Alexander Novakovskiy</b> (15')</li> </ul> 10:45–11:00 <b>Coffee Break</b>
11:00–12:30 <b>Session 4-B: Best Practice – Data Sharing &amp; Observing Networks</b> (Session chair: Ryosuke Shibasaki) <b>Keynote:</b> <i>Current Data Practises in Polar Institutions and Networks: A Case Study with the HIACMS Project.</i> <b>Paul Berkman</b> (20') <ul style="list-style-type: none"> <li>• <i>EXPEDITION: An Integrated Approach to Expose Expedition Information and Research Results.</i> <b>Roland Koppe</b> (15')</li> <li>• <i>Assembling an Arctic Ocean Boundary Monitoring Array.</i> <b>Takamasa Tsubouchi</b> (15')</li> <li>• <i>Building on the IPY: Discovering Interdisciplinary Data through Federated Search.</i> <b>Siri Jodha Singh Khalsa</b> (15')</li> <li>• <i>GEO Cold Regions—The Interface To GEOSS For Polar And Mountainous Cold Region Observations.</i> <b>Yubao Qui</b> (15')</li> </ul> 12:30–14:00 <b>Lunch and Poster Session</b> (See list of posters for details)

<p>14:00–15:15 <b>Session 5: Data Centre/Service Provider Accreditation</b> (Session chair: Lesley Rickards)</p> <p><b>Keynote:</b> <i>Certification and Accreditation of Data Centres and Services</i>. <b>Michael Diepenbroek</b> (30')</p> <ul style="list-style-type: none"> <li>• <i>ICPSR and the Data Seal of Approval: Accreditation Experiences and Opportunities</i>. <b>Jared Lyle</b> (15')</li> </ul> <p><b>Panel discussion</b> (30')</p> <ul style="list-style-type: none"> <li>• <b>Michael Diepenbroek</b> (PANGAEA, MARUM Center for Marine Environmental Sciences, Bremen University. ICSU World Data System)</li> <li>• <b>Toshihiko Iyemori</b> (World Data Centre for Geomagnetism)</li> <li>• <b>Akira Kadokura</b> (National Institute of Polar Research)</li> <li>• <b>Jared Lyle</b> (Interuniversity Consortium for Political and Social Research. Data Seal of Approval)</li> </ul>
<p>15:15–15:45 <b>Coffee break and Poster Session</b> (See list of posters for details)</p>
<p>15:45–17:15 <b>Session 6: The Future of Scientific Polar Data Management</b> (Session co-chairs: Kim Finney &amp; David Hik)</p> <p><b>Keynote:</b> <i>International Polar Data Management</i>. <b>Mustapha Mokrane</b> (30')</p> <ul style="list-style-type: none"> <li>• <i>Towards an International Polar Data Coordination Network: An Arctic Perspective</i>. <b>Peter Pulsifer</b> (15')</li> <li>• <i>Creating Web of Data for Science</i>. <b>Hideaki Takeda</b> (15')</li> </ul> <p><b>Panel discussion</b> (30')</p> <ul style="list-style-type: none"> <li>• <b>Paul Berkman</b> (Bren School of Environmental Science and Management, University of California, Santa Barbara)</li> <li>• <b>Taco de Bruin</b> (Royal Netherlands Institute for Sea Research)</li> <li>• <b>Mustapha Mokrane</b> (ICSU World Data System)</li> <li>• <b>Mark Parsons</b> (Research Data Alliance)</li> <li>• <b>Hideaki Takeda</b> (National Institute of Informatics, Japan)</li> </ul>
<p>17:15–17:45 <b>Session 7: Polar Forum Communique</b></p> <p><b>Facilitator:</b> <b>Rorie Edmunds</b> (ICSU World Data System)</p>
<p>17:45–18:30 <b>ICSU World Data System Town Hall Meeting</b></p>

## Poster Presentations

<b>P1</b>	<i>Hornsund GLACIO-TOPOCLIM Database—Polish Polar Station IPY Legacy</i>	<b>Bartłomiej Luks et al.</b>
<b>P2</b>	<i>Harmonizing Polar Biodiversity Data for Wider Access and Integration: A Collaboration between the Spanish Polar Data Center and GBIF-Spain</i>	<b>Oscar Bermúdez et al.</b>
<b>P3</b>	<i>'Quantarctica': New Standalone GIS Package for Antarctic Research, Operation, and Education using Open-source Software</i>	<b>Kenichi Matsuoka et. al. (Presenting author: Stein Tronstad)</b>
<b>P4</b>	<i>Polar Data Tools at the Integrated Earth Data Applications (IEDA)</i>	<b>Frank Nitsche et al. (Presenting author: Robert Arko)</b>
<b>P5</b>	<i>A Global Environmental Database Project at the National Institute for Environmental Studies and its Contribution to DIAS/GRENE</i>	<b>Hideaki Nakajima et al.</b>
<b>P6</b>	<i>Web-based Technologies in Data/Information Management for Polar Data in Russian World Data Centres</i>	<b>Evgeny Kahrin et al. (Presenting author: Michael Nisilevich)</b>
<b>P7</b>	<i>Polar Magnetic Data at WDC Kyoto—Services under International Collaborations</i>	<b>Toshihiko Iyemori et al.</b>
<b>P8</b>	<i>Cosmic-ray Neutron Data held by WDC for Cosmic Rays</i>	<b>Takashi Watanabe et al.</b>
<b>P9</b>	<i>Observations and Data Handling at International Center for Space Weather Science and Education, Kyushu University, Japan.</i>	<b>Hideaki Kawano et al.</b>
<b>P10</b>	<i>Data and Science Lessons Learned from Atmospheric/Environmental Observations in Japan and Alaska</i>	<b>Yasuhiro Murayama et al.</b>
<b>P11</b>	<i>Japanese Contribution to Super Dual Auroral Radar Network (SuperDARN)</i>	<b>Nozumu Nishitani et al.</b>
<b>P12</b>	<i>Continuous Broadband Seismic Observation on the Greenland Ice Sheet under Greenland Ice Sheet Monitoring Network</i>	<b>Seiji Tsuboi et al.</b>
<b>P13</b>	<i>Current Status of Science Data Archives for the Data Obtained by the Japanese Antarctic Research Expedition</i>	<b>Akira Kadokura et al.</b>



<b>P14</b>	<i>Metadata Management at the Polar Data Center (PDC) of the National Institute of Polar Research (NIPR), Japan</i>	<b>Masaki Kanao et al.</b>
<b>P15</b>	<i>Data Management Plans for the Southern Ocean Observing System (SOOS)</i>	<b>Louise Newman et al.</b>
<b>P16</b>	<i>NICT Science Cloud: Distributed Storage System and Parallel Data Processing Applicable for Polar Research Data</i>	<b>Ken Murata et al.</b>
<b>P17</b>	<i>Operation of Data Acquisition, Transfer, and Storage System for Worldwide Observation Networks</i>	<b>Tsutomu Nagatsuma et al.</b>
<b>P18</b>	<i>IUGONET Data Analysis Software (UDAS) for Upper Atmosphere Study</i>	<b>Yoshimasa Tanaka et al.</b>
<b>P19</b>	<i>Development of a Sensor Observation Service (SOS) Javascript Library</i>	<b>Alex Tate and Paul Breen (Presenting author: Alex Tate)</b>
<b>P20</b>	<i>Outline of Arctic Data Archive System (ADS)</i>	<b>Hironori Yabuki and Takeshi Sugimura (Presenting author: Hironori Yabuki)</b>
<b>P21</b>	<i>The ICSU World Data System: Trusted Data Services for Global Science</i>	<b>Rorie Edmunds</b>

## Side Meetings/Events

---

### SC-ADM-17 (Closed meeting)

The 17th meeting of SC-ADM will take place on 13–14 October in the Large Room (4F) of NMNS. A detailed agenda for the meeting will be prepared by the SC-ADM Co-chairs. Participants are limited to SC-ADM members. A **SC-ADM Dinner** is planned after the daytime meeting on 13 October. Details are to be confirmed, but a restaurant will be chosen in Ueno, near NMNS. More information will be given at the meeting.

### Icebreaker

An **Icebreaker** for the Polar Data Forum will be held on 14 October (18:00–20:00) in NMNS' main restaurant **MOUSEION**, once the SC-ADM meeting has been completed. The fee for the Icebreaker is covered by participants' registration fees.

<http://www.kahaku.go.jp/english/userguide/access/restaurant/index.html>

### Banquet (& Group Photograph)

A **Forum Banquet** will be held at the end of the first day of the Forum, on 15 October (18:30–20:30). The **SEIYOKEN** restaurant (<http://www.seiyoken.co.jp/index.html>) near Ueno Park, which is very famous for its western style menu, has been reserved for the occasion. To join the Banquet, a fee of 3,000 JPY is required from all participants.

Prior to heading to the restaurant, all participants are expected to be available for a **Group Photograph** (17:45–18:00), which will be taken in the Lecture Hall (2F).

### WDS Town Hall Meeting

A **WDS Town Hall Meeting** is planned in Large Room (4F) of NMNS (the poster presentation room) after the closing of the Polar Data Forum. All participants of the Forum are very welcome to join this meeting, which will allow those with an interest in WDS to interact with members of the WDS-SC in a social environment.

### WDS-SC#9 (Closed meeting)

The 9th biannual meeting of the WDS-SC will take place on 17–18 October at the NICT Office near Kojimachi Station, in central Tokyo. A detail agenda for the meeting will be prepared by the WDS-IPO. Participants are limited to WDS-SC members and invited guests.



**Data –  
Lessons Learned  
from IPY**

# Overarching Lessons from the International Polar Year on How to Create a Functional Data Infrastructure

**Mark A. Parsons**<sup>1\*</sup>

<sup>1\*</sup> *Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 12180, USA*  
Email: parsom3@rpi.edu

**Summary.** Immediately following the International Polar Year (IPY), data scientists documented lessons learned about complex, interdisciplinary data management. Have we learned more since then with a greater wisdom of hindsight? Broad, summative lessons from IPY can inform the development and practice polar data management. Data diversity is the central challenge of polar data management. Building relationships and collaborations across the data ecosystem are essential to addressing this challenge. These collaborations are part of improved, iterative, adaptive methods of data curation and system development. It is necessary to extend the collaborations in many directions locally and globally to evolve and sustain an international data infrastructure.

**Keywords.** Data infrastructure, data ecosystem, data stewardship, International Polar Year, data management.

## 1. Introduction

The International Polar Year 2007–2008 (IPY) was a huge and complex project. It presented an incredible data stewardship challenge. Soon after IPY, many data scientists attempted to document lessons learned about complex, interdisciplinary data stewardship. In particular, members of the IPY Data Committee and IPY Data and Information Service (IPYDIS) conducted two major analyses of IPY data management [1, 2].

These were the lessons of the moment, while IPY was still fresh and present. Now, nearly five years after the end of IPY, an attempt will be made to apply additional wisdom of hindsight to examine the deeper, lasting lessons of IPY.

## 2. Building Data Infrastructure

IPY planners laid out a noble and ambitious data management plan for IPY. The IPY Data Committee developed a visionary data policy. Polar data scientists around the world rallied to form the distributed IPYDIS. Polar data management policy and practice advanced immensely, but few would say that IPY has met all the objectives and visions of the planners. A

critical concern at the time was the lack of adequate funding for data management and coordination. This continues to be a concern, but it was perhaps not the core issue. Instead, perhaps, we as a community were somewhat naïve to the challenges of truly *interdisciplinary* data sharing. We assumed we were simply creating a data service from existing components, when we actually needed to be constructing an entire data infrastructure or ecosystem.

Building an infrastructure is a complex and fraught *socio-technical* exercise. It will take time, adaptation, patience, and creativity. That is the central lesson of IPY data stewardship. Four overarching themes will briefly be described that inform the overall process and can guide specific data stewardship activities supporting the development of data infrastructure.

## 3. Overarching Lessons

### 3.1 The challenge in diversity

In IPY, it was found that the greatest data stewardship challenge lies in the diversity of all the data necessary to truly understand complex systems such as the polar regions. Furthermore,

research collections are central to polar research, yet they can be the most disparate and challenging to manage.

Different disciplines also have different data systems, attitudes, and norms of behavior around data sharing that affect how we build integrated systems. For example, centralized metadata registries become unwieldy and imprecise when describing heterogeneous objects to diverse audiences. Instead, a union of specialized portals is looked toward using open web services—a data ‘bazaar’ rather than a ‘one-stop shop’.

### **3.2 Communities and collaboration**

Interoperability, indeed *infrastructure*, is built through relationships. The tacit knowledge of specialization is revealed and shared through relationships. Relationships can build community and collaboration. It was found in IPY that relationships between different data scientists and amongst data scientists, users, and providers improved data systems, documentation, and the data themselves. Great value was found in the creation of a global polar data community, while also integrating data scientists into their local disciplinary communities. Data scientists are often ‘in between’ workers or intermediaries, who can help build community. Data scientist training and career development is an important task.

### **3.3 Methods and training**

Part of a data scientist’s training needs to include instruction on methods, and again relationships are central. It was learned that when developing data systems, it is best to start simple, using proven approaches, and then take an incremental, iterative approach to expand their interconnection. This means that system designers need to work closely with, and be responsive and adaptive to, both data providers and users. Furthermore, user expectations and needs change, and systems always need to be evolving. This requires more than use-case-driven, agile development, it also requires case studies and ethnographic and cognitive science approaches to understanding how people conceive of, produce, and use data.

### **3.4 Globalism and localism**

Infrastructure works across all scales. It must function locally and reach globally. It is important to be constantly building relationships both globally and locally—to act ‘glocally’. For example, the real impact of the IPY data policy was felt when it was enforced by national governments, but the international recognition of the policy pressured the national governments to act. Correspondingly, a union catalogue of data sets could not begin to be built until local data centres were established and functional. In some cases, it took years of cultivating very local partnerships, before they could extend more broadly.

Regional success contributes to global success, which pushes local success. The polar community should continue to foster its own polar and disciplinary communities, while participating in global initiatives. The World Data System and the Research Data Alliance are two good mechanisms for doing this.

## **4. Conclusions**

IPY advanced polar data stewardship. It improved data availability and data science practice. To continue to address the complexities of diverse data, the community needs to grow, continually improve its practice, and build relationships globally and locally within disciplines and regions. Periodic conferences and assessments of the state of polar data practice should continue.

## **References**

1. Parsons, M., Godøy, Ø., LeDrew, E., de Bruin, T., Danis, B., Tomlinson, S. and Carlson, D. A., Conceptual framework for managing very diverse data for complex interdisciplinary science. *J. of Information Science*, 37(6), 555–69, 2011
2. Parsons, M., de Bruin, T., Tomlinson, S., Campbell, H., Godøy, Ø., LeClert, J., The state of polar data—the IPY experience. In Krupnik, I. et al. (Eds). *Understanding Earth’s Polar Challenges: IPY 2007–2008*, CCI Press, Edmonton, 2011

# Arctic Communities, Social Science, and Data Management: Human dimensions of an Arctic Data Coordination Network

**Peter L. Pulsifer**<sup>1\*</sup>

<sup>1\*</sup> National Snow and Ice Data Center, Univ. of Colorado 1540 30th St., Boulder, CO 80301, USA  
Email: pulsifer@nsidc.org

**Summary.** Observed and predicted environmental changes have considerably increased international interest in the Arctic in recent years. To build a comprehensive Arctic Data Coordination Network, it is necessary to understanding the human dimensions of the Arctic region. The International Polar Year 2007–2008 (IPY) was progressive with respect to including Arctic residents, indigenous organizations, and the Social Sciences; however, these individuals and organizations were also faced with the challenges of participating in multidisciplinary, large-scale, international research projects. This paper focuses on the results, lessons learned, and future directions of the still-continuing IPY project: Exchange for Local Observations and Knowledge of the Arctic. Moreover, issues and opportunities related to inclusion of Arctic social sciences within the domain of polar data management are reviewed.

**Keywords.** Arctic, data management, social science, indigenous knowledge, research networks.

## 1. Introduction

The Arctic has been home to indigenous peoples for many generations. Since colonization began, people from all over the world have explored, travelled, inhabited, occupied, and exploited the lands and resources of the Arctic. In recent years, international interest in the Arctic has significantly increased as observed and predicted environmental changes have combined with a dynamic world order, and technological, economic, and socio-cultural shifts in the relationships between the Arctic and other parts of the world [1].

Understanding the human dimensions of the polar regions in general, and the Arctic region in particular, is necessary if a comprehensive Arctic Data Coordination Network (henceforth ‘the Network’) is to emerge over the coming years. If humans are to be an integral part of the Network, individual Arctic communities, indigenous organizations, social scientists, and social science research organizations must be part of the development process.

The International Polar Year 2007–2008 (IPY) marked significant progress with respect to the inclusion of Arctic residents, indigenous

organizations, and the Social Sciences in an IPY science program. This provided an opportunity for indigenous peoples and social scientists to engage in the IPY research process. At the same time, these same individuals and organizations were faced with the challenges of participating in multidisciplinary, large-scale, international research projects; most of which involved the need to perform large-scale data management [2]. Much was learned from this process and many IPY projects continue in one form or another.

This paper focuses first on the results of what was originally IPY project #187: the Exchange for Local Observations and Knowledge of the Arctic (ELOKA) [3]. ELOKA continues in the post-IPY era, and the results of the project, lessons learned, and future directions are discussed. Second, the issues and opportunities related to the inclusion of Arctic social sciences are reviewed within the domain of polar data management.

## 2. Data Management and Arctic Communities

Out of the frozen landscape in which they dwell,

indigenous peoples have carved a productive, vital culture. Through experience and skills passed on from previous generations, their local observations and knowledge tell the story of drastic changes to the Arctic environment and climate—changes that have a global impact. Until recently, the science and policy communities have largely overlooked indigenous local observations and knowledge, and related data. Today, indigenous peoples are acknowledged as investigators, partners, and collaborators, and their local knowledge and observations are being documented for use in Arctic research and policy development.

ELOKA is a research and data management research support service that specializes in working with Arctic communities and researchers in the collection, preservation, and use of local and traditional knowledge and community-based monitoring (CBM) data and information. Through the coproduction of information systems, knowledge holders, indigenous community members, their representative organizations, and university researchers these resources are being made available to a broad audience, including scientists and policy makers.

Here, an international sample is presented of community-oriented information systems including narrative websites, interactive web-based mapping with near real-time user contributed content, CBM databases, and community driven sensor-based applications. The lesson learned and future directions are discussed with respect to application development, interoperability, and practicing ethical data management.

### **3. Data Management and Arctic Social Sciences**

While large-scale, formal data management programs have historically not been prominent in the Arctic social sciences, the need to collaborate across distance and disciplines, demands of IPY, changes in data management policies by funding agencies, and a recognition of the need for long-term preservation of data—

where appropriate—have been key drivers in making data management an important topic of discussion among Arctic social scientists and their representative organizations. Organizations such as the International Arctic Social Sciences Association formally recognize the importance of data management, and the need to engage in the development of collaborative data management initiatives [4]. Specific contributions are presented that can be made by Arctic social sciences in the area of polar data management, and the organizations coordinating related efforts.

### **4. Conclusions**

In conclusion, details will be provided of an existing and expanding network of individuals and organizations with an interest in Arctic community-oriented and social science research and data management, and how this group can engage with the broader Network.

### **References**

1. Shadian, J. M., Tennberg M., *Legacies and Change in Polar Sciences: Historical, legal, and political reflections on the International Polar Year*. MPG Books Group, United Kingdom, 2009
2. Krupnik, I., Hovelsrud, G., Polar Societies and Social Progress. In: Krupnik, I., Allison, I., Bell, R., Cutler, P., Hik, D., López-Martínez, J., Rachold, V., Sarukhanian, E., Summerhayes, C., *Understanding Earth's Polar Challenges: Summary by the IPY Joint Committee, International Polar Year 2007–2008*, 2011
3. ELOKA, <http://eloka-arctic.org> [accessed on: September 2013]
4. IASSA–IASC, Letter of Agreement between the International Arctic Social Sciences Association (IASSA) and the International Arctic Science Committee (IASC), 2008. Online resource: [http://www.iasc.info/files/MOU\\_LOA/IASC\\_IASSA.pdf](http://www.iasc.info/files/MOU_LOA/IASC_IASSA.pdf) [accessed on: September 2013]

# The Polar Data Catalogue: Best Practices for Sharing and Archiving Canada's Polar Data

Julie E. Friddell<sup>1\*</sup>, Josée Michaud<sup>2</sup>, Warwick F. Vincent<sup>2</sup>, Ellsworth F. LeDrew<sup>1</sup>

<sup>1\*</sup> *University of Waterloo, 200 University Avenue West, Waterloo, ON, N2V 3G1, Canada*

<sup>2</sup> *Université Laval, 2325 rue de l'Université, Québec, QC, G1V 0A6, Canada*

Email: julie.friddell@uwaterloo.ca

**Summary.** The Polar Data Catalogue (PDC) is a growing Canadian archive and online public access portal for Arctic and Antarctic research and monitoring data. Through multi-sectoral direction and support over the past decade, the PDC has adopted international standards and best practices to provide a robust infrastructure for reliable security, storage, discoverability, and access to Canada's polar data and metadata. Current efforts focus on seeking new partnerships and implementing new methods for sharing the PDC's archive with other data centres through metadata interoperability protocols.

**Keywords.** Data repository, International Polar Year, best practices, metadata submission, interoperability.

## 1. Introduction

Scientific research in the Canadian Arctic has increased tremendously over the past few decades, especially with development of large research programmes such as ArcticNet and Canada's International Polar Year (IPY). With these programmes also comes the need to efficiently build systems to manage the collected data and ensure proper preservation, stewardship, and access, while respecting confidentiality requirements and researchers' rights to publication [1]. A related challenge to building such a system is to accommodate the vast amounts of data and the huge diversity of topics and fields represented.

## 2. The Polar Data Catalogue

To meet these challenges, the Network of Centres of Excellence—ArcticNet—initiated a partnership with the Canadian Cryospheric Information Network (CCIN) at the University of Waterloo to jointly develop the Polar Data Catalogue (PDC) [2]. The PDC has built one of the largest repositories of polar data in Canada, in collaboration with the Government of Canada Program for IPY 2007-2008, Department of Fisheries and Oceans Canada, Environment

Canada, Noetix Research Inc., GeoConnections, Centre d'études nordiques, Inuit Tapiriit Kanatami, Northern Contaminants Program, Circumpolar Biodiversity Monitoring Program (CBMP), Beaufort Regional Environmental Assessment, and Canadian Polar Data Network (CPDN).

To effectively manage these data, the multi-sectoral PDC Management Committee has developed a management plan with three phases. The first phase consisted of developing a secure infrastructure—including a database and online applications—to facilitate metadata and data archiving and discovery, as well as a set of best practices for optimal metadata and data management. The second phase consisted of providing a unique online presence for archived datasets through the use of Digital Object Identifiers (DOIs). The third phase is to extend Network partnerships and collaboration with other research programs, and polar data and archiving centres, to ensure sustainability and interoperability.

## 3. Best Practices

The PDC has produced a best practices document for submitted data [3] based on best practices for environmental data [4]. The eight



critical steps are identified in Table 1, and range from creating metadata to properly citing datasets.

**Table 1.** Best practices.

Best Practice	Objective
Create metadata	Provide the what, where, and when of data, by whom
Assign descriptive titles	Be as descriptive as possible and include the time period/location
Use constant and stable data formats	Format should be readable far into the future and independent of application changes
Define the content of data files	Provide adequate information to fully understand content of datasets, including describing variables and units
Use consistent data organization	Favour common and understandable arrangement of data rows and columns
Perform basic quality assurance	Provide datasets that are free of errors
Provide documentation	Provide information for a user who is unfamiliar with the data
Cite a dataset	Provide a constant citable format for data

Providing documentation with data is critical. The PDC provides a README template with specific questions to help data providers properly describe the data they are submitting.

#### 4. DOIs

DOIs are International Organization for Standardization (ISO) standard online identifiers that provide long-term links to data; improving the discoverability, accessibility, and citability of the data to which they are assigned. As partners in CPDN, which provides a national cloud-based deep archive of the data, the PDC is working closely with the Canada Institute for Scientific and Technical Information (Canada's member of DataCite International) to assign DOIs to PDC datasets. DOIs are viewed by researchers as an incentive to provide their data to the PDC, and thus make the PDC an attractive database for polar researchers.

#### 5. Network and Interoperability

During IPY, the PDC partnered with the US National Snow and Ice Data Center and the

Norwegian Meteorological Institute to share IPY-related metadata records via the Open Archives Initiative Protocol for Metadata Harvesting. In the intervening years, additional interoperability has been established with CBMP, Yukon Research Centre, Ontario Council of University Libraries, and Inuit Qaujisarvingat. Further discussions are currently underway with organizations in the United Kingdom, Sweden, and Japan to initiate metadata interoperability, and additional sharing protocols are being implemented, including Web Map Service and Catalogue Service for the Web via GeoNetwork. Development is proceeding at CCIN to facilitate discovery of the shared metadata records. PDC metadata records currently follow the Federal Geographic Data Committee standard, but metadata will be soon converted to the North American Profile of the ISO 19115 metadata standard to increase interoperability with other data centres.

#### 6. Conclusions

The success of the Polar Data Catalogue in Canada is due to a management plan that focuses on adherence to international standards and best practices for archiving data; incentives for researchers to contribute metadata and data, including project funding and annual renewal being contingent on entering and updating PDC entries; and the growing Network of partners involved in and contributing to its initial and ongoing development.

**Acknowledgments.** We would like to thank Leah Braithwaite and other members of the PDC Management Committee for guidance; ArcticNet and other partners for funding support; and the hundreds of researchers who willingly provide their time and data for archiving and sharing.

#### References

1. Vincent, W. et al., Data Management. In: Vincent, W. et al. (Eds.), *Impacts of Environmental Change in the Canadian Coastal Arctic: A Compendium of Research Conducted during ArcticNet Phase I (2004–2008)*, ArcticNet Inc., Québec City, Canada,

19-20, 2010

2. CCIN, [www.polardata.ca](http://www.polardata.ca) [accessed on: September 2013]
3. CCIN, [www.polardata.ca/pdcinput/public/PDC\\_Best\\_Practices\\_FULLL.pdf](http://www.polardata.ca/pdcinput/public/PDC_Best_Practices_FULLL.pdf) [accessed on: September 2013]
4. Hook, L. et al., *Best Practices for Preparing Environmental Datasets to Share and Archive*. Oak Ridge National Laboratory Distributed Active Archive Center, updated 2007

# Preservation and Publication of IPY Data – A Collaborative Effort of PANGAEA, ICSTII, and ICSU-WDS

**Michael Diepenbroek<sup>1\*</sup>, Amelie Bucker<sup>2</sup>, Hannes Grobe<sup>2</sup>**

<sup>1\*</sup> Center for Marine Environmental Science (MARUM), Bremen University, Hochschulring 18, Bremen, 28359, Germany

<sup>2</sup> Alfred-Wegener-Institute for Polar and Marine Research (AWI), Am Handelshafen 12, Bremerhaven, 27570, Germany

Email: mdiepenbroek@pangaea.de

**Summary.** PANGAEA, ICSU-WDS, and ICSTII have worked together to extract over 1000 IPY data sets from more than 340 science articles for preservation and publication. Keywords relevant to IPY projects are used to filter out supplementary data from articles and linked materials, which are then checked for quality and annotated with metadata. These data can be accessed through the PANGAEA and ICSU-WDS data portals, and can be cross-referenced with science articles. However, large amounts of IPY data are still neither archived nor openly accessible, and this initiative has highlighted the need for early involvement of organizations such as ICSU-WDS in projects to help better plan and fund data management efforts.

**Keywords.** International Polar Year, long-term preservation, data publication, open access, data management.

## 1. Introduction

In a collaborative effort between PANGAEA – Data Publisher for Earth and Environmental Science, the ICSU World Data System (ICSU-WDS), and the International Council for Scientific and Technical Information (ICSTII), scientific data resulting from projects related to the International Polar Year (IPY) have been extracted from literature for long-term preservation and publication.

## 2. IPY Data Preservation

In a first step, ICSTII staff compiled a bibliography using keywords relevant to IPY projects. This bibliography referencing more than 1200 articles served as a basis for PANGAEA staff to filter out journal articles having supplementary data that could be extracted either from the articles or from supplementary materials supplied with the publication. In the end around 500 articles with data or data supplements were identified, which could be archived in PANGAEA.

Preparation of data included a technical quality control of data (check for outliers,

correctness of geocoding, precision of values, etc.) and - using the editorial system of PANGAEA - annotation with metadata. Data and metadata were imported into the PANGAEA archive (relational database; RDB).

Due to harmonization of content and structure of data and metadata, archived datasets are efficiently usable.

## 3. IPY Data Access

Data can be retrieved via the PANGAEA data portal [1] or via the ICSU-WDS Data Portal [2] using the keyword 'IPY'.

Due to collaborations of PANGAEA with science publishers, data can be cross-referenced with science articles. In some cases, data are directly accessible from the article's splash page (e.g., Reference [3]).

## 4. Discussion

Up until the end of August 2013 data supplements of more than 340 publications have been archived through PANGAEA, resulting in

around 1000 datasets accessible in machine readable format.

The initiative presented here describes a data management effort *following* the IPY. Notwithstanding the overall positive effect of the work performed, the initiative also exemplifies the limitations of data management efforts that are not synchronized from the very beginning with science activities. Not only can it be assumed that large parts of existing data from IPY projects are not archived and will probably never be openly accessible, it also must be noted that during the initiative presented here, data producers could not be included into the editorial and final assessment of data.

## 5. Conclusions

The long-term storage of and free access to scientific data is an indispensable basis for scientific knowledge. Early involvement of

organizations like ICSU-WDS and of WDS members in the process of project planning and funding would allow for better data management activities. This not only leads to more collaboration and synergies between projects during the runtime of a science program but also improves the data basis for on-going and future programs and projects. It is therefore one of the declared aims of ICSU to establish a closer context between science programs and the World Data System.

## References

1. PANGAEA, <http://www.pangaea.de/search?q=project%3Aipy> [accessed on: July 2013]
2. ICSU World Data System, <http://www.icsu-wds.org/services/data-portal> [accessed on: July 2013]
3. ScienceDirect (Elsevier), <http://www.science-direct.com/science/article/pii/S0012821X11001233> [accessed on: July 2013]



# **Best Practice – Data Administration**

# Administering Antarctic Data – A Practical Use-case

**Kim Finney**<sup>1\*</sup>

<sup>1\*</sup> *Australian Antarctic Division, Australian Antarctic Data Centre, Kingston, 7050 Tasmania, Australia*  
Email: kim.finney@aad.gov.au

**Summary.** Scientific data administration is performed to ensure that data are curated in a manner that supports their qualified reuse. Curation usually involves actions that must be performed by those who capture or generate data and a facility with the capability to sustainably archive and publish data beyond an individual project's lifecycle. The Australian Antarctic Data Centre is such a facility. How this Centre has approached the administration of Antarctic science data is described in the following paper and serves to demonstrate key facets necessary for undertaking polar data management in an increasingly connected global data environment.

**Keywords.** International polar data management, scientific data curation, data administration, long-term preservation, open-access.

## 1. Introduction

The Australian Antarctic Data Centre (AADC), which has been operating for 16 years as the primary data repository for the Australian Antarctic Science program (AAp), has been gradually refining its policy base, cementing its integration into the science program workflow and upgrading under-pinning infrastructure. All of this is aimed at improving data administration services that are available to Antarctic researchers with the goal of lifting the volume and type of data that are publicly accessible for re-use.

AAp is a competitive research program involving scientists from the Australian Antarctic Division, Commonwealth Scientific and Industrial Research Organization, State/Federal government agencies, the university sector and international institutions. AADC coordinates the archiving and publication of data derived from AAp Antarctic and Southern Ocean-based research according to the open data principles of the Antarctic Treaty System [1]. In performing its functions the Centre works as part of the international network of Antarctic Data Centres, co-ordinated under the auspices of the Scientific Committee on Antarctic Research and was admitted to the International Council for Science: World Data System (ICSU-WDS) in 2011. ICSU-

WDS is an international federation of global data centres and data service providers. Australia's ability to contribute to such global systems and to reuse data within AAp and beyond is dependent upon scientists paying adequate attention to data management tasks that need to be performed within individual science projects and upon easy researcher access to core data management infrastructure. This paper describes how AADC has been approaching polar data administration and developing infrastructure to support AAp science.

## 2. Science Application Process and AAp Data Policy

In 2010, AADC conducted an audit of the data it had received from past science projects implemented under the umbrella of the Australian Science program in all of its previous guises, since the establishment of the Antarctic Division in 1980. There was a specific focus on those projects that commenced post the creation of AADC (in 1996). Not surprisingly, a large number of projects had not submitted any data for archiving, despite there being a long-standing policy that 'all data should be deposited with the AADC'. Three critical issues were

identified as contributing to this poor level of compliance:

- A lack of applied penalties associated with non-compliance.
- No prior understanding by AADC of what datasets should be delivered from approved projects.
- An inadequate set of utilities available for AADC to administer policy compliance, and limited tools and assistance for scientists to comply with many of the data policy obligations.

Recognising that reforms were necessary, development of the new 2011–2021 Antarctic Science Strategic Plan offered an opportunity to strengthen the AAp Data Policy [2] and more closely align it with the science project assessment process, as well as begin a process of upgrading the AADC toolset.

Since the introduction of the new Strategic Plan, a public biennial call is made for science proposals. These are assessed by peer review under a Ministerial approved assessment process. This process provides for rating proposals based on a range of criteria. A researcher's previous history of data submission is now taken into account in the scoring of assessment criteria. Since the program is highly competitive, the relatively few points allocated to 'data submission history' still has the capacity to make the difference between a proposal being supported, or not. This is a good incentive for scientists to make sure that they have allocated sufficient effort and resources to preparing data for archiving and subsequent reuse. The first milestone in all new successful AAp projects now requires the submission of a Data Management Plan, to be delivered to AADC within the first six months of receiving project approval. In this Plan, scientists must identify what datasets they intend to collect and when they think these data will be ready for submission to AADC. All investigators must have submitted data by a project's end-date and are encouraged to formally cite their data in publications using

persistent addressing (i.e., digital object identifiers).

### **3. MyScience**

To manage the data and information flow associated with administering the AAp Data Policy, AADC had to redevelop its various and separate information systems. With a keen desire to minimize application maintenance overhead, it was decided that the primary tool that would be used by the AADC to administer compliance with the Policy would also be a tool that could be used by AAp scientists to manage their individual project-based resources (i.e., metadata, data, publications, and data management plan). The web-based application developed to fulfil this function is called MyScience.

MyScience is accessible via secure login and is available to any scientist with an Internet connection, a browser, and one or more registered AAp projects (past or current). The application uses a range of controlled vocabularies with which scientists can tag content. Investigators are encouraged to supply new vocabulary terms where there are deficiencies. These vocabularies are then reused in other AADC core infrastructure to mark-up data for exchange within global data networks.

Importantly, MyScience was introduced to AAp in conjunction with the allocation of Science (project) Liaison Officers. These Officers routinely monitor clusters of projects and provide individualized support to scientists in terms of helping them to comply with the Data Policy.

### **4. Conclusions**

The data administration activities outlined in this paper have already resulted in long outstanding datasets being deposited within AADC. The Centre is now far better placed to administer the AAp Policy, and scientists are being supported to comply with Policy obligations. Investigator cooperation is helping to build better infrastructure, which as a result, is now more

closely meeting scientific data publication, discovery, access, and integration requirements.

## **References**

1. Antarctic Treaty Secretariat, Antarctic Treaty System, <http://www.ats.aq/e/ats.htm> [accessed on: August 2013]
2. Australian Antarctic Program Data Policy, [https://data.aad.gov.au/aadc/about/data\\_policy.cfm](https://data.aad.gov.au/aadc/about/data_policy.cfm) [accessed on: August 2013]



# Establishing Korean Polar Data Policy and its Future Directions

**Dongmin Jin<sup>1\*</sup>, Min Choul Lee<sup>1</sup>**

<sup>1\*</sup> Korea Polar Research Institute, 26, Songdomirae-ro, Yeonsu-gu, Incheon, 406-840, South Korea  
Email: djmin@kopri.re.kr

**Summary.** Korea implemented its Antarctic research program in 1987 and diversified to the Arctic in 2002. Since the development of the Joint Committee on Antarctic Data Management, Korea has acknowledged the importance of data management in the Antarctica. The launch of Korea Polar Research Institute (KOPRI) in 2004 also saw establishment of the Korea Polar Data Center (KPDC), which outlines and executes a Polar Data Policy that is applied to all KOPRI-supported projects and includes submission of a data management plan with all project proposals. KPDC has set up an information technology infrastructure and developed a metadata management system for efficient data management. But, there is still a long way to go; especially in terms of raising researcher recognition on the importance of data management for improving data registration and sharing. This presentation will share KPDC's experiences during its establishment and data gathering activities, and will introduce major components of its data policy, as well as future directions.

**Keywords.** Polar data, data policy, data management plan.

## 1. Introduction

The Korea Polar Research Institute (KOPRI) is the national operator of the Korean Polar Program, and established the Korea Polar Data Center (KPDC) in 2010. KPDC's role is to efficiently manage and collaboratively share polar data produced by the Korean Polar Program. Korea implemented its Antarctic research program in 1987 and diversified into the Arctic in 2002.

The Scientific Committee on Antarctic Research and the Council of Managers of National Antarctic Program initiated the Joint Committee on Antarctic Data Management in 1997 to seek out the best way forward for Antarctic data management. However, this was before KOPRI was set up as an autonomous and affiliated research body in the Korean Ocean Research and Development Institute (KIOST; previously KORDI), and active discussions—and subsequent concrete preparations—for establishing KPDC began in 2010.

The establishment of KPDC led to the adoption of a Polar Data Management Policy within KOPRI, along with regulations and guidelines prescribing definitions and procedures for handling polar data.

Interesting contents of the Data Policy are as follows. First, not only Antarctic but also Arctic data are included due to Korean bi-polar activities. Second, the policy mandates that researchers include a data management plan (DMP) when submitting a project proposal; this plan is evaluated as a part of the proposal. Third, to maintain data quality and minimize any losses, researchers should upload all data to KPDC within three months of acquisition. Lastly, to enhance cooperative use of data, metadata are made open after registration and raw data are made open after a three-year exclusivity period.

The data policy is implemented as a research institute regulation without any underpinning national legal basis, however, which results in some limitations on its execution. Certain domestic laws such as the Marine Scientific Research Act might be applicable to handling polar data, but can involve controversies. A Consultative Meeting of the Antarctic Treaty underlined the importance of data sharing, and with the International Arctic Science Committee now adopted a data policy emphasizing open use of data, KPDC is now being asked to respond in a timely and appropriate manner to such international trends and fill the domestic gap.

## 2. KPDC Establishment and Outline of its Polar Data Policy

Endeavours to establish KPDC began immediately after the launch of KOPRI as autonomous institute within KIOST in 2004. Lack of manpower and budget initially delayed progress, but as Korean investment in polar research activities increased, KOPRI secured sufficient finances to found KPDC in 2010. The setting up of KPDC was carried in two directions: development of a data policy and of infrastructure, including software. KPDC outlined a Korean Polar Data Policy reflecting international requirements and in doing so increased researcher awareness on the importance of polar data. Moreover, KPDC installed a system with storage provision and metadata management tools.

### 2.1 Considerations when developing Korean Polar Data Policy

Korean polar research activities are not limited to Antarctica but also extend to the Arctic. Since the Antarctic Treaty states that scientific observations from Antarctica shall be made freely available, all Antarctic data should be correctly managed.

Integrated data management has not been resolved in Korea. Even domestic laws such as the Marine Scientific Research Act that enforce researchers to submit data acquired during research activities are not fully complied with.

In contrast, international communities have underlined the importance of polar data in aiding global climate change efforts, and efficient use and prompt sharing of data are trends worldwide.

With Korea now operating the RV Araon icebreaker in both polar regions and the completion of the second Antarctic research station in the 2014 austral summer season, the Korean polar program is expected to produce increasingly more data.

### 2.2 Key content of the Polar Data Policy

KPDC data covers not only Antarctic but also Arctic data. On the basis of the Antarctic Treaty and the statement of principles and practices for Arctic data management, these data will be (in principle) open for free access and use after a

predefined exclusive-use period (see Table 1).

**Table 1.** Registration and open access periods.

	Registration	Open Access
Metadata	Immediately after acquisition	As soon as confirmed by administrator
Raw Data	Within three months of acquisition	Three years after date of acquisition

When researchers present research proposals, they must add a DMP as a part of the proposal. KPDC then requests researchers to register and submit relevant data according to the submitted DMP (see Table 2).

**Table 2.** Number of annual registrations.

	2010	2011	2012	Total
Counts	17	201	105	323

## 3. Limitations

### 1. Researcher awareness

Datasets acquired during research activities have been historically treated as personal property and many researchers are still reluctant to submit data. It will take time to change this attitude.

### 2. Lack of legal obligation

The implemented data policy has been built on KOPRI's internal regulations, which have limited applicability. Furthermore, it may conflict with domestic laws for handling polar data.

### 3. Development of systems and experts

Polar data management is not a short-term consideration. Continual system development and training of experts are required.

## 4. Conclusions

KPDC's Data Policy is not seen as perfect, and it will explore the following in the future.

1. Development of schemes to enhance further researcher awareness on polar data.
2. Cooperation with the Korean government to develop national legislation.
3. Development of software for efficient management.
4. Seeking of a stable budget and manpower.

# Design and Implementation of the Brazilian Information System on Antarctic Environmental Research

**Rocío M. Zorrilla C.<sup>1\*</sup>, Maira Poltosi<sup>1</sup>, Luiz M. R. Gadelha<sup>1</sup>,  
Fábio Porto<sup>1</sup>, Ana Maria de C. Moura<sup>1</sup>**

<sup>1\*</sup> *National Laboratory for Scientific Computing, Av. Getúlio Vargas 333, Petrópolis, Rio de Janeiro, 25651-075, Brazil*  
Email: romizc@gmail.com

**Summary.** Data generated by environmental research activities in Antarctica are essential for evaluating how its biodiversity and environment are being affected by changes that are happening at a global scale and that are triggered by ever-increasing human activities. In this work, we describe the Brazilian Information System on Antarctic Environmental Research (BrISAntar), which enables the acquiring, storing and querying of research data generated by activities of the Brazilian National Institute for Science and Technology on Antarctic Environmental Research. Even though BrISAntar's data model is organized around field gathering operations, to facilitate querying, it is straightforward to export data stored in BrISAntar to standards commonly used standards in the field.

**Keywords.** Antarctic environmental research, scientific data management.

## 1. Introduction

Global environments have been extensively changed by humans, which has impacted on the biodiversity of these environments. Antarctica is no exception to this trend, having seen increases in air temperature and reduction in its glaciers. To determine more precisely the extension and rate of biodiversity change, it is essential to gather, archive, and analyze data on spatial and temporal distribution of species along with information about their surrounding environment [1–2]. It is also important to use data integration techniques in order to make these data discoverable and queryable.

In this work, the Brazilian Information System on Antarctic Environmental Research (BrISAntar) is presented; an information system that enables the acquiring, storing, and querying of research data generated by the Brazilian National Institute for Science and Technology on Antarctic Environmental Research [3].

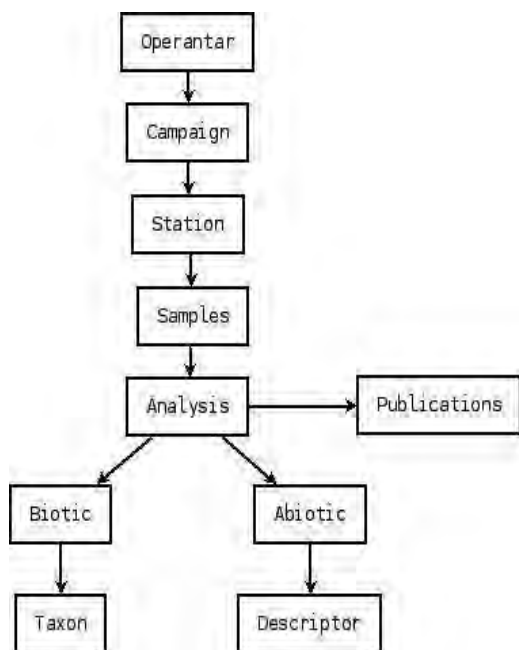
## 2. BrISAntar Overview

Here, the architecture and a simplified view of the

conceptual data model of BrISAntar are presented. The architecture of BrISAntar is organized into the following layers: Presentation, Application, Persistence, and Database.

The Presentation layer contains the logic for rendering the user interface that includes, for instance, functionality for uploading data about collected samples and performed analyses. The Application layer is responsible for managing transactions related to the application logic. The Persistence layer manages data on a relational database.

The conceptual data model of BrISAntar is organized around the concepts of operations and campaigns, which are related to actual field expeditions for collecting samples and performing measurements. Figure 1 characterizes the order that must be followed when populating the database, where an Operantar represents the beginning of an annual expedition that is given by several campaigns in the Antarctic region. These campaigns are executed at a specific station; a geographical region with fixed points where sample collection is carried out.



**Figure 1.** Proposed data model.

With these collected samples, different types of analysis are performed and the results registered. Results are classified in two types: biotic or abiotic. Biotic results are stored following the structure of a known taxonomic database, whereas abiotic results are stored as a set of descriptors and values. The results produced by an analysis may lead to a scientific publication and in this case, information about the publication should also be registered in the system.

### 3. Conclusions

In its current state, BrISAntar facilitates data acquisition, storage, and querying; providing a valuable tool to the Brazilian community on Antarctic environmental research. Additional functionalities currently being developed include: a data visualization and analysis module, where data can be visualized in maps or through charts;

data publication modules for exporting data using Darwin Core [4] and Ecological Metadata Language [2] data standards to enable integration of the data available in BrISAntar with those in global infrastructures such as the Global Biodiversity Information Facility [5] (in particular, the Antarctic Biodiversity Information Facility [6]) and the Data Observation Network for Earth [7]— a scientific workflow module for enabling scientists to automate their analysis routines.

**Acknowledgments.** This project is funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico.

### References

1. Hardisty, A., Roberts, D., A Decadal View of Biodiversity Informatics: Challenges and Priorities. *BMC Ecology*, 13, 16, 2013
2. Michener, W. K., Porter, J., Servilla, M., Vanderbilt, K., Long Term Ecological Research and Information Management. *Ecological Informatics*, 6(1), 13–24, 2011
3. Valentin, Y. Y., Dalto, A. G., Lavrado, H. P. (Eds), INCT-APA Annual Activity Report 2012. Editora Cubo, 2012. Available online at: <http://www.biologia.ufrj.br/inct-antartico> [accessed on: September 2013]
4. Wiczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., Vieglais, D., Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE*, 7(1), e29715, 2012
5. Global Biodiversity Information Facility, <http://www.gbif.org> [accessed on: August 2013]
6. Belgian Biodiversity Platform, <http://www.biodiversity.aq> [accessed on: August 2013]
7. Data Observation Network for Earth, <http://www.dataone.org> [accessed on: August, 2013]

# The Arctic Cooperative Data and Information System: Data Management Support for the NSF Arctic Research Program

**James A. Moore<sup>1\*</sup>, Mark C. Serreze<sup>2</sup>, Don Middleton<sup>1</sup>, Mohan Ramamurthy<sup>3</sup>, Lynn Yarmey<sup>2</sup>**

<sup>1\*</sup> National Center for Atmospheric Research, Boulder, Colorado, USA

<sup>2</sup> National Snow and Ice Data Center, Boulder, Colorado, USA

<sup>3</sup> University Corporation for Atmospheric Research, Boulder, Colorado, USA

Email: jmoore@ucar.edu

**Summary.** The US National Science Foundation (NSF) funds the Advanced Cooperative Arctic Data and Information System (ACADIS). This system is operational, and serves the growing and increasingly diverse data management needs of NSF's arctic research community. The ACADIS Gateway is located at: <http://www.aoncadis.org/>. The ACADIS investigator team combines experienced data managers, curators, and software engineers from the National Snow and Ice Data Center, University Corporation for Atmospheric Research, and National Center for Atmospheric Research. ACADIS fosters scientific synthesis and discovery by supplying a service that provides a secure long-term data archive to NSF investigators. The system offers preservation services for data, as well as discovery and access to Arctic-related data from this and other archives. This paper describes the technical components of ACADIS, the implementation of best practices, the value of ACADIS to the community, and the major challenges facing this archive in the future for handling the diverse data coming from US NSF Arctic investigators.

**Keywords:** NSF Arctic data management, Arctic data stewardship, data management best practices.

## 1. Introduction

The Advanced Cooperative Arctic Data and Information System (ACADIS) provides sustainable data management, data stewardship services, and leadership for the US National Science Foundation (NSF) Arctic research community through open data sharing, adherence to best practices and standards, capitalizing on appropriate evolving technologies, and community support and engagement. ACADIS leverages other pertinent projects, taking advantage of appropriate emerging technologies and participating in emerging cyberinfrastructure initiatives.

The key elements of ACADIS that support the NSF Arctic community include:

- *Data Services:* Fundamental technical services offered through ACADIS;
- *Support for Datasets with Special Requirements:* ACADIS handling of unique datasets;

- *Metadata Documentation, Sharing, and Usability:* All aspects of metadata generation, updating and utilization in ACADIS;
- *Interoperability and Initiatives:* Activities to enhance, develop, and improve data management practices and technology;
- *Science Support:* Other community support activities as a result of expanding ACADIS acceptance and usage.

## 2. Technical Components

The data services offered by ACADIS focus on (1) meeting the needs of the data providers, (2) providing dynamic search capabilities to peruse ACADIS and related cryosphere data repositories, (3) efficient data download, and (4) special services including dataset reformatting and visualization. These services are built around of the following key technical elements:

1. The ACADIS Gateway housed at the National Center for Atmospheric Research has been

developed to support NSF Arctic data originating from the Arctic Observing Network, and now broadly across the Division of Polar Programs/Arctic Sciences organization and related archives.

2. The Arctic Data Explorer (ADE) developed at the National Snow and Ice Data Center (NSIDC) is an integral service of ACADIS, bringing the rich archive from NSIDC together with catalogues from ACADIS and international partners in Arctic research.
3. *Rosetta* and the Digital Object Identifier (DOI) generation scheme are tools available to the community to help publish and utilize datasets in integration and synthesis.

### 3. Challenges

To provide a self-service data publishing platform that requires minimal curation oversight, while maintaining rich metadata for discovery, access, and preservation is challenging. Search tools such as ADE are limited by inconsistency, poor form, and content gaps in metadata records. Although metadata standards are a first step towards consistent content, more work is necessary. Moreover, the ability to discover a dataset does not imply that the data can be directly downloaded. The complementary and unique capabilities of the ACADIS Gateway and ADE are intended to offer the user choices for data discovery and access, with the clear objective of increasing discovery and use of all Arctic data.

ACADIS is embarking on the critical activity of generating DOIs for all its holdings. This is a big job, involving development of consistent identification across many disciplines before implementing this across thousands of datasets in the archive.

Metadata is at the core of ACADIS activities—from capturing metadata at the point of data submission, to ensuring interoperability with other larger initiatives, providing data citations, and supporting data discovery. ACADIS metadata efforts include: (1) evolution of the ACADIS metadata profile to increase flexibility in search, (2) documentation guidelines, and (3) metadata

standardization. A major activity is now underway to ensure consistency in the metadata profile across all archived datasets.

ACADIS accepts that some special handling may be required as a normal part of its community support services. Its biggest challenge is to make sure that these special datasets (e.g., large volume, many data files) can be accessed via the ACADIS Gateway. Development is in progress to implement automated flexible procedures for uploading of data, metadata, and documentation capable of dealing with a wide variety of data types and formats.

ACADIS is engaged in multiple collaborations and partnerships to enhance, develop, and improve data management practices, technology, and cyberinfrastructure within the Arctic science communities, as well as include external partners participating in the activity.

ACADIS user-support services remain the vital link to our science community. The three key components of this support are data (plus metadata and documentation) submission to the archive, effective discovery tools for located data in ACADIS and related polar archives, and providing efficient access procedures for downloading data needed for science analysis activities.

### 4. Conclusions


This project offers a strategy as to how to develop and implement an agile and robust data catalogue with access, visualization, and long-term archive capabilities. It is believed to be a viable strategy for developing international data exchange protocols through best practices; use of standards; and open access to data, metadata, and documentation. Data managers should ideally be involved early in the planning process so that they can develop a plan to support an observing network. Part of this process is to understand the specific funding agency requirements that might drive system design and implementation (e.g., real-time data availability; access to complementary operational data; and

agreement on formats including parameter names, units, and resolution). Data managers take this information and—by employing best practices—design a process/system that meets community needs, stakeholder requirements, and public expectations.

**Acknowledgements:** ACADIS is supported through a collaborative grant to NSIDC and UCAR/NCAR from the US NSF, Grant #ARC1016048.







# **Best Practice – Publication & Citation**

# Best Practice—Data Publication and Citation

**Karen Visser**<sup>1\*</sup>

<sup>1\*</sup> Australian National Data Service, WK Hancock Building, The Australian National University, Acton, ACT 0200, Australia

Email: karen.visser@ands.org.au

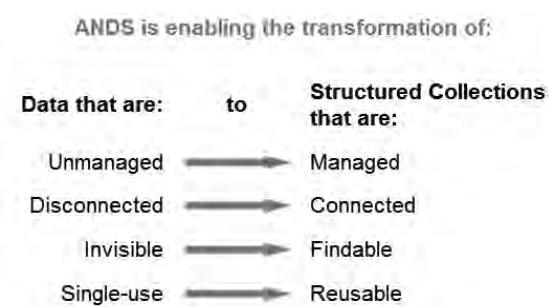
**Summary.** Data citation is rapidly becoming an accepted, and indeed in some cases expected, scholarly publication practice. Research data is now not only a reference within scholarly articles but also a first class scholarly output in its own right. But what can researchers do to ensure their research data outputs are acknowledged and cited in publications? The Australian National Data Service is an active player in the global data citation arena and aims to forge significant advantages in data publication for research communities. This keynote will examine international research data publishing trends, as well as provide best practice exemplars of associated data citation practices. It will also explore how researchers can position themselves and their data to take advantage of these new data publication opportunities.

**Keywords.** Data management, data citation, data administration, publishing, open-access.

## 1. Introduction

Ensuring that Australian research data are accurately cited and correctly acknowledged is seen as a critical component of a new paradigm of data-intensive research.

The Australian National Data Service (ANDS) [1] is supported by the Australian Government, and its primary goal is to develop infrastructure that enables the transformation of Australian research data (see Figure 1).



**Figure 1.** Transformation of Australian research data by ANDS.

ANDS is building Research Data Australia [2], which is an Internet-based discovery service designed to provide rich connections between data, projects, researchers, and institutions, as

well as promote visibility of Australian research data collections in search engines.

ANDS has been working closely with the Australian Antarctic Data Centre (AADC) to connect their data records into the Australian Research Data Commons. Additionally, AADC has been an early adopter of both the principles and scholarly practices associated with data citation, including: ensuring quality metadata to enable accurate citation, minting Digital Object Identifiers (DOIs), taking part in the Thomson Reuters Data Citation Index Trial, and being a leading institution in the Australian Data Citation Community of Practice facilitated by ANDS.

ANDS works with over 79 research organizations to ensure that they and their researchers understand the benefits of data citation and have the capability to take advantage of the benefits it offers now, but more importantly, in the future.

## 2. The Data Citation Rubik's Cube

Components of the increasingly interlinked emergent global trends associated with data management may be viewed as akin to a Rubik's Cube, where each component is related to the others and only when all facets are in alignment

does the full impact of that alignment become apparent. The six major components of this Data Citation Rubik's Cube include:

- Governments wishing to encourage innovative responses to the 'Big Questions' of science and to increase returns on research investments.
- Publishers embracing new publishing paradigms and developing products to measure the impact of research data.
- Social media opportunities, which aid in the diffusion of research data and may indicate early identification of seminal datasets.
- Funders, research organizations, and researchers who see benefits from shared access to research data.
- Adaptation of existing scholarly linkage and identification mechanisms to enable data to be recognized as a first class research output.
- National bodies, such as ANDS, who work with all of the other facets to ensure standards and interoperability of metadata.

Together, these trends and mandates send a clear message to researchers and research institutions that the publishing (and funding) landscape is changing and that data management practices need to change to leverage new opportunities.

### **3. First, Manage My Data**

Aligning global trends to individual or group research datasets and harnessing the benefits of such alignments, however, is premised on the notion that the research data is managed and well described, including: good descriptors (metadata) when depositing research data in institutional or discipline repositories, appropriate research codes, funding information, application of open licenses where possible, insurance of ethics clearances, inclusion of underlying data, and the obtaining of a DOI.

Well-managed and well-described data can hot-house citation and collaboration opportunities through linking data, research materials, journal publications, associated software, derived datasets, conference proceedings, and grants.

### **4. Find and Connect My Data**

Researchers who do not align the facets of their own research data practices within this new paradigm may find that opportunities for scholarly publishing diminish in the near future. There are many exemplars and practices that support a researcher in increasing their academic profile through the routine publication of their data outputs.

### **5. Finally, Cite My Data**

A growing body of literature indicates that publications including citations to underlying data have higher citation rates than articles without reference to the research data [3]. Adding a DOI will help formal citation index services and social media altmetric services to track who is reusing research data, and in which field of research.

### **6. Conclusions**

Use and reuse of research datasets is logical, time and cost efficient, and will enable innovative responses to problems afflicting our embattled planet. Data citation is pivotal to this process.

The work of ANDS and the exemplars provided by the AADC and other Australian research institutions are supporting data citation scholarly practices internationally. This work is also contributing to the emerging reward systems for data publication.

### **References**

1. Australian National Data Service, <http://ands.org.au/cite-data/index.html> [accessed on: September 2013]
2. Research Data Australia <http://researchdata.ands.org.au/> [accessed on: September 2013]
3. Piwowar, H. A., Day, R. S., Fridsma, D. B., Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS ONE*, 2(3), e308, 2007

# From Data to Publications: The Polar Information Spectrum

**Shannon Vossepoel<sup>1\*</sup>**

<sup>1\*</sup> Arctic Institute of North America, University of Calgary, ES 1040, 2500 University Drive, NW, Calgary, Alberta, T2N 1N4, Canada

Email: shannonv@ucalgary.ca

**Summary.** The Arctic Institute of North America is keenly aware of the need to develop better connections between polar publications and research project information contained in its databases and other types of polar information held by other organizations. Here, an overview is presented of the six types of information, the ways in which they are used, the importance of ensuring that this information is interconnected, and a vision for the future of connecting polar information in Canada.

**Keywords.** Information management, databases, future, sharing, Canada.

## 1. Introduction

The polar information spectrum consists of information that describes at least six different types of entities: researchers, organizations, research facilities, research projects, research datasets, and publications. The management of polar research datasets has been the subject of enormous attention in recent years, but it is only one aspect of the polar information puzzle. Escalating interest in polar regions has prompted a great appetite for all six types of polar information and an increasing need to ensure that the six types are interconnected in data management systems.

## 2. Canada's Data Management Systems

In Canada, different data management systems (databases) are responsible for maintaining different types of polar information. The Canadian Polar Commission is responsible for the management of databases that contain information about polar researchers and polar research facilities. The Polar Data Catalogue at the University of Waterloo manages a database that contains metadata describing polar research datasets. The Arctic Science and Technology Information System (ASTIS) at the Arctic Institute of North America manages one main database and several subset databases that contain

information about polar publications and polar research projects. At present, there is no database in Canada that manages information about polar organizations.

The Canadian Polar Commission, founded in 1991, keeps track of information about individual polar researchers in Canada through the Researcher's Directory, available for free online [1]. They also keep track of information about Canadian polar research facilities through the Northern Research Facilities database, which is also available for free [2]. The Polar Data Catalogue (accessed freely online [3]) handles metadata for the research datasets created by the Canadian International Polar Year (IPY) programme, ArcticNet, the Beaufort Regional Environmental Assessment (BREA), the Climate Change Adaptation Programme, Aboriginal Affairs, and Northern Development Canada's Northern Contaminants Programme. Moreover, they have been designated to handle metadata for the planned Canadian High Arctic Research Station.

ASTIS is the oldest of the three data management systems. In operation since 1978, the ASTIS database currently contains 78,000 records describing publications and research projects about northern Canada and the circumpolar Arctic. ASTIS is responsible for the Canadian IPY Publications Database, which is part

of the international IPY Publications Database. ASTIS also covers all publications produced by ArcticNet, BREA, Aboriginal Affairs, and Northern Development Canada's Northern Contaminants Program, as well as publications about northern Canada from many other sources. ASTIS also contains 16,000 research project descriptions from the three Canadian northern territories based on information collected by the agencies that license all field research. The ASTIS database is available for free online [4] and a full list of all ASTIS subset databases is also available [5].

### 3. Interconnecting Information

Publications and research project descriptions typically act as a gateway to polar information for individuals who do not work directly with raw data. Publications, in particular, provide context and meaning for research data. Publications contain the knowledge that researchers have extracted from their data and present that knowledge in a form that can be used by others. As gateway information sources, it is particularly important to ensure that information about publications and research projects is managed effectively in data management systems and that it is linked to information about research datasets, researchers, research facilities, and organizations.

Linking these information types will provide a more holistic view of polar information, and

enable users to access all the information they need quickly and easily.

### 4. Conclusions

The Arctic Institute of North America is keenly aware of the need to develop better connections between polar publications and research project information contained in its databases and other types of polar information held by other organizations. It envisions the future of polar information as one in which all types of polar information are interconnected and easily accessible by everyone.

### References

1. Canadian Polar Information Network, [http://www.polarcom.gc.ca/index.php?page=resource-library&hl=en\\_US](http://www.polarcom.gc.ca/index.php?page=resource-library&hl=en_US) [accessed on: September 2013]
2. Canadian Polar Information Network, [http://www.polarknowledge.ca/index.php?page=northern-research-facilities&hl=en\\_US](http://www.polarknowledge.ca/index.php?page=northern-research-facilities&hl=en_US) [accessed on: September 2013]
3. Canadian Cryospheric Information Network, <http://www.polardata.ca/> [accessed on: September 2013]
4. Arctic Institute of North America, <http://www.aina.ucalgary.ca/astis/> [accessed on: September 2013]
5. Arctic Institute of North America, <http://www.arctic.ucalgary.ca/databases> [accessed on: September 2013]

# Inter-university Upper Atmosphere Global Observation Network (IUGONET) Metadata Database and Possible Application for 'Polar Data Activities'

**Akiyo Yatagai<sup>1\*</sup>, Tomoaki Hori<sup>2</sup>, Shuji Abe<sup>3</sup>, Yoshimasa Tanaka<sup>4</sup>, Atsuki Shinbori<sup>1</sup>, Yukinobu Koyama<sup>5</sup>, Norio Umemura<sup>2</sup>, Yuka Sato<sup>4</sup>, Manabu Yagi<sup>6</sup>, Satoru UeNo<sup>7</sup>, Noriko O. Hashiguchi<sup>1</sup>, Naoki Kaneda<sup>7</sup>**

<sup>1\*</sup> Research Institute for Sustainable Humanosphere, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

<sup>2</sup> Solar-Terrestrial Environment Laboratory, Nagoya University, Furou-cho, Chikusa-ku, Nagoya 464-8601, Japan

<sup>3</sup> International Center for Space Weather Science and Education, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

<sup>4</sup> National Institute of Polar Research, 10-3, Midori-cho, Tachikawa, Tokyo 190-8518, Japan

<sup>5</sup> Data Analysis Center for Geomagnetism and Space Magnetism, Graduate School of Science, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan

<sup>6</sup> Planetary Plasma and Atmospheric Research Center, Graduate School of Science, Tohoku University, 6-3 Aramaki Aza-Aoba, Aoba-ku, Sendai Miyagi 980-8578, Japan

<sup>7</sup> Kwasan and Hida Observatories, Graduate School of Science, Kyoto University, Kurabashira, Kamitakara-cho, Takayama, Gifu 506-1314, Japan

Email: akiyo\_yatagai@rish.kyoto-u.ac.jp

**Summary.** An overview of the Inter-university Upper Atmosphere Global Observation Network project is presented. This is a Japanese programme to build a metadata database for ground-based observations of the upper atmosphere, in which metadata of observatories, instruments, datasets, and so on, obtained from both polar regions by various atmospheric radars and photometers are archived. The project began four-and-a-half years ago and is registering historical data, including analogue data stored at the National Institute of Polar Research. By using the metadata database, researchers throughout the world can access the data files and information of data holders. The system is also used for other database systems in the field of Earth and Planetary Sciences and the interoperability is being developed of the database with other databases.

**Keywords.** Database, metadata, upper atmosphere, Earth and planetary sciences.

## 1. Introduction

The Inter-university Upper atmosphere Global Observation Network (IUGONET) [1–2] is an inter-university programme that was started in 2009 by the National Institute of Polar Research, Tohoku University, Nagoya University, Kyoto University and Kyushu University with the purpose of building a database for ground-based observations of the upper atmosphere. The IUGONET institutions archive data at various altitudes from the Earth's surface to the Sun observed by radars, magnetometers, photometers, radio telescopes and helioscopes, and so on.

## 2. Metadata Database

Systems have been developed to facilitate searching of the metadata, for further updating of the system, and inclusion of additional metadata. The IUGONET development team adopted the Space Physics Archive Search and Extract metadata model [2] to describe the upper atmosphere data. This model is used as the common metadata format by the Virtual Magnetospheric Observatory and other virtual observatories for solar-terrestrial physics. It includes metadata referring to each data file (called a 'Granule'), which enable searching for data files, as well as datasets. Further details are

described in [2] and [3].

Currently, three additional Japanese institutions are being incorporated in IUGONET: the National Institute of Information and Communication Technology, the Solar Observatory/National Astronomical Observatory of Japan and the Kakioka Magnetic Observatory/Japan Meteorological Agency. That such institutes are keen to join IUGONET shows that it has become a common and indispensable database system for upper atmospheric study, especially in Japan. Furthermore, metadata of observations of the troposphere (such as X-band radar and radiosonde observations) taken at the middle and upper atmosphere radar observatories at Shigaraki and the Meteor radar in Indonesia, have been incorporated. These additions will contribute to efficient interdisciplinary scientific research.

At the start of 2013, registration of the 'Observatory' and 'Instrument' metadata was completed, which simplifies searching of the metadata database. The number of registered metadata until the end of July 2013 totalled 8.8 million, including 793 observatories and 878 instruments. Many important observatories/instruments registered in IUGONET are located in and around the polar regions of Antarctica, the Arctic, and the 'Third Pole Environment' (i.e., the Tibetan Plateau). The search system is now being distributed to better manage the increasing number of metadata, meaning that each institution holds its own metadata on its own server.

### 3. Analysis Software

IUGONET does not regulate the format of data. Instead, iUgonet Data Analysis Software (UDAS) is being developed to handle all types of formatted data on the same platform. UDAS is a plug-in software for the THEMIS Data Analysis Software Suite (TDAS), which is written in Interactive Data Language (IDL). By using UDAS/TDAS/IDL, it is possible to easily download, plot, and analyze observed data. Further details are described in [4].

### 4. Further Developments

It is important to increase the interoperability and metadata exchange between the database development groups. A Memorandum of Understanding has been signed with the European Near-Earth Space Data Infrastructure for e-Science project, which has similar objectives to IUGONET in regard to a framework for formal collaboration.

Furthermore, observations by satellites and the International Space Station are being incorporated with a view to creating/linking metadata databases. The development of effective data systems will contribute to the progress of scientific research on solar terrestrial physics, climate, and the geophysical environment.

### 5. Conclusions

IUGONET has developed (and continues to develop) a metadata database for ground-based observations of the upper atmosphere, including both polar regions. All types of cooperation, metadata input, and feedback are welcomed—especially those concerning database connection in the near future.

**Acknowledgements.** The IUGONET project was initially supported by the Special Budget (Project) for FY2010 and in more recent years by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

### References

1. Inter-university Upper atmosphere Global Observation Network, <http://www.iugonet.org/> [accessed on September 2013] (in Japanese)
2. Hayashi, H. et al., Inter-university Upper Atmosphere Global Observation Network (IUGONET). *Data Sci. J.*, 12, WDS179–184, 2013.
3. King, T. et al., SPASE 2.0: A standard data model for space physics. *Earth Sci. Inform.*, 3, 67–73, 2010.

4. Hori, T., et al., Development of IUGONET metadata format and metadata management system. *J. Space Sci. Info. Jpn.*, 105–111, 2012 (in Japanese)
5. Tanaka, Y. et al., Development of IUGONET data analysis software. *J. Space Sci. Info. Jpn.*, 91–98, 2012 (in Japanese)



# Biodiversity.aq: Online Tools for Antarctic Biodiversity Data Discovery and Publication

**Anton P. Van de Putte<sup>1\*</sup>, Nabil Youdjou<sup>1</sup>, Bruno Danis<sup>2</sup>**

<sup>1\*</sup> Royal Belgian Institute for Natural Sciences, Brussels, B-1000, Belgium

<sup>2</sup> Université libre de Bruxelles, Av. F. D. Roosevelt, 50, Brussels, B-1000, Belgium

Email: Antonarctica@gmail.com

**Summary.** www.biodiversity.aq is a portal that offers free and open access to Antarctic biodiversity data. It takes the form a set of web platforms giving access to a distributed network of contributing databases, according to the principles of the Global Biodiversity Information Facility. Continuing the legacy of the Scientific Committee on Antarctic Research's Marine Biodiversity Information Network and the Antarctic Biodiversity Information Facility, it is building a set of new data discovery tools using an advanced technical infrastructure, capable of linking with many potential data resources, and of presenting them in many different contexts. Here an overview is presented of the different features, and will showcase how its components can be used to access existing data, but also to publish and present information online in a coherent way. For example, the Data Paper concept will be introduced as a feature of the Integrated Publishing Toolkit, which is a component of the infrastructure.

**Keywords.** Biodiversity, Integrated Publishing Toolkit, Scientific Committee on Antarctic Research, Antarctic, Southern Ocean.

## 1. Introduction

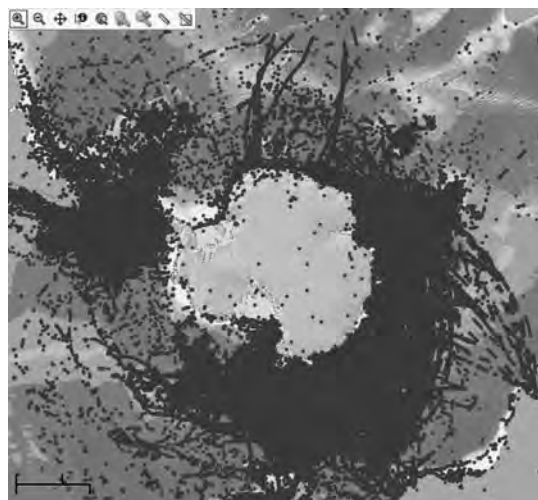
Biodiversity.aq is an open-access platform designed for scientists to publish and share baseline scientific data on Antarctic biodiversity (see Figure 1). Its roots can be traced back to the International Polar Year (IPY) [1] and the Scientific Committee on Antarctic Research – Marine Biodiversity Information Network (SCAR–MarBIN).

Initially, the SCAR-MarBIN initiative was launched during the Census of Antarctic Marine Life (CAML) and IPY timeframe to provide the community with the tools to establish a web-based inventory of Antarctic marine biodiversity, from micro-organisms to whales, that would be openly accessible to all. However, while SCAR-MarBIN was aimed at marine biodiversity data, biodiversity.aq integrates both marine and terrestrial biodiversity data.

## 2. Integrated Publishing Toolkit

The Integrated Publishing Toolkit (IPT), created by the Global Biodiversity Information Facility (GBIF)

and Pensoft publishers, allows its users to standardize, document and share their primary biodiversity data using a simple web interface. It's a one-stop shop for opening up data, which can be crucial in the face of the extremely rapid environmental change that is currently challenging Antarctic ecosystems.



**Figure 1.** An example of data available through biodiversity.aq.

Once data have standardized and shared, by following eight simple steps, the IPT can be used to generate a readable data paper that can be submitted to a growing number of data publishers.

An IPT dedicated to Antarctic biodiversity data can be found at [ipt.biodiversity.aq](http://ipt.biodiversity.aq), and offers a data hosting and publishing service to nations or research institutes that lack such facilities. This enables all researchers to easily publish their data, thus making it available for future generations of scientists.

### 3. Data

The biodiversity.aq network provides access to data from both the marine and terrestrial realms. Data.biodiversity.aq not only provides access to data submitted through the IPT it also aggregates Antarctic biodiversity data from various providers, including the German PANGAEA and the Australian Antarctic Data Centre.

The network also feeds information into global biodiversity initiatives such as the Ocean Biogeographic Information system (OBIS), which offers access to marine species datasets from around the world, and the Global Biodiversity Information Facility (GBIF). All such publicly available Antarctic biodiversity data can be searched and retrieved through this data platform.

### 4. Collaborations

To attract more users and showcase the usefulness of the SCAR information networks, a series of data derived products have been developed also using a community based approach. For instance, biodiversity.aq collaborates on the development of the Antarctic Field Guides (AFG) and the Biogeographic Atlas of the Southern Ocean (BASO) projects.

AFG is a collaborative tool offering free access to information that can help users identify Antarctic organisms. At [afg.biodiversity.aq](http://afg.biodiversity.aq) users can build a tailor-made customized guide to be taken into the field or simply browsed. The pages are generated instantly from the contents of

authoritative, quality-controlled data resources, and ensure user access to up-to-date information about the group of organisms of interest. High-quality pictures and species descriptions provided by experts allow identification of Antarctic organisms. Any user can create a personalized field guide.

The multi-authored BASO will provide an up-to-date synthesis of Antarctic and sub-Antarctic biogeographic knowledge and a new comprehensive online resource for visualization, analysis, and modelling of species distribution. The major objectives of BASO are to establish a new synthesis of the biogeography of the Southern Ocean (patterns and processes), covering benthos, zooplankton, nekton, birds, and seals to provide a benchmark of current biogeographic knowledge.

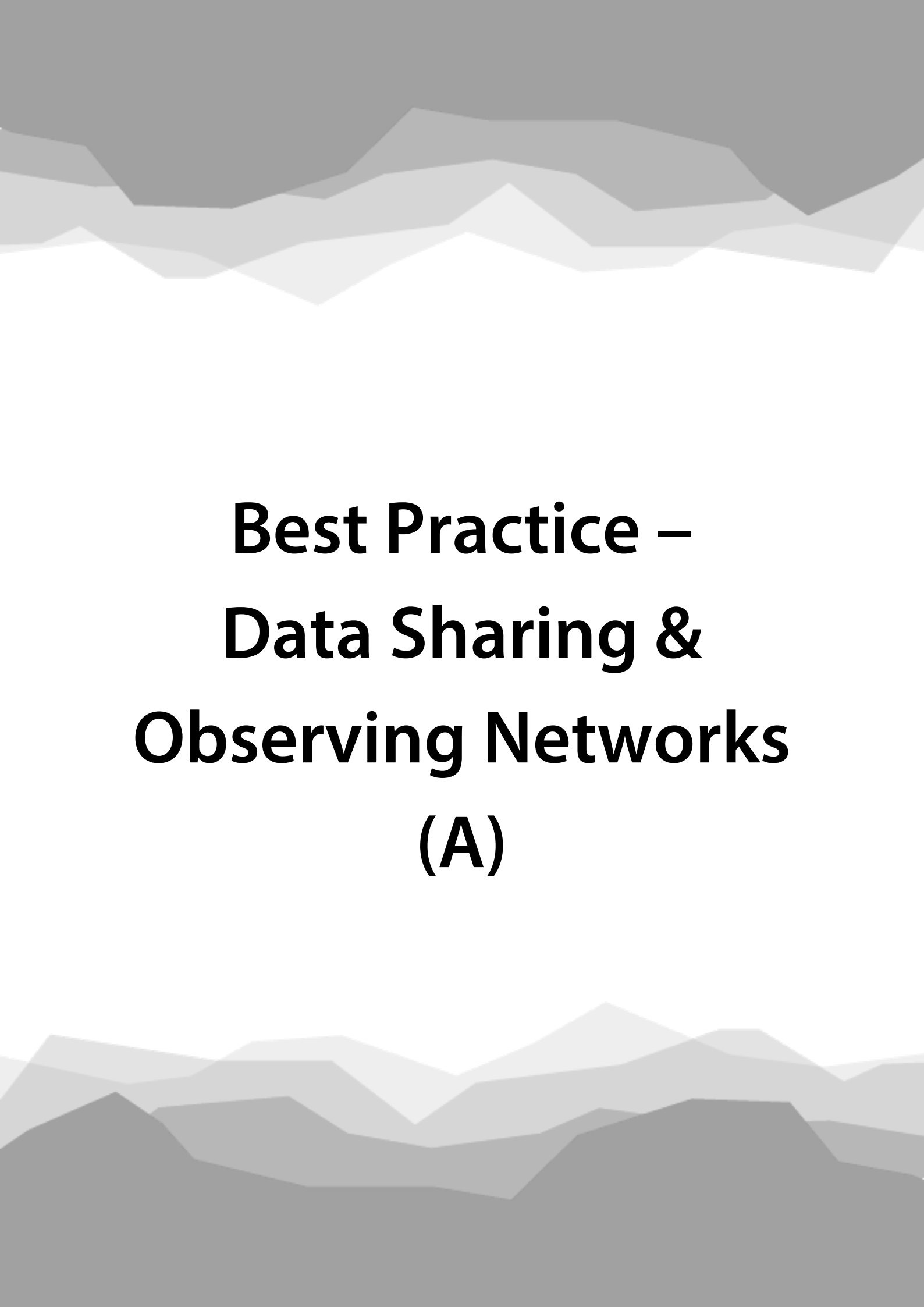
### 5. Conclusions

Biodiversity.aq is an open-access platform designed for scientists to publish and share baseline scientific data on Antarctic biodiversity. It acts as the Antarctic Regional node for global biodiversity networks such as GBIF and OBIS.

**Acknowledgments.** The authors wish to thank the vast community involved in the biodiversity information networks, in particular, the Scientific and Follow-up Committees of SCAR-MarBIN and ANTABIF/biodiversity.aq. The two initiatives are funded by the Belgian Science Policy Office and are implemented within the Belgian Biodiversity Platform. We would also like to thank the different sponsors, including: the Alfred P. Sloan Foundation, CAML, the Alfred Wegener Institute, the German Research Foundation, SCAR, the Australian Antarctic Division, and the Netherlands Organization for Scientific Research.

### References

1. Danis B., Van de Putte A. P., Renaudier S, Griffiths H., Connecting Biodiversity Data During the IPY: The Path Towards e-Polar Science. *Adaptation and Evolution in Marine Environments*, 2, 21–32, 2013



**Best Practice –  
Data Sharing &  
Observing Networks  
(A)**

# Data Sharing and Observing Networks—How can we do better?

**Lesley J Rickards<sup>1\*</sup>**

<sup>1\*</sup> *Permanent Service for Mean Sea Level (PSMSL), National Oceanography Centre, Joseph Proudman Building, 6 Brownlow Street, Liverpool, L3 5DA, UK  
Email: ljr@bodc.ac.uk*

**Summary.** Science in general and environmental science in particular involves the collection of data, often from observing networks. These data are precious: often irreplaceable; they are always unique, if only in the timing of collection. In the polar regions, their spatial and temporal coverage is quite sparse—and they are extremely expensive to collect. But they are fundamental to the understanding of the processes that control our natural environment—and the importance of polar regions in this regard has been long recognised. Case studies and examples will be used to highlight current good practice in polar observing networks and associated data sharing. However, there remain missing elements which will be considered together with some ideas on how these can be resolved. The advent of the ICSU World Data System will also be highlighted as an important element of the solution for data sharing.

**Keywords.** Data sharing, observing networks, best practice.

## 1. Introduction

It has long been realized that observations in the polar regions are essential to improve our understanding how our planet works. In fact, environmental observations underpin many of the activities we undertake including scientific research, modelling, monitoring and assessment. The resulting data are a unique resource, both now and in the future, and over the years a variety of databases have been compiled bringing together data from many different sources.

The data collected from observing networks are often irreplaceable; they are always unique, if only in the timing of collection. Even when considering all of the data collected, spatial and temporal coverage is quite sparse. Data from the polar regions data can also be extremely expensive to collect. In recent years there has been need for access to more interdisciplinary and integrated data sets to further our knowledge and understanding and to better manage the polar regions. There is also an increasing requirement for operational data in near-real-time for forecasting.

## 2. Data Policies and Data Sharing

Data sharing will be discussed both in the context of data policies (e.g., International Council for Science (ICSU), World Meteorological Organization, Intergovernmental Oceanographic Commission, and Antarctic Treaty) and in terms of what needs to be shared and how best to do this.

As far as possible, and for maximum use of data, the aim should be to make data freely available. A strong public domain for scientific data and information promotes greater return from investment in research. It stimulates innovation and enables more informed decision-making. The talk will comment on the success or otherwise of these policies.

The international Argo profiling float programme is a good example of a project with free and open access to all the data collected, real-time data are available within 24 hours and quality controlled data on a longer time scale. Against this, it is important to ensure intellectual property rights are not compromised and scientific papers produced by those responsible for the data collection. In addition it is important to give proper credit to the data collectors and data must be properly referenced or cited.

Once the policies have been agreed, then technical decisions made to ensure that data can be shared both now and in the future. Use of standards is the key to further progress. Examples will be used to illustrate present practice for data discovery, data exploration and visualisation, and data delivery. Some shortcomings of the current situation are discussed, balanced by highlighting some of the recent advances and successes made in providing access to integrated data sets. The ICSU World Data System, with its 60 accredited data centres and data analysis services has the potential to play a major role in these data sharing activities.

### **3. Observing Networks**

Efficient and effective observing networks are essential in the polar regions to enhance our understanding of these regions and their global influence. This talk will present several case studies to illustrate best practice in polar observing systems from data collection to data sharing. Selected examples, taken from the Southern Ocean Observing System and Sustaining Arctic Observing Networks amongst others, will be examined. This will include observing networks of autonomous vehicles or robots (e.g., Argo profiling floats, marine mammals, and oceanographic underwater gliders), biodiversity networks (e.g., Scientific Committee on Antarctic Research – Marine Biodiversity Network, and Arctic Biodiversity Monitoring Program) and terrestrial observing networks (e.g., International Network Terrestrial Research and Monitoring in the Arctic).

### **4. Funding**

One issue that cannot be ignored is that of funding. Although funding usually comes from national sources, international cooperation is

essential. However, the funding must go beyond the cost of the observation networks themselves, expensive though they are, to include long term data stewardship and data dissemination. There are plenty of examples where observing networks are operational, but the data difficult to find or where much work is needed to make the data easily useable, both now and for future generations. Good practice will be considered together with a discussion of lessons learned from projects where full funding has been difficult to obtain.

### **5. Conclusions**

Increasingly environmental data from observing networks in polar regions and elsewhere are being recognised as a valuable resource. Open access and sharing of data are important. There are many initiatives underway to make observations and share data—these may be national, project or programme based or international. However there are still major barriers to the efficient re-use of data and to overcome these, and make the best use of the new technologies available, a culture change is required. The big issue of sharing data is not one just for those who are data managers, but also scientists and other data collectors who must provide information alongside their data to ensure maximum benefit.

Finally, it is also necessary to be aware of developments outside of the polar arena that interact and overlap with it and with whom we need to develop and maintain collaboration. It is essential to bring in and build upon other related activities and not work in isolation—not reinventing what has already been developed or is under development, but adapting and adopting as appropriate.

# Antarctic Space Weather Data Managed by IPS Radio and Space Services of Australia

**Kehe Wang<sup>1\*</sup>, Colin Yuile<sup>1</sup>, Dave Neudegg<sup>1</sup>, Andrew Kelly<sup>1</sup>**

<sup>1\*</sup> *IPS Radio and Space Services, Bureau of Meteorology, Australia, Level 15, Tower C, 300 Elizabeth Street, Surry Hills, NSW, 2010, Australia*  
Email: K.Wang@bom.gov.au

**Summary.** The Ionospheric Prediction Service (IPS) has an extensive collection of data from field instruments in the Antarctic, the oldest data being ionospheric recordings dating back to the 1930s. The IPS sensor network (IPSNET) spans the Australasian and Antarctic regions to collect information on space weather. In the Antarctic, sensors include ionosondes, magnetometers, riometers, and cosmic ray detectors. Those data collected in the Antarctic region flow into the IPS World Data Centre at Sydney, mostly in near real time. These data, combined with data from other stations in IPSNET, provide a basis for developing space weather reports for Antarctica. The majority of the archived data are available to domestic and international clients online. This paper summarizes the datasets collected from Antarctic stations, and their data management within IPS.

**Keywords.** Antarctic, space weather, data management, AWK, jqPlot.

## 1. Introduction

Since the 1930s, nations worldwide have established greater than 30 space weather observation stations in the Antarctic region, both independently and in collaboration with other countries. Currently, many of those stations are still in daily operation.

There are four Australian stations in the Antarctic and sub-Antarctic operated by Australian Antarctic Division: Casey, Davis, Mawson, and Macquarie Island. From these, and the Scott Base (operated by the New Zealand Antarctic Programme with an ionosonde from The University of Canterbury), the Ionospheric Prediction Service (IPS) collects five different types of space weather related data: ionogram, magnetometer, riometer, cosmic ray, and ionospheric scintillation data. All data are transferred to Sydney and archived in the IPS World Data Centre, and are then published via the IPS official website [1] and a File Transfer Protocol site (ftp.ips.gov.au). Thus the data can be downloaded, and after the data file format has been converted into Extensible Markup Language or JavaScript with PHP or AWK, may also be

visualized using online plot tools: Ptplot [2] and jqPlot (jqPlot is under test internally and will be available soon).

## 2. Ionogram Data

Ionosondes measure reflected high-frequency radio signals from the layers in the ionosphere. Raw data files from ionosondes at Antarctic stations are automatically cleaned locally and then transferred to the IPS office in Sydney for further processing, such as regional ionospheric mapping and manual scaling.

Raw data files also are saved on compact discs and digital versatile discs locally, and shipped to the World Data Centre in Sydney each polar summer. Scaled (International Union of Radio Science parameters) ionospheric data are published via the IPS website [1]. Scaled ionogram data have been used by IPS to study the complicated dynamics of the polar cap ionosphere [3–5].

## 3. Magnetometer Data

Magnetometers measure variations in the geomagnetic field. At present, there are five

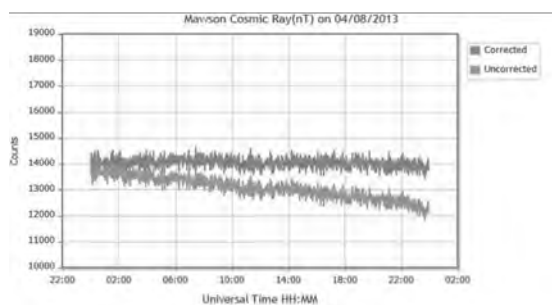
magnetometers operating at the four Australian Antarctic stations—out of the 11 magnetometer stations across IPSNET.

#### 4. Riometer Data

Riometers are 'Relative Ionospheric Opacity Meters', and measure the absorption of high-frequency radio waves by the lowest D-region of the ionosphere during geomagnetic storm events ('polar cap absorption'). IPS has riometer stations at all four Australian Antarctic stations.

#### 5. Cosmic Ray Data

The only cosmic ray station in the Antarctic is Mawson Station; apart from this, another cosmic ray station is located at Kingston, near Hobart in Tasmania. Figure 1 shows an example of cosmic ray data plotted with jqPlot.



**Figure 1.** Cosmic Ray data of Mawson on 4 August 2013.

#### 6. Ionospheric Scintillation Data

Ionospheric scintillation is a rapid fluctuation of radio-frequency signal phase and/or amplitude, generated as a signal passes through the ionosphere. Macquarie Island is the only Antarctic station out of the six ionospheric scintillation monitoring stations across IPSNET.

#### 7. Conclusions

The Antarctic region is a very important area for observing, monitoring, and detecting space weather, due to the focusing of energy onto the polar ionosphere by the near-vertical

geomagnetic field, which maps out the boundary of the Earth's magnetosphere with the solar wind. Disturbances in the ionosphere from the polar regions travel equator-wards and affect the mid-latitude ionosphere. The data collected is irreplaceable in ionospheric mapping, high-frequency forecasting, and other various applications of space weather research and forecasting.

**Acknowledgments.** We are deeply grateful to the colleagues who have made contributions to IPSNET's establishment and maintenance in the past decades. We are particularly thankful to Campbell Thomson, Lianne Grant, Richard Marshall, Mike Hyde, and Michael Terkildsen.

#### References

1. IPS Radio and Space Services, Bureau of Meteorology, Australia, [www.ips.gov.au](http://www.ips.gov.au) [accessed on: August 2013]
2. Wang, K., Yuile, C., The application of an online data visualization tool, Ptplot, in the World Data Centre (WDC) for Solar-Terrestrial Science (STS) in IPS Radio and Space Services, Australia. *Proceedings of 1st ICSU-WDS Conference Global Data for Global Science, Data Science Journal*, Kyoto, Japan, 86–90, 2012
3. Neudegg, D., Terkildsen, M., Wang, M., Long-term median foF2 variations in the Antarctic polar cap and the competing effects of solar EUV, magnetospheric precipitation and ionisation transport. *39th COSPAR (ICSU Committee on Space Research) Scientific Assembly 2012, Mysore, India, 2012*
4. Neudegg D., Wang, M., Francis, M., Bouya, Z., Kelly, A., Median foF2 variations in the Antarctic polar cap 1957–2001 and the effects of solar EUV, magnetosphere–ionosphere coupling and thermospheric transport. *Asia–Oceania Geophysics Society (AOGS) Conference, Brisbane, Australia, 2013*
5. Neudegg, D., Antarctic polar cap ionosphere and effects of solar EUV, magnetosphere–ionosphere coupling and thermospheric transport. *Australian Space Science Conference ASSC-13, UNSW, Sydney, Australia, 2013*

# TRANSMIT Prototype: Cross-institutional Network Approach from Geophysical Database to User Application for GNSS Science and Industry

**H. Sato<sup>1\*</sup>, N. Hlubeck<sup>1</sup>, M. Aquino<sup>2</sup>, P. Kieft<sup>2</sup>, E. Plakidis<sup>3</sup>, S. Mushini<sup>4</sup>, the TRANSMIT team<sup>5</sup>**

<sup>1\*</sup> German Aerospace Centre (DLR) Institute of Communications and Navigation, Neustrelitz, Germany

<sup>2</sup> Institute of Engineering Surveying and Space Geodesy, University of Nottingham, Nottingham, UK

<sup>3</sup> Istituto Nazionale di Geofisica e Vulcanologia (INGV), Rome, Italy

<sup>4</sup> IEEA, France

<sup>5</sup> [coordination@transmit-ionsphere.net](mailto:coordination@transmit-ionsphere.net)

Email: [hiroatsu.sato@dlr.de](mailto:hiroatsu.sato@dlr.de)

**Summary.** Training Research and Applications Network to Support the Mitigation of Ionospheric Threats (TRANSMIT) is a European research and training program for Global Navigation Satellite Systems data ranging from polar to lower latitude regions. The participants of the project use data archives from data hosting institutions, and process the data for end-users in science and industry. To demonstrate the results from multiple processors in a single platform, a cross-institutional data flow network model named the TRANSMIT prototype has been developed. This model and current results are presented here, with particular focus on the roles and responsibilities of hosting institutions, such as placement and storage of raw/processed data, operation of processors, and a user interface to connect the network components.

**Keywords.** Global Navigation Satellite Systems, ionosphere, scintillation, data flow, data sharing.

## 1. Introduction

Global Navigation Satellite Systems (GNSS) are a crucial component in countless modern systems, for example, in telecommunication, navigation, remote sensing, and precision timing. The main threat to reliable and safe operation of GNSS is the variable propagation conditions encountered by GNSS signals as they pass through the Earth's upper atmosphere (the ionosphere).

The formation of irregularities in the ionosphere causes scintillation phenomena characterized by sudden signal fluctuations that can cause GNSS operational outages. These phenomena are frequent in the polar and auroral regions, as well as in equatorial latitudes.

Here, a European research and training project is presented for mitigating the effects of the ionosphere. Within the frame of this project, a prototype web service will be implemented. This service is designed to highlight the results of the individual research projects in such a way, that a user of that web service can interact with the

software of those projects, namely, initiate reruns of the software on stored data with parameter sets of choice. Data storage, processing, and the user's web display are performed at different institutions. Thus, the whole service is strongly network centric.

## 2. Overview of Project Objectives

TRANSMIT [1] aims to expand the European knowledge base and critical mass in a multidisciplinary, intersectoral and industry-led training programme. It will also set up a prototype of the proposed Ionospheric Perturbation Detection and Monitoring network and associated service. One of the main objectives is to incorporate solutions into this system that respond to all end-user needs and that are applicable in all geographical regions of European interest (polar, high and mid- latitudes, equatorial regions).



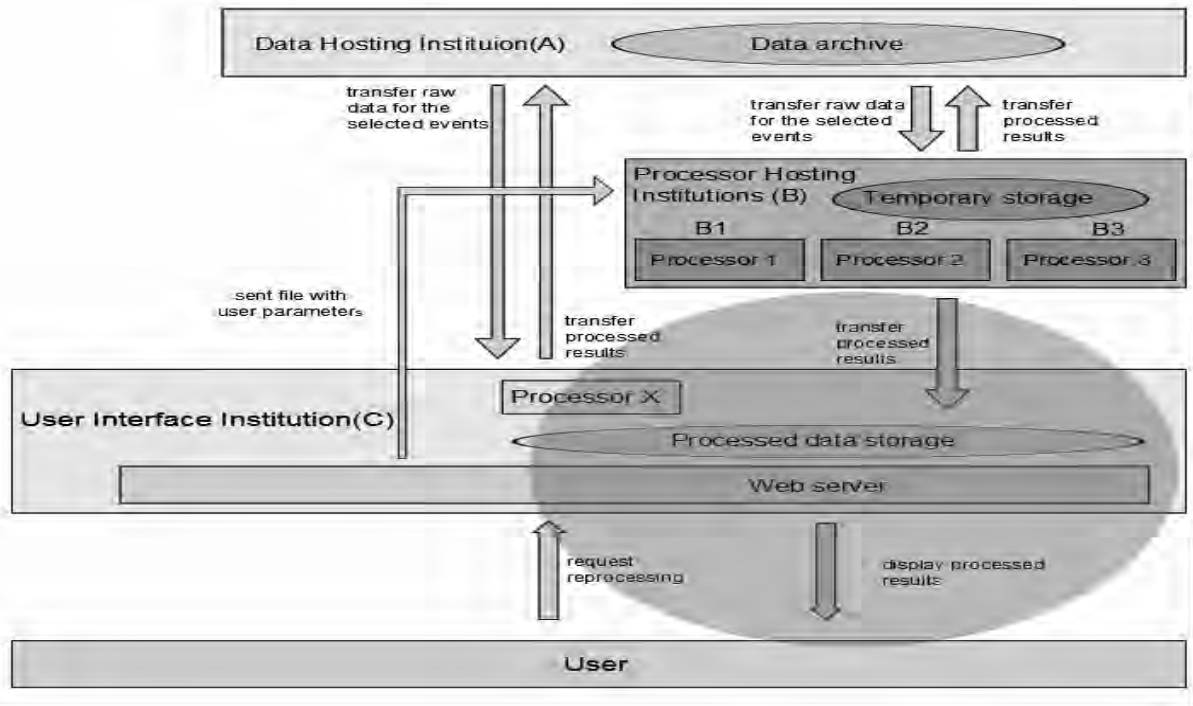


Figure 1. Data flow design for the TRANSMIT prototype.

### 3. TRANSMIT Prototype: Data Flow and Roles of Institutions

Research institutions in the project have developed processors that retrieve data from GNSS and other geophysical data archives. The processors can run individually or in a mutually connected manner, and provide results for GNSS-related scientific and industrial partners. The results are designed to be demonstrated in a single-user interface via a web service called the TRANSMIT prototype. Figure 1 shows the data flow chart and responsibilities of the hosting institutions in the system. During the current phase of the project, the system is formed by institutions divided into the following three categories:

- (A) Data archive institution
- (B) Processor hosting institutions (B1, B2...)
- (C) Web-user service institution

(A) is responsible for storing and archiving raw data, and also secures backup storage of processed results when user requests are transferred through the processors in (B). Institutions in (B) host processors, and provide storage space to save raw data while processing is performed. They continuously check for new

parameter files from (C). When new files become available, (B) institutions rerun the processors and upload the results to (C), which provides software for data transfer and store processed data. (C) sends a parameter file to (B), and displays the (re)processed data on a web page while validating the results.

### 4. Conclusions

The TRANSMIT prototype has been developed to demonstrate processed results from data archives through multiple institutions and processors via a web service. This prototype service is to be extended for mutual interaction of processors and use of the data system being developed by data management communities.

**Acknowledgments.** TRANSMIT is an FP7 Marie Curie Initial Training Network supported by European Commission.

### References

1. TRANSMIT Science Management Office, <http://www.nottingham.ac.uk/transmit> [accessed on: September 2013]

# Hydrometeorological Database (HMDB) for Practical Research in Ecology

**Alexander Novakovskiy<sup>1\*</sup>, Vladimir Elsakov<sup>1</sup>**

<sup>1\*</sup> *Institute of Biology, Komi science centre, Ural division, Russian academy of science,  
Kommunisticheskaya st., 28, Syktyvkar, Komi Republic, 167928, Russia  
Email: novakovsky@ib.komisc.ru*

**Summary.** The regional HydroMeteorological DataBase (HMDB) was designed for easy access to climate data via the Internet. The HMDB contains various climatic parameters (temperature, precipitation, pressure, humidity, wind strength and direction) for 168 meteorological stations in Russia and bordering countries for the periods of instrumental observations (over 100 years). The storage, queries and analytical blocks were also developed with Internet access.

**Keywords.** HMDB databases, climate, temperature, precipitation.

## 1. Introduction

One of the important tasks of modern ecological research in the Russian North is the analysis of the reactions of biodiversity parameters to climate fluctuations and monitoring of common trends of vegetation changes. Meteorological observations for the North of Russia show a steady trend of warming beginning since the 1970s and receiving the greatest speed of development in 1990s ('the modern warming of climate'). It is very important to use instrumental observations such as temperature, precipitation, and other climate characteristics collected on weather stations for accurate assessment of climate change. A large amount of these data are now available (the volume depending on the weather station), but the average duration of observations is 100–120 years and over. The volume and temporal period of collection makes these data very difficult to process manually. It is not appropriate to use spreadsheets (e.g., Microsoft Excel) for data manipulation. In our opinion, the best solution is to use a database management system, but this approach requires considerable effort to design the database structure and an appropriate user interface. A useful prospective direction in database utilization is the harnessing of Internet-

oriented databases, because in this case, users can access the data from anywhere in the world.

Hence, our goal was to create and test an information resource that is accessed via the Internet, linked to databases containing meteorological information (temperature, precipitation, pressure, humidity, wind strength and direction, etc.) of more than 100 years temporal coverage, with observations of daily resolution.

## 2. Entered Data

The basis of any information system is its population with concrete data. As of today, there is information about 168 weather stations from Russia and bordering countries in the HydroMeteorological DataBase—HMDB. For each station the average temperature and amount of precipitation had been entered. The typical observation time for each weather station was about 100 years. The earliest data observation was 1882, the last –2012.

The following open Internet resources were used to populate the database: <http://aisori.meteo.ru/ClimateR—All-Russian> Research Institute of Hydrometeorological Information—World Data Centre, <http://rp5.ru—weather forecast>).

### 3. Data Processing Algorithms

The set of most commonly used algorithms for meteorological data analysis were embedded in the HMDB system. For temperature data, the following algorithms were implemented: the calculation of average temperatures for different periods of time (ten-days, months, winter, summer, and annual); the sum of summer temperatures, dates of stable transition across 0, 5, 10 °C and the duration of these periods; the sum of effective temperatures (temperatures above 0, 5, 10 °C); the number of extreme hot (above 20, 25 and 30 °C) and extreme cold (below -20, -25 and -30 °C) days.

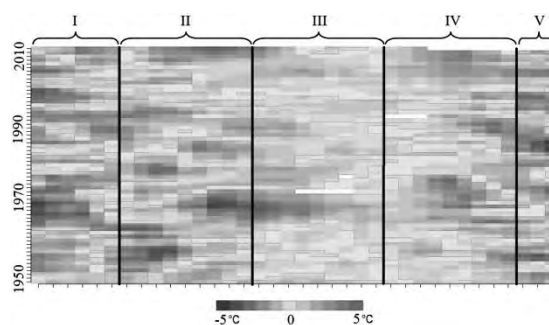
For the other climate parameters (precipitation, pressure, humidity) only monthly, summer, winter, and annual sums were implemented. For indicators of strength and wind direction, construction of the wind diagrams was provided.

### 4. Examples of using the HMDB

The developed database was used in conjunction with the R (open source) statistical package for analyses of temperature data of weather stations located in the Subarctic zone of Russia. In general, an increase of annual temperature over the past 20 years was observed for all weather stations. However, this increase is quite heterogeneous. From Figure 1, the greatest increase is noted for the spring (March, April) and autumn (September, October) months. In contrast, decrease of average temperatures was marked for late summer (August) and late winter (February).

Another focus of the database is a joint analysis of temperature series and satellite images time series (AVHRR, MODIS). This analysis permitted determination of the nature and extent of vegetation change in the arctic region, according to observed climatic fluctuations.

The reaction of northern Eurasian ecosystems to the 'warming' is heterogeneous and has regional diversities that are connected with the different increases of annual temperature, proximity of the ocean, the presence of permafrost, and altitude.



**Figure 1.** Deviation of annual decade temperature of Hoseda-hart meteorological station (during the period 1950–2012). The horizontal axis denotes ten-day periods and the vertical axis each year. The seasons are presented by numerals: I and V—winter, II—spring, III—summer, and IV—autumn.

Firstly, the results show that 56.6% of the Russian Subarctic region is characterized as an area of little change, and trends of biomass production are related to July temperature growth [1]. Most stable parameters were noted for vegetation cover in Central Siberia. An essential growth of productivity is observed for 19.9% of the territory. Mostly this growth is connected with an increase of shrub proportions in the vegetation cover. A decrease of productivity is noted for 23.5% of the area.

### 5. Conclusions

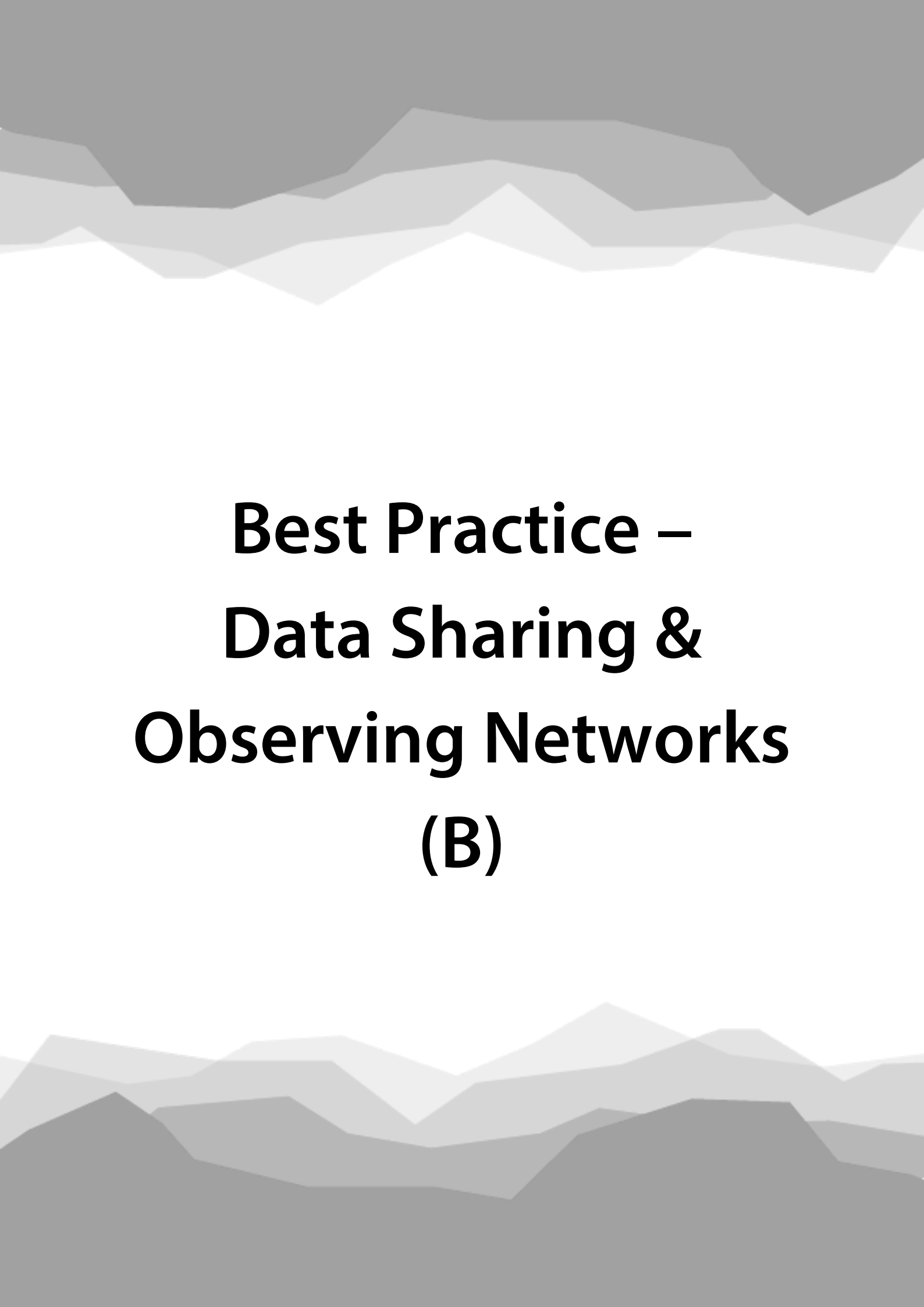
We hope this database will be useful for environmental researchers. The system is located at <http://ib.komisc.ru/climat> (in Russian only, at present; to obtain the password, please contact the authors).

**Acknowledgments.** This work was done within the frames of the UrD RAS scientific program 'Reaction of ecosystems of the European North and Western Siberia cryolithozone to climatic fluctuations of the last tens of years' (12-C-4-1018).

### References

1. Elsakov, V. V., The satellite data in chlorophyll index investigation at tundra communities. *The Earth research from Space*, 1, 60–70, 2013 (in Russian)





**Best Practice –  
Data Sharing &  
Observing Networks  
(B)**

# Current Data Practises in Polar Institutions and Networks: A Case Study with the HIACMS Project

**Paul A. Berkman<sup>1,2\*</sup>, Lawson W. Brigham<sup>3</sup>, Frank W. Davis<sup>1,4</sup>, Jean-Claude Gascard<sup>5</sup>,  
Benjamin S. Halpern<sup>1,4</sup>, Christine Provost<sup>5</sup>, Oran R. Young<sup>1</sup>**

<sup>1\*</sup> Bren School of Environmental Science and Management, University of California, Santa Barbara, CA 93106, USA

<sup>2\*</sup> Digital Integration Technology Limited (DigIn), 6 Caxton House, Broad Street, Great Cambourne, Cambridge, CB23 6JN, UK

<sup>3</sup> School of Natural Resources & Agricultural Sciences, University of Alaska Fairbank, PO Box 858140, Fairbanks, Alaska 99885, USA

<sup>4</sup> National Center for Ecological Analysis and Synthesis (NCEAS), 835 State Street, Suite 300, Santa Barbara, CA 93101, USA

<sup>5</sup> LOCEAN, Université Pierre et Marie Curie, 4, place Jussieu (CC 100), Tour 45/46 - 513A, 85252 Paris cedex 5, France

Email: berkman@bren.ucsb.edu

**Summary.** Holistic Integration for Arctic Coastal-Marine Sustainability (HIACMS) is a new three-year research programme that will tackle the themes of the Arctic Science Engineering and Education for Sustainability program. HIACMS is designed to add value through partnerships, and has already established links with projects in the US and Europe. Data management is a primary consideration of HIACMS—especially since data are typically from ‘unstructured sources—and Geographic Information System tools and approaches will be employed alongside more novel methods to aggregate and integrate diverse layers of Arctic coastal-marine data. These and other features of the HIACMS project will be introduced as a case-study to build accessible data practises for governance, infrastructure development and sustainable operations in Arctic coastal-marine environments.

**Keywords.** Big Data, relational schema, sustainability, infrastructure, integration.

## 1. Introduction

Interests are awakening globally to take advantage of extensive energy, shipping, fishing, and tourism opportunities associated with diminishing sea ice in the Arctic Ocean. To address these diverse and expanding interests—Holistic Integration for Arctic Coastal-Marine Sustainability (HIACMS) is a new three-year research project to develop and demonstrate an international, interdisciplinary and inclusive process that will enhance the practice of governance for sustainability in Arctic coastal-marine systems, balancing:

- National interests and common interests;
- Environmental protection, social equity and economic prosperity; and
- Needs of present and future generations.

To achieve these holistic objectives—as a component of the Arctic Science Engineering and Education for Sustainability (ArcSEES) program coordinated by the US National Science Foundation and French Centre Nationale de la Recherche Scientifique [1]—HIACMS will carry out tasks addressing the four themes of the ArcSEES program:

1. Natural and Living Environment;
2. Built Environment;
3. Natural Resource Development; and
4. Governance

## 2. Collaborative Links

To enhance its cost-effectiveness, HIACMS also has established links to the Study of Environmental Change (SEARCH) [2] and Arctic

Climate Change, Economy and Society (ACCESS) [3] projects that are supported extensively within the United States and Europe, respectively.

In addition, HIACMS is designed to add value through partnerships, especially those involving interdisciplinary data aggregation and integration to reveal sustainable development scenarios around the Arctic Ocean. Data aggregation requires open collaboration, stimulating individuals to continuously update the resource collection with 'creative commons' approaches, such as Wikipedia. Data integration requires user-defined manipulation of digital resources, beyond programmer-constrained solutions, based on individual lines of inquiry. The data involved with HIACMS will include numeric and experimental results generated by Arctic observing networks (e.g., [4]) and research platforms, as well as published syntheses and other natural language products that together reflect the 'structured' and 'unstructured' resources underlying the 'Big Data' challenges we face today.

### 3. Data Management

A central consideration is data management, which currently includes library and archival strategies that involve the content of resources, as well as the context of records, respectively. Standardized data management includes metadata, markup, and database strategies. However, these strategies are limited in their application; recognizing that more than 90% of the digital information produced is considered to be 'unstructured' and unmanaged. More to the point, all resources with meaning—whether hardcopy or digital—also have structure, and the unique advantage of digital over hardcopy is that the structure can be manipulated automatically and objectively for the purposes of information management and knowledge discovery.

### 4. Tools and Approaches

HIACMS will involve Geographic Information System (GIS) tools and approaches, including marine spatial planning and ecosystem-based

management, to aggregate and integrate diverse layers of Arctic coastal-marine data, taking into account existing GIS platforms and information networks (see Table 1) on a pan-Arctic basis. HIACMS also will utilize innovations that leverage the inherent structure of digital resources, which been deployed earlier through the National Science Digital Library [5] in conjunction with the Committee on Data for Science and Technology [6] and International Research on Permanent Authentic Records in Electronic Systems (InterPARES) programme [7].

### 5. Conclusions

HIACMS (Holistic Integration for Arctic Coastal-Marine Sustainability) is a new three-year research project with funding from the United States and France, involving 'international, interdisciplinary and inclusive' networking (Table 1). HIACMS provides a case study of information integration strategies to reveal decision-maker options for sustainable operations in the Arctic Ocean.

### References

1. National Science Foundation, [www.nsf.gov/pubs/2012/nsf12553/nsf12553.htm](http://www.nsf.gov/pubs/2012/nsf12553/nsf12553.htm) [accessed on: September 2013]
2. Arctic Research Consortium of the United States, [www.arcus.org/search](http://www.arcus.org/search) [accessed on: September 2013]
3. Arctic Climate Change, Economy and Society, [www.access-eu.org](http://www.access-eu.org) [accessed on: September 2013]
4. Arctic Observing Summit, [www.arcticobservingsummit.org](http://www.arcticobservingsummit.org) [accessed on: September 2013]
5. National Science Digital Library, [www.nsdlib.org](http://www.nsdlib.org) [accessed on: September 2013]
6. Berkman, P. A., Morgan, G. J., Moore, R., and Hamidzadeh, B. Automated Granularity to Integrate Digital Information: The "Antarctic Treaty Searchable Database" Case Study. *Data Science Journal*, 5, 84–99, 2006
7. International Research on Permanent Authentic Records (InterPARES) Project, [www.interpares.org/](http://www.interpares.org/) [accessed on: September 2013]
8. Berkman, P. A. and Vylegzhanin, A. N. (Eds.), *Environmental Security in the Arctic Ocean*. Springer, Dordrecht, 2013

Table 1. International participation in Arctic organizations and information networks.<sup>1</sup>

STATES <sup>2</sup>	ARCTIC ORGANIZATION <sup>3,4</sup>																
	AC <sup>5</sup>	AMEC	BEAC <sup>6</sup>	FARO	IASC	MOPP	NACG	NAFO <sup>7</sup>	INC	NEAF <sup>7</sup>	INF	OSPA	PB	SAR	SCAP	SPIT	NATO <sup>8</sup>
Afghanistan																X	
Albania																X	
Argentina																X	
Australia																X	
Austria																X	
Belgium <sup>9</sup>							X					X				X	
Bulgaria <sup>9</sup>								X								X	
Canada	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Chile																X	
China				X	X					X						X	
Croatia																	X
Cuba								X									
Czech Republic <sup>9</sup>										X						X	
Denmark <sup>9,10</sup>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Dominican Republic																X	
Egypt																X	
Estonia <sup>9</sup>							X									X	
Finland <sup>9</sup>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
France <sup>9</sup>	X	X	X	X	X		X				X					X	
Germany <sup>9</sup>	X	X	X	X	X		X				X					X	
Greece <sup>9</sup>																X	
Hungary <sup>9</sup>																X	
Iceland	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
India																X	
Ireland <sup>9</sup>							X					X				X	



STATES <sup>2</sup>	ARCTIC ORGANIZATION <sup>3,4</sup>																
	AC <sup>5</sup>	AMEC	BEAC <sup>6</sup>	FARO	IASC	MOPP	NACG	NAFO <sup>7</sup>	NC	NEAF <sup>7</sup>	NF	OSPA	PB	SAR	SCAP	SPIT	NATO <sup>8</sup>
Italy <sup>9</sup>			X	X	X											X	
Japan			X	X	X		X									X	
Latvia <sup>9</sup>							X										X
Lithuania							X										X
Luxembourg <sup>9</sup>											X						X
Monaco															X		
Netherlands <sup>9</sup>	X		X	X	X		X				X				X		X
New Zealand															X		
Norway	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Poland <sup>9</sup>	X		X	X			X								X		X
Portugal <sup>9</sup>							X			X					X		X
Republic of Korea (South)				X	X			X								X	
Romania																X	
Russian Federation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Saudi Arabia															X		
Serbia															X		
Slovakia																	X
Slovenia																	X
South Africa															X		
Spain <sup>9</sup>	X						X				X				X		X
Sweden <sup>9</sup>	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Switzerland												X			X		
Turkey																	X
Ukraine								X									
United Kingdom <sup>9</sup>	X	X	X	X	X		X				X					X	X
United States	X	X	X	X	X	X	X	X			X		X	X	X	X	X

STATES <sup>2</sup>	ARCTIC ORGANIZATION <sup>3,4</sup>																
	AC <sup>5</sup>	AMEC	BEAC <sup>6</sup>	FARO	IASC	MOPP	NACG	NAFO <sup>7</sup>	NC	NEAF <sup>7</sup>	NF	OSPAA	PB	SAR	SCAP	SPIT	NATO <sup>8</sup>
Venezuela																X	
<b>Number of States</b>	20	4	15	17	16	8	20	13	5	8	8	15	5	8	8	42	28

<sup>1</sup> Updated from [8]

<sup>2</sup> Among the 52 states in this table, the eight Arctic states are highlighted.

<sup>3</sup> Highlighted organizations include all of the Arctic states.

<sup>4</sup> AC (1996 Arctic Council); AMEC (1996 Arctic Military Environmental Cooperation Programme); BEAC (1993 Barents Euro-Arctic Council); FARO (1998 Forum of Arctic Research Operators); IASC (1990 International Arctic Science Committee); MOPP (2013 Agreement on Cooperation on Marine Oil Pollution, Preparedness and Response in the Arctic); NACG (2007 North Atlantic Coast Guard Forum); NAFO (1978 Convention on Future Multilateral Cooperation in the Northwest Atlantic Fisheries); NATO (1949 North Atlantic Treaty); NC (1952 Nordic Council); NEAF (1980 Convention on Future Multilateral Cooperation in North-East Atlantic Fisheries); NF (1991 Northern Forum); OSPA (1992 Convention for the Protection of the Marine Environment of the North-East Atlantic); PB (1973 Agreement on the Conservation of Polar Bears); SAR (2011 Agreement on Cooperation on Aeronautical and Maritime Search and Rescue in the Arctic); SCAP (1994 Standing Committee of the Conference of Arctic Parliamentarians), SPIT (1920 Treaty Concerning the Archipelago of Spitsbergen, and Protocol).

<sup>5</sup> Arctic Council (AC)—In addition to the eight Arctic Member States and the twelve non-Arctic Observer states listed in the table, there are six Permanent Participants from Arctic indigenous peoples organizations (Arctic Athabaskan Council, Aleut International Association, Gwich'in Council International, Inuit Circumpolar Council, Russian Arctic Indigenous Peoples of the North, and Saami Council). The Arctic Council also involves nine intergovernmental and inter-parliamentary organizations (International Federation of Red Cross & Red Crescent Societies, International Union for the Conservation of Nature, Nordic Council of Ministers, Nordic Environment Finance Corporation, North Atlantic Marine Mammal Commission, Standing Committee of the Parliamentarians of the Arctic Region, United Nations Economic Commission for Europe, United Nations Development Program, United Nations Environment Program) as well as eleven non-governmental organizations (Advisory Committee on Protection of the Seas, Arctic Circumpolar Gateway, Association of World Reindeer Herders, Circumpolar Conservation Union, International Arctic Science Committee, International Arctic Social Sciences Association, International Union for Circumpolar Health, International Work Group for Indigenous Affairs, Northern Forum, University of the Arctic, World Wide Fund for Nature-Global Arctic Program). The European Union has applied to the Arctic Council for Permanent Observer status. In addition, the Arctic Council involves expert groups and task forces along with its six working groups: Arctic Contaminant Action Program (ACAP), Arctic Monitoring and Assessment Program (AMAP), Conservation of Arctic Fauna and Flora (CAFF), Emergency Prevention, Preparedness and Response (EPPR), Protection of the Marine Environment (PAME) and Sustainable Development Working Group (SDWG).

<sup>6</sup> Barents Euro-Arctic Council—Permanent Members (Denmark, Finland, Iceland, Norway, Sweden and Russian Federation with the European Commission) and other states are observers.

<sup>7</sup> Includes European Economic Community or European Union.

<sup>8</sup> It is a question whether NATO, like other North Atlantic organizations (e.g., NACG, NAFO, NEAF, OSPA) has an Arctic remit.

<sup>9</sup> Member of European Union.

<sup>10</sup> Includes Greenland (which is not a member of the European Union) and the Faroe Islands as autonomous areas.

# EXPEDITION: An Integrated Approach to Expose Expedition Information and Research Results

**Ana Macario<sup>1</sup>, Roland Koppe<sup>1\*</sup>, Hans Pfeiffenberger<sup>1</sup>**

<sup>1\*</sup> Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Am Handelshafen 12, Bremerhaven, 27570, Germany  
Email: Roland.Koppe@awi.de

**Summary.** The portal EXPEDITION offers an integrative ‘one-stop shop’ framework for discovery and reuse of scientific content originating from research platforms operated by the Alfred Wegener Institute (AWI). This information sharing framework is designed for interoperability and can be extended to various information systems worldwide. The framework is based on open technologies and access is freely available for scientists, funding agencies, and the public. Because AWI’s research is known to be focused on both polar regions, access to various ready-to-use data products from the Arctic Ocean, Southern Ocean and Antarctic, as well as AWI-operated observing networks, will be offered.

**Keywords.** Data management, data sharing, information integration.

## 1. Introduction

The first data and publication discovery services offered by the Alfred Wegener Institute (AWI) with a focus on the polar regions dates back about 10 years [1]. With the recently refactored EXPEDITION portal [2, 3], a data access framework has been launched based on international standards defined by the International Organization and Open Geospatial Consortium, and supporting interoperability with various international initiatives (e.g., SeaDataNet II, EUROFLEET II, etc.).

As depicted in Figure 1, EXPEDITION integrates content from distinct research platforms (vessels, ocean-based stations, land-based stations and aircraft) into one portal. Map-supported browsing, faceted searching, and interactive relationship graphs are some of the current functionality offered to support efficient content discovery.

The EXPEDITION framework is based on three core components:

- Expedition Catalogue—contains metadata describing each AWI research platform and respective expeditions.
- DSHIP—the in-progress data acquisition system on board of, for example, Polarstern

provides access to raw data from vessel-mounted devices and event logging (station books).

- Generalized track lines—used to display map-based validated routes, as well as overlays with geo-referenced data.

## 2. Information Discovery and Access

In addition to its core components, EXPEDITION offers comprehensive discovery and access to various information systems containing expedition-related results and products. While to the main information providers at this time are PANGAEA (validated primary data) and EPIC (reports and publications), extensions to other providers are possible (e.g., access to AWI bathymetry database is planned).

A number of interoperability and core services have been developed to facilitate content discovery and access:

Common vocabularies and Gazetteers—SeaDataNet vocabularies for describing ship operation status and activity, topic/theme/discipline and ocean regions (Gazetteer) are used in the metadata catalogue and in the discovery services.

Near real-time data—raw (non-validated) data

from meteorological station and thermosalinograph on board of AWI vessels, as well as 24-h sea ice coverage images, are graphically displayed on the portal homepage.

### 3. Future Work

The EXPEDITION catalogue is currently being extended so as to fulfill concrete requirements from AWI's Directorate and platform coordination. These requirements are, amongst others, the need to track scientific output associated with various measurements on board and map support to track line planning of future cruises in order to optimize resources. For example, new tracks should be placed in areas where the vessels have not yet been or over old tracks so as to be able to search for trends or extend knowledge about a specific region.

In addition, EXPEDITION will soon offer semi-automatic generation of SeaDataNet Cruise Summary Reports. Discipline-oriented collections of Global Information System based map layers are also planned (e.g., ready-to-use data products in PANGAEA such as gridded sea ice concentration).

We are also planning to address the complex issue of accessing and visualizing Big Data for which unrestricted access is not practicable (e.g., bathymetry, multi-channel seismic). As proof-of-concept, we plan to use the existing multibeam bathymetry database as a provider for EXPEDITION. Other Big Data candidates are audio and video material from AWI land-based and ocean-based stations.

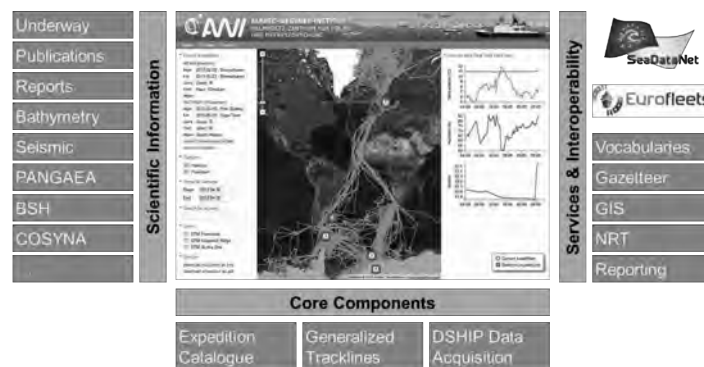
### 4. Conclusions

The new EXPEDITION portal offers a 'one-stop shop' framework for discovery and access of scientific content related to AWI research platforms, in particular, those in the polar regions. A similar open-access framework designed for interoperability could be extended to other information systems worldwide, which would help bringing together scientific content from the polar regions.

Given that a wide range of statistics could additionally be generated using the harvested information, monitoring and reporting of publications and data resulting from research programmes—institutional or global, such as IPY—should be straightforward to implement.

### References

1. Macario, A., Pfeiffenberger, H., Reinke, M., A Discovery Service for Knowledge Related to Research Platforms. XXVIII SCAR Open Science Conference 2004, Bremen, Germany, 25–31 2004
2. EXPEDITION, <http://expedition.awi.de> [accessed on: September 2013]
3. Koppe, R. and MANIDA Scientific Team, MaNIDA: Integration of marine expedition information, data and publications: Data Portal of German Marine Research. EGU General Assembly 2013, Vienna, Austria, hdl:10013/epic.4117, 2013



**Figure 1.** EXPEDITION architectural elements. Scientific information (left) is harvested from various sources and integrated via core components (bottom). Additional services and interoperability issues are embedded (right).

# *Assembling an Arctic Ocean Boundary Monitoring Array*

**Takamasa Tsubouchi<sup>1\*</sup>, Sheldon Bacon<sup>1</sup>**

<sup>1\*</sup> *National Oceanography Centre, Southampton, European Way, Southampton, SO14 3ZH, UK*  
Email: tt2r07@noc.ac.uk

**Summary.** The Arctic Ocean boundary is defined as four ocean gateways: Fram, Davis and Bering Straits, and the Barents Sea Opening. Each gateway contains a suite of long-term moored installations, and coast-to-coast hydrographic measurements are regularly made at these. The approach to Arctic Ocean boundary measurements outlined in this paper is generating significant scientific outcomes: the first quasi-synoptic net heat and freshwater transports in summer 2005, the first directly-estimated quasi-synoptic inorganic nutrient budget, and a dissolved inorganic carbon budget. However, Arctic data are not always easy to access; there is no pan-Arctic data access agreement, and the IPY Data Policy was never formally enacted. As a result, it is not unusual for originators to retain control, even some years after the making of the measurements. Arctic science would progress faster with more open access.

**Keywords.** Arctic Ocean boundary, hydrographic data, mooring data, IPY Data Policy.

## **1. Introduction**

Recently, the UK Natural Environment Research Council has largely been delivering funding for research in the Arctic via the 'Research Programme' mode. UK Arctic marine physics—encompassing both sea ice and ocean—has been developed first under the Arctic Synoptic Basin-wide Observations (ASBO) project (2006–2010; Principle Investigators (PIs): Prof. Laxon, Dr. Bacon), and currently under the The Environment of the Arctic: Climate, Ocean and Sea Ice (TEA-COSI) project (2011–2015; Dr. Bacon).

At the heart of these two projects is the perception that Arctic ice and ocean fluxes could, for the first time, be objectively determined using inverse modelling. The boundary measurements define a closed box (including coastline), enabling application of conservation constraints. This, in turn, means that real oceanic transports (i.e., independent of arbitrary reference values) can be calculated.

## **2. Science**

The Arctic Ocean boundary is defined as four ocean gateways (plus coastline): Fram, Davis and Bering Straits, and the Barents Sea Opening (BSO). Each gateway contains a suite of long-term

moored installations, and coast-to-coast hydrographic measurements are regularly made at these.

The approach to Arctic Ocean boundary measurements in the ASBO and TEA-COSI projects is generating significant scientific outcomes: the first quasi-synoptic net heat and freshwater transports in a single month from summer 2005 [1], the first directly-estimated quasi-synoptic inorganic nutrient budget [2], and a dissolved inorganic carbon budget [3].

Current work is to define a full annual (summer-to-summer) cycle of monthly net heat and fresh water (FW) transports during 2005–06. This is ongoing work under the TEA-COSI project. The main data sources are direct moored array observations of temperature, salinity, and velocity obtained by 135 moored instruments. Also considered are sea ice export and sea surface current variability across the defined boundary, based on satellite measurements [4]. An important goal of this particular project is not just to calculate an annual cycle of fluxes, but also to determine the adequacy of the instrumental configuration, as presently deployed, to the task. Are any parts of the boundary in need of additional instrumentation?

### 3. Data

The ambition was (and is) to assemble disparate (meaning separate and distinct) datasets from around the Arctic boundary into a coherent Arctic boundary monitoring array. The data were generated by and obtained from a number of different PIs, institutions, and agencies. Indeed, five institutes in the world contribute to sustain the Arctic boundary observation lines: the University of Washington (UW) in the US for Davis Strait (PI: Dr. Lee) and the US side of Bering Strait (PI: Dr. Woodgate); the Norwegian Polar Institute (NPI) in Norway (PI: Dr. Dodd) and Alfred Wegener Institute (AWI) in Germany (PI: Prof. Schauer) for the western and eastern side of Fram Strait, respectively; the Institute of Marine Research (IMR) in Norway (PI: Dr. Ingvaldsen) for BSO, the University of Alaska Fairbanks (UAF) in the US for the Russian side of Bering Strait (PI: Prof. Weingartner). Data accessibility is different across the institutes and can be categorized as follows:

- Open-access databases (e.g., International Council for the Exploration of the Sea, World Ocean Database);
- PIs who held their own data, but maintained an open-access policy;
- PIs who held their own data, and where negotiation was required to enable access.

The data accessibility statistics are summarized in Table 1. Data was hardest to access when starting to gather conductivity, temperature, and depth (CTD) and mooring data in summer 2005 for the first heat and FW transports [1]. Ease of data access has been increasing as time has passed and the value of this approach has been recognised (i.e., gathering data around the Arctic Ocean

boundary to draw a comprehensive picture [2-3].

### 4. Conclusions

Arctic data are not always easy to access; there is no pan-Arctic data access agreement, and the IPY Data Policy was never formally enacted. It is not unusual for originators to retain control, even some years after the taking measurements. This can arise from the personal to the national level. Arctic science would progress faster with greater open access (while still recognizing a normal degree of priority for data originators).

**Acknowledgments.** The Arctic boundary CTD and mooring data are provided by Dr. Lee at UW (Davis Strait), Dr. Hansen at NPI and Dr. Fahrbach at AWI (Fram Strait), Dr. Ingvaldsen at IMR (BSO) and Dr. Woodgate at UW and Prof. Weingartner at UAF (Bering Strait).

### References

1. Tsubouchi, T., Bacon, S., Naveira-Garabato, A. C., Aksenov, Y., Laxon, S. W., Fahrbach, E., Beszczynska-Moller, A., Hansen, E., Lee, C. M., Ingvaldsen R. B., The Arctic Ocean in summer: a quasi-synoptic inverse estimate of boundary fluxes and water mass transformation. *J. Geophys. Res.*, 117, 28, 2012
2. Torres-Valdes, S., Tsubouchi, T., Bacon, S., Naveira-Garabato, A. C., Sanders, R., McLaughlin, F. A., Petrie, B., Kattner, G., Azetsu-Scott, K., Whitley, T. E., Export of nutrients from the Arctic Ocean, *J. Geophys. Res.*, 118, 20, 2013
3. MacGilchrist, G. A., Naveira-Garabato, A. C., Tsubouchi, T., Bacon, S., Torres-Valdes, S., Azetsu-Scott, K., Arctic Ocean carbon transports. Submitted to *Deep Sea Res.*, 2013
4. Giles, K. A., Laxon, S. W., Ridout, A. L., Wingham, D. J., Bacon, S. Western Arctic Ocean freshwater storage increased by wind-driven spin-up of the Beaufort Gyre. *Nature Geoscience*, 2012

**Table 1.** Summary of data accessibility for each outcome

(%)	Physical transports		Biogeochemical transports			Annual cycle	
	CTD	Mooring	Nutrients	DIC	$\delta_{18}O$	CTD	Mooring
(a)	40	20	25	75	25	40	20
(b)	0	0	25	25	50	40	60
(c)	60	80	50	0	25	20	20

# Building on IPY: Discovering Interdisciplinary Data through Federated Search

**Lynn Yarmey<sup>1</sup>, Siri Jodha Singh Khalsa<sup>1\*</sup>**

<sup>1\*</sup> *National Snow and Ice Data Center, University of Colorado, Boulder, CO, 80309, USA*  
Email: lynn.yarmey@colorado.edu

**Summary.** The rich legacy of the International Polar Year (IPY) includes advances in open data and meaningful progress towards interoperability of data, systems, and standards. Building on the IPY notion of the ‘union catalogue’, the Advanced Cooperative Arctic Data and Information Service has partnered with Earthcube and the Earth and Space Science Informatics Laboratory to support and promote interdisciplinary data discovery and reuse through federated data search. Realized by metadata brokering technologies and the growing adoption of international metadata standards, federated data search welcomes the diversity in Arctic data and increased efforts in data management. Recognizing the value of data expertise associated with targeted community data repositories, federated search enables specialized data holdings to be discovered by broader audiences. Federated data search compliments the role of metadata registries such as the Global Change Master Directory, and provides a useful mechanism for interoperability across the Arctic web-of-repositories. Here, the technologies of metadata brokering and federated search are presented, the experiences of putting these technologies into practice are discussed, and the challenges and next steps are articulated for this valuable, user-centric continuation of the IPY legacy.

**Keywords.** Information infrastructure, metadata and systems interoperability, federated data search, metadata brokering, interdisciplinary data discovery.

## 1. Introduction

Modern polar research benefits from the complex history of Arctic and Antarctic science. Observations over time, and across geospatial scales and domains, are crucial for setting baselines and understanding the rapid changes in these key regions. The International Polar Year (IPY) played a critical role in promoting data sharing and offering central coordination for observations. The IPY Data and Information Service (IPYDIS) identified early the unique needs of interdisciplinary international research, and introduced the ‘rigorous yet collaborative’ approach of a *union catalogue* [1]. This catalogue would allow for disparate search strategies and interfaces to access IPY data and resources. The experiences from IPYDIS’ union catalogue design and implementation are well documented [2].

Despite these significant steps, data discovery and access challenges remain. Many data

repositories—funded by different agencies, nations, and operational and industry groups—maintain separate catalogues and systems. Given this complex and diverse landscape of resources, those looking to reuse data need a good deal of insider knowledge, time, and luck to discover data. Leveraging IPY contributions, as well as recent technical advances, metadata brokering addresses this problem by mediating across distributed systems. Metadata brokering tools work with different access mechanisms, update schedules, and metadata standards to enable users to search many catalogues simultaneously. Federated search portals, such as the Advanced Cooperative Arctic Data and Information Service (ACADIS) Arctic Data Explorer, then leverage brokered metadata so as to offer searching across many distributed sources. Federated search presents an opportunity to advance polar cyberinfrastructure through coordination of standards, infrastructure, and resources to

support critical polar science.

## 2. Discussion

Federated data search honours the rich legacy of polar research by allowing diversity in contributing repositories. Polar data repositories and services often form in response to community-specific needs; monitoring stations have different service requirements than seasonal biological surveys or one-time soil moisture projects. Communities have different metadata and data content standards and encodings, vocabularies, and access protocols. Metadata brokering allows all of this diversity, and minimizes the burden of participation. Contributing repositories are not required to send data to a central system nor manage their data in a pre-defined manner. Federated data search through metadata brokering bridges the distributed legacy of polar science.

### 2.1 Technology—Metadata brokering

Brokering technologies can access metadata through a variety of exchange protocols (e.g., Open Archives Initiative Protocol for Metadata Harvesting or Thematic Real-time Environmental Distributed Data Services) and aggregate them into a central system. Alternatively, queries can be distributed ‘on the fly’ and results aggregated. Metadata are translated from native standards into a common schema using crosswalks. Once harmonized, a single search query formulates a set of returns based on a defined relevance ranking algorithm. Resource links in the metadata enable users to view the original metadata record for a dataset of interest, with the potential to build additional functionality—such as download or transformation—based on web services. The Earth and Space Science Informatics Laboratory’s GICat tool is the brokering application used by the ACADIS Arctic Data Explorer.

### 2.2 Experiences with federated data search

The ACADIS Arctic Data Explorer experience has shown that to have a successful, sustainable, open source metadata brokering product, a few points should be considered. Most important is recognizing that the biggest challenges in

metadata brokering are not necessarily technical. For example, relationships among developers across participating institutions are key, and consistent, honest, and timely communication with stakeholders is required. Stakeholders should include representatives from groups including: scientists providing data, scientists searching for data, developers, data curators, system architects, technical operations staff, project managers, metadata experts, web usability experts, other federated data search efforts, and funders. In addition to working closely with stakeholders, it is vital to identify the primary audience for the brokered application and proactively seek usability feedback from them early and often.

### 2.3 Challenges and next steps

Significant progress has been made towards comprehensive federated data search, though challenges remain. Experiences thus far have highlighted additional technical, social, and socio-technical elements necessary to achieve scientific goals. Examples include: long-term maintenance and resourcing system scaling, lack of governance structures, meaningful relevance ranking of thousands of search results to help searchers find what they need, standardization of key metadata content and best practices across communities, and building trust into systems.

## 3. Conclusions

Federated data search made possible by metadata brokering technologies begins to address the problem of finding data of interest in a myriad of diverse, isolated repositories. The ACADIS Arctic Data Explorer experiences with federated data search in the Arctic can inform transpolar data discovery and access efforts. Ongoing communication, coordination, research, and development are needed to address challenges.

**Acknowledgments.** ACADIS (NSF Award ARC 1016048) partners the University Corporation for Atmospheric Research with the National Center for Atmospheric Research and the National Snow



and Ice Data Center.

## References

1. Parsons, M. A., et al., International Polar Year Data Management Workshop 2006. *Glaciological Data Series*, GD-33, 2006
2. Parsons, M. A., Godøy, Ø., LeDrew, E., de Bruin, T. F., Danis, B., Tomlinson, S., Carlson, D., A Conceptual Framework for Managing Very Diverse Data for Complex Interdisciplinary Science. *Journal of Information Science*, 37(6), 2011

# GEO Cold Regions—The Interface to GEOSS for Polar and Mountainous Cold Region Observations

**Yubao Qiu**<sup>1\*</sup>

<sup>1\*</sup> *Group on Earth Observations, 7 bis, avenue de la Paix, Geneva, CH-1211 Geneva 2, Case postale 2300, Switzerland*  
Email: yqiu@geosec.org

**Summary.** Society is dependent on Earth Observation data to understand changes to the Earth System. The Group on Earth Observations (GEO) was created to support decision making by exploiting the growing potential of Earth Observations through international collaboration. In particular, the recently created GEO Cold Regions initiative works with collaborators on many levels and has the aim of coordinating observations of Earth's fragile 'cold regions'. The issues in these regions are strongly related to the nine 'Societal Benefit Areas' of the Global Earth Observation System of Systems (GEOSS). Thus, GEO Cold Regions is a cross-cutting programme that is able to utilize the brokerage and interoperability of the GEOSS Common Interface.

**Keywords.** Cold regions, GEOSS, data interoperability, polar, mountains, coordination.

## 1. Introduction

The Earth system has undergone severe global environment challenges—especially over the last ten years—along with the specific challenge of global climate change, resulting in sea-level rise and increased frequency of extreme weather events. Society is dependent on Earth Observation data by in-situ, satellite, and modelling methods to understand the changing factors of the Earth System. Earth Observations promote the study, monitoring, assessment, prediction, and information management activities used to describe and understand the changes in the global environment that may alter the capacity of the Earth to sustain life.

## 2. Group on Earth Observations

The Group on Earth Observations (GEO) was launched in response to calls for action by the 2002 World Summit on Sustainable Development and by the G8 (Group of Eight) leading industrialized countries. These high-level meetings recognized that international collaboration is essential for exploiting the growing potential of Earth Observations to

support decision making in an increasingly complex and environmentally stressed world.

## 3. GEO Cold Regions

Recently, GEO's aim has been a global effort to tackle Earth's fragile 'cold regions' by focusing on coordination of observations over the Poles (North Pole, South Pole, and the Third Pole of the Tibetan Plateau and Alpines mountain areas). The GEO Cold Regions initiative keeps these cold regions as the target area for observations, which relate strongly to frozen water in its various forms, and are associated with the influence of water (cryosphere), ecosystem, biodiversity, health, energy, disaster, climate, weather, and agriculture issues. The issues are thus relevant to the nine 'Societal Benefit Areas' of the Global Earth Observation System of Systems (GEOSS), and therein the implementation of 26 associated tasks (60 components). Hence, GEO Cold Regions has a cross-cutting nature within the GEOSS framework.

Under the GEO umbrella, GEO Cold Regions works with collaborators on many levels: existing international organizations/networks, national or

international programmes, national/regional GEOSS efforts, projects, observation stations, individual experts, and the private sectors.

#### **4. Vision**

The vision of GEO Cold Region's coordination is to:

- Highlight the cold regions as the most fragile ecosystems on Earth;
- Fill observation gaps in remote cold regions by satellite and in-situ data through the global platform of GEO;
- Implement an observation strategy, provide a proactive framework for the development of information-related services for the 'Future Earth' research initiative, and assess the

sustainable development ability addressed in Rio+20;

- Strengthen partnerships with policy-makers, stakeholders, and funders across the cold region ecological and engineering fields;
- Facilitate global coordination of national infrastructure and capacity building, especially over the nine Societal Benefit Areas.

#### **5. Conclusions**

GEO Cold Regions is characterized by the brokerage and interoperability offered via the GEOSS Common Interface, and the aim is to provide services to the Societal Benefit Areas by enabling data access for free or at minimum cost.





# **Data Centre/Service Provider Accreditation**

# Certification and Accreditation of Data Centres and Services

**Michael Diepenbroek<sup>1\*</sup>, Ingrid Dillo<sup>2</sup>, Lesley Rickards<sup>3</sup>, Mary Vardigan<sup>4</sup>**

<sup>1\*</sup> Center for Marine Environmental Science (MARUM), Bremen University, Hochschulring 18, Bremen, 28359, Germany

<sup>2</sup> Data Archiving and Networked Services (DANS), Anna van Saksenlaan 10  
2593 HT Den Haag, The Netherlands

<sup>3</sup> Permanent Service for Mean Sea Level (PSMSL), National Oceanography Centre, Joseph Proudman Building,  
6 Brownlow Street, Liverpool L3 5DA, UK

<sup>4</sup> Inter-university Consortium for Political and Social Research (ICPSR), University of Michigan  
Institute for Social Research, P.O. Box 1248, Ann Arbor, MI 48106-1248, USA  
Email: mdiepenbroek@pangaea.de

**Summary.** Certification is the base for accreditation of data centres and services. Certification and accreditation is beneficial for researchers, funders, as well as for the accredited facility itself. Different approaches have been accomplished during the last decade. Community independent, well organized, and practically applied are the International Council for Science – World Data System certification and the Data Seal of Approval. Together, both organizations have initiated an interest group (IG) on certification in the context of the Research Data Alliance. The working groups linked to the IG aim at improving certification catalogues, as well as at a common organization of certification procedures. The next step is to invite more organizations and to establish a common certification authority.

**Keywords.** Accreditation, certification, trustworthiness.

## 1. Introduction

Sharing of data and related services stimulates large scale and complex approaches in science. It leads to more efficiency in research. Similar to the peer review system for science literature an element is needed that indicates the quality of usable data and services.

## 2. Certification and Accreditation of Data Centres and Services

A prerequisite for accreditation of data services is certification, which should supply an objective base for evaluating data archives, as well as more specialized service providers. The overall goal is to supply an estimate on the overall trustworthiness of data services. Certification and accreditation is beneficial for different stakeholders:

1. For researchers, who want to be confident that the integrity and authenticity of the data in archives are protected and the data remain

accessible, usable, and meaningful over time, and that data services provided can be safely used.

2. For funders, who want reassurance that their investments in the production of valuable research data are not wasted, but will catalyze future research. Certification will make it easier for funders to assess service providers and to make informed investment decisions.
3. For data centres and services, for which certification provides a quality indicator to show to funders and users. In addition, the certification process will provide them with advice on where improvements are needed and useful.

Different approaches have been accomplished during the last decade. Among those are the Open Archival Information System (OAIS) [1], Network of Expertise in Long-Term Storage of Digital Resources (nestor) seal [2], International Organization for Standardization

(ISO) standard 16363 [3], certification of the World Meteorological Organization Information System (WMO-IS) [4], Data Seal of Approval (DSA) [5], and International Council for Science – World Data System (ICSU-WDS) certification of members [6]. Except OAIS, all approaches supply a catalogue for certification criteria. Some of the approaches are community independent; others are more adapted to a specific science field. Also, there are differences with respect to the efforts needed to conduct a certification—for reviewers, as well as for applicants.

Except those for WMO-IS and ICSU-WDS, all approaches concentrate on a data repository function. However, due to the often complex nature of data services other functions should be included. The repository function is often only part of a service. In this regard, certification for the WMO-IS is proprietary, whereas ICSU-WDS has a more generic approach for other service functions.

Eventually, certification is a mixture of the certification schema or catalogue, the certification procedure, and the reviewer's expertise. Certification practices within ICSU-WDS have shown that evaluations are highly dependent on the background of reviewers.

More recently, ICSU-WDS and DSA have agreed to cooperate and have founded an interest group (IG) on certification that is also endorsed by the Research Data Alliance (RDA). Under the umbrella of the IG, two working groups (WGs) have been established: (1) a WG on a global registry for certified data centres and services, and (2) the ICSU-WDS/DSA WG on certification. The goals of the latter group are the establishment improved certification catalogues and procedures, as well as a common test bed.

Both organizations within in the WG are certification authorities (CAs). ICSU-WDS uses its

Scientific Committee and DSA its Board to conduct reviews. The aim for the next step is invite more organizations and to establish a common CA.

### 3. Conclusions

Owing to the mostly complex nature of data hosting and serving facilities, accreditation is based on generic, as well as specialized, certification criteria. Conducting certifications requires experienced experts in various fields. To raise the impact of accreditations—and thus the trustworthiness of data facilities—organizations active in this field need to collaborate to foster a common CA.

### References

1. Reference Model for an Open Archival Information System (OAIS; corresponds to ISO 14721:2003), <http://public.ccsds.org/publications/archive/650x0b1.pdf> [accessed on: September 2013]
2. nector Working Group, Catalogue of Criteria for Trusted Digital Repositories: Version 2, Deutsche Nationalbibliothek, Frankfurt (Main), Germany, <http://nbn-resolving.de/urn:nbn:de:0008-2010030806> [accessed on: September 2013]
3. ISO 16363:2012—Space data and information transfer systems—Audit and certification of trustworthy digital repositories, [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=56510](http://www.iso.org/iso/catalogue_detail.htm?csnumber=56510) [accessed on: September 2013]
4. WMO-IS, [http://www.wmo.int/pages/themes/wis/index\\_en.html](http://www.wmo.int/pages/themes/wis/index_en.html) [accessed on: September 2013]
5. DSA <http://www.datasealofapproval.org/> [accessed on: September 2013]
6. ICSU-WDS, Certification of WDS members, [http://icsu-wds.org/images/files/WDS\\_Certification\\_Summary\\_11\\_June\\_2012.pdf](http://icsu-wds.org/images/files/WDS_Certification_Summary_11_June_2012.pdf) [accessed on: September 2013]

# ICPSR and the Data Seal of Approval: Accreditation Experiences and Opportunities

*Jared Lyle<sup>1\*</sup>, Mary Vardigan<sup>1</sup>*

<sup>1\*</sup> *Inter-university Consortium for Political and Social Research (ICPSR), P.O. Box 1248, Ann Arbor, MI, 48106, U.S.A.*

Email: lyle@umich.edu

**Summary.** The Inter-university Consortium for Political and Social Research (ICPSR), a domain repository with a 50-year track record of archiving social and behavioural science data, applied for and acquired the Data Seal of Approval (DSA) in 2010. The DSA was a non-intrusive, straightforward approach to assessing organizational, technical, and operational infrastructure. The DSA assessment helped ICPSR improve transparency, monitor high-level archival processes, and raise awareness within the organization and beyond.

**Keywords.** Assessment, certification, data repository, trusted repository, Data Seal of Approval.

## 1. Introduction

As data repositories and dissemination platforms proliferate, assessment of repository quality and trustworthiness grows in importance. Assessment promotes trust that the data will be available for the long term, provides a transparent view into the workings of the repository, and improves processes and procedures through measurement against a community standard.

Common elements of assessment include review of the organizational framework (e.g., governance, staffing, policies, finances), technical infrastructure (e.g., system design, security), and treatment of data (e.g., access, integrity, process, preservation).

## 2. The Data Seal of Approval

The Data Seal of Approval (DSA) was initiated by the Data Archiving and Networked Services (DANS) in 2009 'to safeguard data, to ensure high quality and to guide reliable management of research data for the future without requiring the implementation of new standards, regulations or high costs' [1]. There are 16 guidelines to the DSA—three target the data producer, three the data consumer, and ten the data repository.

Self-assessments are completed online with ratings and then peer-reviewed by a DSA Board member. Approximately 20 repositories have been granted the DSA since 2011.

The DSA has minimal requirements in comparison to other assessments, such as the Trustworthy Repositories Audit and Certification (TRAC) [2], the Trusted Digital Repository Checklist (ISO/DIS 16363) [3], and the Digital Repository Audit Method Based on Risk Assessment [4].

## 3. DSA Experience of Inter-university Consortium for Political and Social Research

The Inter-university Consortium for Political and Social Research (ICPSR) applied for and acquired the DSA in 2010. ICPSR found the DSA to be a non-intrusive, straightforward approach to assessing high-level organizational, technical, and operational infrastructure. The DSA is less labour- and time-intensive than other assessments; completion of documentation took a few days rather than months.

The DSA assessment process helped ICPSR improve transparency, monitor high-level archival processes, and raise awareness within the



organization. For instance, in documenting the DSA guidelines, ICPSR staff recognized the need to make policies more public, including posting past versions of Terms of Use agreements. Displaying the DSA logo is also a visible sign to ICPSR website visitors that the repository has achieved trusted status.

Since ICPSR previously had undertaken a TRAC self-audit, the DSA did not uncover significant flaws in the system, but it did help the organization continue to sharpen its processes and procedures.

#### **4. Conclusions**

The Data Seal of Approval provided an inexpensive, relatively quick, and straightforward accreditation process. The results of the DSA process helped ICPSR to continue to refine processes and procedures. The DSA provides a low barrier of entry for repositories to certify that

they are trustworthy, while helping them to improve their own systems. The Seal carries meaning that is easily recognized, especially as more repositories complete the assessment and as more producers and consumers recognize the value added.

#### **References**

1. Data Seal of Approval, <http://datasealofapproval.org/?q=about> [accessed on: September 2013]
2. Trustworthy Repositories Audit and Certification, [http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf) [accessed on: September 2013]
3. Trusted Digital Repository Checklist, <http://public.ccsds.org/publications/archive/652x0m1.pdf> [accessed on: September 2013]
4. Digital Repository Audit Method Based on Risk Assessment, <http://www.repositoryaudit.eu/> [accessed on: September 2013]





# **The Future of Scientific Polar Data Management**

# International Polar Data Management

**Mustapha Mokrane**<sup>1\*</sup>

<sup>1\*</sup> *ICSU World Data System International Programme Office, Koganei, 184-8795 Tokyo, Japan*  
Email: mustapha.mokrane@icsu-wds.org

**Summary.** The scientific research landscape has dramatically changed since the first international data-driven research programmes of the late 19th and early 20th centuries. Changing scientific practices, increased societal expectations, and technical advances with impact on the volumes of data generated, are making international coordination efforts for research data management needed more than ever before. Convergence between domain-specific data activities should be encouraged through such efforts. The priority should be set for the scientific community to enable multidisciplinary approaches needed to respond to modern global challenges. The polar research community has always been at the forefront of multidisciplinary research—as demonstrated during the last International Polar Year (IPY)—and is therefore uniquely positioned to adopt this approach and coordinate its data activities.

**Keywords.** International polar data management, scientific data curation and stewardship, long-term preservation, open-access.

## 1. Historical Background

Since the last IPY in 2007–2008, the international scientific data management landscape has slowly improved, in particular, in the field of polar data. Historically, IPY built on the successful past models of the 1957–58 International Geophysical Year (IGY) and even older IPY in 1882–1883 [1]. IGY was a good example of an internationally coordinated research programme, which influenced the global data management landscape. One of its clear successes, and possibly only institutional legacy, are the so-called International Council for Science (ICSU) – World Data Centres. This was one of the first international data-driven efforts to preserve and make scientific data openly and freely accessible [2]. When seen from a purely scientific perspective, IPY—and its predecessor models—were undoubtedly successful multidisciplinary scientific research endeavours. However, from a scientific data management point of view, they have also revealed how challenging it is for the international scientific community to coordinate the management, preservation, and dissemination of scientific data, including for some of the IPY datasets [3].

## 2. Evolution of the Global Data Management Landscape

Learning from the IPY discomfiture, global scientific organizations (such as ICSU) and intergovernmental organizations (such as the World Meteorological organization (WMO)) have already upgraded existing, or started new, data management initiatives. For example, the ICSU World Data System (ICSU-WDS) [4] and the WMO Information System (WMO-IS) [5] will contribute to a stronger international scientific data management framework by enhancing the international coordination efforts and providing the basis of common infrastructures. However, it is critical for the success of these initiatives that scientific data practitioners are effectively engaged. Because the scientific challenges at hand are multidisciplinary by nature, these global efforts need also to be widely promoted and adopted by domain-specific scientific communities, including polar data practitioners and researchers.

In the area of polar research data, the Antarctic community has a long-standing international data management effort operating

under the umbrella of ICSU's Scientific Committee for Antarctic Research and the Antarctic Treaty [6]. Recently, the Arctic polar data community is increasingly concerned with data preservation and sharing, and efforts have started under the umbrella of the ICSU International Arctic Science Committee and the Arctic Council to increase awareness about data issues [7]. These programmes and initiatives need to coordinate with each other to maximize the benefits of existing investments, but also strengthen their ties with global coordination efforts like ICSU-WDS, WMO-IS, and others.

For example, data policies and good practises already adopted in data management at the global level by other communities can be adopted or transferred and adapted, where appropriate. The concept of data publication—including the use of permanent identifiers such as Digital Object Identifiers—has also gained a lot of international traction [8]. However, publishing data is far from being widely adopted in the developing polar data management networks, even though this mechanism offers clear benefits in terms of accessibility and usability of datasets, as well as scientific reward for the data practitioners in the form of scientific publications and citations.

Certification of data repositories is another mechanism by which global initiatives are promoting good practises. A number of synergetic certification procedures co-exist, ranging from the heavier International Organization for Standardization certifications to the more balanced community-based norms like ICSU-WDS membership and the Data Seal of Approval. Polar data repositories and services should proactively adopt certification procedures to align their capacities with similar capacities in other domains.

### 3. The Way Forward

So far, the weaknesses and relative lack of coordination in global scientific data systems are hindering the full realization of societal benefits expected from taxpayer-funded research. The

reasons behind this slow progress are diverse, and range from the relatively easy to solve—technical issues like interoperability, to the more difficult to tackle—political obstacles and barriers to transnational coordination. Another important barrier is the lack of recognition in the scientific community for data management practitioners' work on the one hand, and the lack of new and sustainable funding mechanisms to support the internationally coordinated data e-infrastructure needed by the scientific community on the other hand. Much remains to be solved in order to build an internationally coordinated research data infrastructure that provides openly accessible and usable scientific data. The ICSU World Data Centres demonstrated to some extent how a flexible international coordination mechanism—based solely on national capacities—can provide successful long-term data preservation and accessibility. Today, the scientific endeavour and the societal challenges we are facing, the amount of funding available, and the volumes of data produced, have dramatically changed. This new landscape requires new models, where domain-specific data communities continue to advance their activities and at the same time, open up to and link with other domains under the umbrella of international coordination efforts.

### References

1. Chris R. et al, A Framework for the International Polar Year, 2007–2008. ICSU, Paris, 2004
2. Aronova E., Baker K. S., Oreskes N., Big Science and Big Data in Biology: From the International Geophysical Year through the International Biological Program to the Long Term Ecological Research (LTER) Network, 1957–Present. *Historical Studies in the Natural Sciences*, 40(2), 183–224
3. David, C., A lesson in sharing. *Nature*, 469, 293, 2011
4. ICSU World Data System, <http://www.icsu-wds.org> [accessed on: July 2013]
5. World Meteorological Organization Information System, <http://www.wmo.int/pages/prog/www/WIS/> [accessed on: July 2013]

6. Scientific Committee on Antarctic Research – Standing Committee on Arctic Data Management, <http://scadm.scar.org/index.php> [accessed on: July 2013]
7. International Arctic Science Committee's Data Statement, <http://www.iasc.info/home/iasc/data> [accessed on: July 2013]
8. OECD, OECD Principles and Guidelines for Access to Research Data from Public Funding, 2007

# Towards an International Polar Data Coordination Network: An Arctic Perspective

**Peter L. Pulsifer<sup>1\*</sup>, Lynn Yarmey<sup>1</sup>, Julie Friddell<sup>2</sup>, William Manley<sup>3</sup>, Amos Hayes<sup>4</sup>, Allison Gaylord<sup>5</sup>, Scot Nickels<sup>6</sup>**

<sup>1\*</sup> National Snow and Ice Data Center, Univ. of Colorado 1540 30th St., Boulder, CO 80301 USA

<sup>2</sup> Canadian Cryospheric Information Network, Univ. of Waterloo 200 University Ave. W, Waterloo, ON, N2L 3G1, Canada

<sup>3</sup> Institute of Alpine and Arctic Research, Univ. of Colorado 1560 30th St., Boulder, CO 80303 USA

<sup>4</sup> Geomatics and Cartographic Research Centre, Carleton University 1125 Colonel By Dr., Ottawa, ON, K1S 5B6

<sup>5</sup> Nuna Technologies, PO Box 1483, Homer, AK, 99603, USA

<sup>6</sup> Inuit Quajisarvingat, 75 Albert St. Suite 1101, Ottawa, Ontario, K1P 5E7, Canada

Email: pulsifer@nsidc.org

**Summary.** Many elements of an international Arctic Data Coordination Network currently exist. However, active network-building activities are required to achieve this goal, including better understanding of how networks develop, recognizing the importance of scale, establishing a diverse resource base, and developing community standards.

**Keywords.** Arctic, data management, research networks, standards, funding.

## 1. Introduction

Research data management is not new—the challenges of discovering, accessing, and using data have existed for centuries. In the domain of polar research, the World Data Centre system was developed more than fifty years ago by the International Council for Science to manage data resulting from the International Geophysical Year of 1957–58 [1]. More recently, significant progress towards establishing an international polar data management Network was made during the International Polar Year 2007–2008 (IPY) [2].

A sustained international Network does not yet exist, however. In this paper, it is argued that the fundamental building blocks for such a Network exist, and that the time is right to move forward. Realizing such a network will require a number of elements: (i) an organic but nurtured community of practice, supported by a participatory framework that recognizes both the power of the network and the need for autonomy [3–4]; (ii) recognition and engagement of actors at all scales from local to global, including indigenous peoples of the Arctic; (iii) a sustained yet diverse resource model; (iv) community data

sharing standards.

The benefits are recognized of developing a Network that includes both Arctic and Antarctic components; however, based on the expertise of the authors, an Arctic-focus perspective is presented.

## 2. Pieces of the Puzzle

A review will be provided of existing Arctic data management projects and programmes; including, but not limited to: the IPY Data Information Service and its legacy, Advanced Cooperative Data and Information Service, Arctic Observing Viewer, Exchange for Local Observations and Knowledge of the Arctic, Polar Data Catalogue, Inuit Knowledge Centre, Geomatics and Cartographic Research Centre, Polar View, Circumpolar Biodiversity Monitoring Program, Arctic Portal, and World Meteorological Organization Information System.

## 3. Putting the Pieces Together

While IPY programs made significant progress towards establishing a Network, the period of

planning and the duration of IPY were relatively short. Researchers now have the lessons learned from IPY to help us move forward. The authors argue that examining the theory and practice of building networks is useful in considering how to establish the Network.

### 3.1 Networks

Broad social or organizational networks require strong and numerous bonds; however, ‘weak ties’—occasional versus frequent interaction, surface versus in-depth collaboration—and network hubs can facilitate network building.

Effective participatory mechanisms must be in place, even in very formal organizations and networks. Ensuring that all voices are heard is necessary to develop sustained connections at a global scale.

### 3.2 Scale

Recently, the power of ‘grassroots’ movements and the ‘local’ have been demonstrated through phenomena such as Citizen Science and democratic movements. Whilst the power of these movements is recognized, it is argued that what is sometime negatively characterized as a top-down approach (overarching policy development, Big Science etc.) is also important and valuable in building networks.

This section on scale is concluded by arguing that the ‘regional’ or ‘middle’ scale is critically important in connecting a strong Network.

### 3.3 Diverse Resourcing

Establishing adequate sustained resourcing (human, financial, infrastructure) is a central concern of Network development. Recent advances in technology make the world smaller than ever and increase the possibilities for a collaborative effort that maximizes the efficiency of resources. Achieving this in the real world—where individual projects and institutions have their own objectives, deliverables, personalities, and so on—is challenging.

Here, a model of ‘crowdsourcing’ will be discussed that builds on the tenets of sound open source software development projects, cloud computing, and emerging organizational and

economic vehicles such as the virtual organization.

### 3.4 Interoperability

A core driver for developing a Network is the desire for interoperability—the ability of a system to work with other systems without special effort on the part of the user. Interoperability for information sharing fundamentally relies on the creation and adoption of community-based data description standards and data delivery mechanisms. There is a broad range of interoperability frameworks and specifications available; however, these need to be adapted for polar community needs. Progress towards Network interoperability will be reviewed, and a prototype distributed data systems demonstrated. The remaining challenges will also be explored.

## 4. Conclusions

Establishing a Network will be challenging, with possibly the most difficult challenge being the development of a model that enables participants to gain from the collective, while maintaining their autonomy.

## References

1. Ruttenburg, S., The ICSU World Data Centers. *Eos*, 73(46), 494–495, 1992
2. Parsons, M. A., de Bruin, T., Tomlinson, S., Campbell, H., Godøy, Ø, LeClert, J., The State of Polar Data—the IPY Experience. In: Krupnik, I., Allison, I., Bell, R., Cutler, P., Hik, D., López-Martínez, J., Rachold, V., Sarukhanian, E., Summerhayes, C., *Understanding Earth’s Polar Challenges: Summary by the IPY Joint Committee, International Polar Year 2007–2008*, 2011
3. Wenger, E., McDermott, R., Snyder, W. M., *Cultivating Communities of Practice*. Harvard Business Press, 2002
4. Yarmey, L., Baker, K., Towards Standardization: A Participatory Framework for Scientific Standard-Making. *International Journal of Digital Curation*, 8(1), 2013



# Creating Web of Data for Science

**Hideaki Takeda**<sup>1\*</sup>

<sup>1\*</sup> *National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo, Japan*  
Email: takeda@nii.ac.jp

**Summary.** The web can be the infrastructure for sharing and reusing scientific data. This is enabled by Linked Open Data (LOD) technologies, which are extensions of the web into the data domain. With LOD technologies, data is shareable, reusable, and linkable. Building schemata for domain specific data and identifier systems for key objects provide greater semantics on the top of the LOD infrastructure.

**Keywords.** Linked Open Data, World Wide Web, Digital Object Identifier, Open Researcher and Contributor ID.

## 1. Introduction

The invention of the World Wide Web (the web) provided humanity with such a revolutionary change in communication among people that almost no-one can imagine our society without the web now. A similar impact has recently begun in data communication due to what is called Linked Open Data (LOD), which are the natural extension of the web into the data domain and is enhanced with Semantic Web Technology. It is expected that LOD will change both the publication and use of data, and thus their value.

## 2. Value of the Web

The web is not only a convenient tool for communication but has also changed the way of communicating. The most significant change is that the web has created a global information space where everyone can place, retrieve, and link information without permission, in contrast to traditional information systems that require control for these actions to some extent. This change was enabled by the use of very simple techniques, such as, Uniform Resource Identifiers (URIs), Hypertext Transfer Protocol (HTTP), and Hypertext Markup Language (HTML).

URIs enable everyone to create unique locations for their information with little permission. Each URI is entirely unique so that information with a URI can be shared worldwide.

HTTP supplies a simple access method to information with a URI, and HTML offers a simple representation method for information linked to other information with a URI.

In its entirety, pieces of information added by people all over the world are formed as a unique information sphere. Anyone can reach any piece of information freely, and also find others by tracing hyperlinks.

## 3. Linked Open Data

LOD are the natural extension of the web into the data domain, specifically, they form a unique global data sphere. The key LOD technologies are similar to those for the web—URIs, HTTP, and the Resource Description Framework (RDF).

URIs and HTTP take the same roles as for the web, ensuring the uniqueness of identifiers and access to data, respectively. However, instead of HTML to represent information, RDF is used to describe data.

RDF represents data via a simple syntax: resource–property–value (a triple). The resource must be a URI, whereas the other two are either URIs or are literally stated. The difference between HTML and RDF is that the links are not labelled in HTML, but properties are used as labels in RDF. Since data are mostly processed by computers, the need for labelling links is of greater importance.

Data are described as URIs and RDF and are accessed via HTTP so that all data in datasets published by anyone can be accessed and linked to each other.

RDF Schema (RDFS) produce structures in RDF. Schemes are defined with RDFS and published as RDF so that anyone can use the published schemes and combine certain of them.

Tim Berners-Lee, the inventor of the web, defined the principles for Linked Open Data [1];

- Use URIs as names for things.
- Use HTTP URIs so that people can look up those names.
- When someone looks up a URI, provide useful information using standards (RDF, SPARQL).
- Include links to other URIs so that people can discover more things.

Linked Open Data are accepted by people in various domains and more than 200 datasets are now published as LOD; forming the LOD cloud shown in Figure 1. In this figure, the lower-right area indicates the datasets in the Life Science, where targets like genes, proteins, chemicals, and diseases are shared among different research activities. Hence, LOD are useful for the different activities to connect data in the datasets.

#### **4. Other Key Technologies for Web of Data for Science**

LOD provide the basic infrastructure to publish and share data. However, more technologies are required to enhance them, especially in science.

The first technology is the building of shareable schemata in each domain. Each science domain has some sort of schemata for data. These need to be converted into shareable, extendable, and combinable schemata that can be used as schemata in LOD.

The second is the building of permanent identifiers for key objects. URIs merely provide a

syntax for unique identifiers but do not offer semantics for them. Contrary to the freedom in the web, identifiers for some objects should be controlled and maintained if they are critical within the domains.

A good example is the Digital Object Identifiers (DOIs) mainly used for academic digital objects like scientific papers. These are maintained by a system organized by a single organization; the International DOI Foundation, which ensures not only the uniqueness but also permanent accessibility of the identifiers.

Another example is the Open Researcher and Contributor ID project, which can supply a unique and permanent identifier for any researcher in the world and is also maintained by a system organized by a single organization. Such schemata and identifier systems contribute greater semantics on the top of the LOD infrastructure.

#### **5. Conclusions**

Traditionally, data in science have been shared and reused. The web provides a new way for such sharing and reuse. With a number of scientific domains having already adopted this new method, it is reasonable to apply it in the Polar Sciences too. In particular, since various research activities from different domains produce data in the Polar Sciences, the merits of the web of data would be maximized. While some challenges still exist such as the treatment of Big Data as a web of data, these will be solved by researchers in both the Computer Sciences and other sciences working together.

#### **References**

1. W3C, <http://www.w3.org/DesignIssues/LinkedData.html> [accessed on: September 2013]

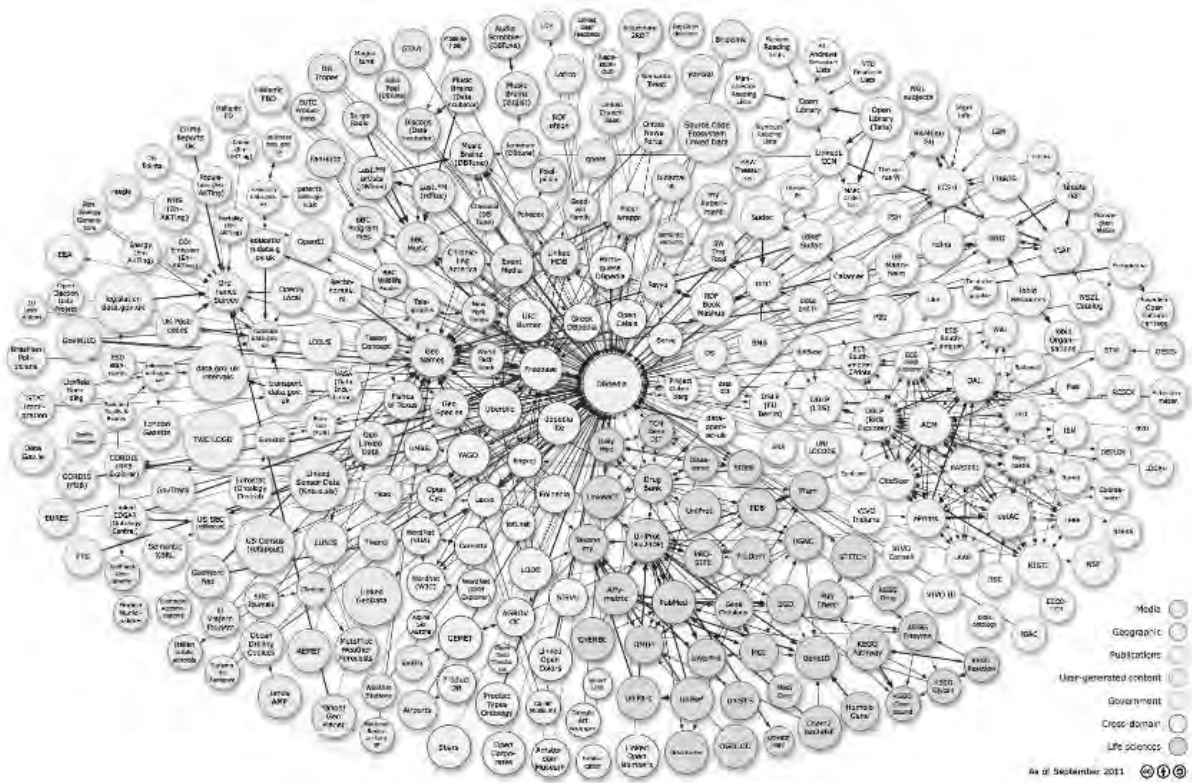


Figure 1. LOD Cloud.





# **Poster Presentations**

# Hornsund GLACIO–TOPOCLIM Database—Polish Polar Station IPY Legacy

**Bartłomiej Luks<sup>1\*</sup>, Sebastian Sikora<sup>2</sup>, Dariusz Puczko<sup>3</sup>, Tomasz Budzik<sup>4</sup>, Tomasz Wawrzyniak<sup>1</sup>**

<sup>1\*</sup> *Institute of Geophysics Polish Academy of Sciences, Ksiecia Janusza 64, Warsaw, 01-452, Poland*

<sup>2</sup> *University of Wrocław, Kosiby 8, Wrocław, 51-621, Poland*

<sup>3</sup> *Institute of Biochemistry and Biophysics Polish Academy of Sciences, Pawinskiego 5a, Warsaw, 02-106, Poland*

<sup>4</sup> *University of Silesia, Bedzinska 60, Sosnowiec, 41-200, Poland*

Email: luks@igf.edu.pl

**Summary.** The purpose of creating the GLACIO–TOPOCLIM database was to provide a tool for sharing and exchange of meteorological, climatological, and glaciological data among scientists working in surrounding of Hornsund fjord during the third International Polar Year 2007–2008. Launched in 2010, the Hornsund database became a popular tool for a number of scientists interested in the polar regions, and is broadly used by the climatological community.

**Keywords.** Meteorology, glaciology, database.

## 1. Introduction

Dynamic changes in the environment are particularly evident in the Arctic. The observation and continuous monitoring of selected geophysical parameters is an objective fulfilled by staff of the Polish Polar Station in Hornsund, Svalbard. These measurements are a continuation of research started during the International Geophysical Year 1957–1958, when the Polish built an all year-around research station in Isbjornhamna Bay in Hornsund fjord, Spitsbergen. This station has operated continuously since 1978. The Institute of Geophysics of the Polish Academy of Sciences, who owns the station, deals with data acquisition from monitoring programmes. In addition to the base research and monitoring program, the station's staff help in research carried out by scientists from different Polish universities and research institutions.

## 2. Database

The third International Polar Year 2007–2008 (IPY) prospered, with a number of research projects realized in cooperation with the Polish Polar Station. Meteorological and glaciological projects

were the main interest in the Hornsund fjord surroundings. Since these projects were conducted by research teams from different institutions, and the number of people handling and processing data grew as time went on, a common platform for data sharing and exchange was essential.

The Hornsund GLACIO–TOPOCLIM database was founded in 2010. In the beginning, it was used for sharing the outputs of two large IPY projects: Topoclim and Glaciodyn. Those outputs were mainly meteorological datasets from Automatic Weather Stations localized in the Hornsund surroundings, and the datasets formed the core of the Hornsund GLACIO–TOPOCLIM database. Since then, the database has gradually been supplemented with Hornsund Station meteorological datasets and the results of completed and ongoing glaciological research from the area. The database is open to all scientists wishing to contribute meteorological and glaciological data from Svalbard.

The database can be accessed through the [www.glacio-topoclim.org](http://www.glacio-topoclim.org) website; however, registration is required to retrieve the collected data. In contrast, access to metadata, a list of

publications connected with Hornsund fjord, and monthly meteorological reports are permitted without registration. To date, the database has over 120 active users, and is broadly known within the Polish polar science community.

### **3. Conclusions**

In conclusion, the Hornsund GLACIO–TOPOCLIM database provides:

- A data exchange and sharing platform.
- Open access to metadata, a list of publications, and monthly weather reports.
- Access to datasets free-of-charge.

**Acknowledgments.** Establishing of the monitoring system was possible thanks to the financial support of the Glaciodyn and Topoclim IPY projects by the Ministry of Science and Higher Education, Republic of Poland. Development and maintenance of the system has been granted by the projects: ice2sea (Seventh Framework Programme; No. 226375) and SvalGlac (NCBiR/PolarCLIMATE-2009/2-2/2010).

# Harmonizing Polar Biodiversity Data for Wider Access and Integration: A Collaboration between the Spanish Polar Data Center and GBIF-Spain

**Oscar Bermúdez<sup>1\*</sup>, Virginia González-Álvaro<sup>2</sup>, Francisco Pando<sup>2</sup>, Antonio Barragán<sup>1</sup>**

<sup>1\*</sup> National Polar Antarctic Data Centre (NPDC), Spanish Geological Survey, Ríos Rosas 23, Madrid, 28003, Spain

<sup>2</sup> GBIF Spain, Coordination Unit, Royal Botanical Garden-CSIC, Plaza Murillo 2, Madrid, 28014, Spain

Email: o.bermudez@igme.es

**Summary.** Collaboration between the Scientific Committee on Antarctic Research (SCAR) and Global Biodiversity Information Facility (GBIF) made polar biodiversity data accessible through GBIF. Moreover, the current polar biodiversity data available through GBIF was reverted into the National Polar Data Center. Only a few datasets were chosen initially to establish a methodology that could be later applied to the other datasets. The data profile used by GBIF was Darwin Core. To adapt the polar datasets already published within GBIF for the SCAR databank, only metadata was a concern since there is no standardized profile for biodiversity data within the National Polar Data Center. The procedure defined for connecting two large data avenues such as SCAR and GBIF has been expressed in a way that can be easily expanded and replicated.

**Keywords.** Biodiversity, primary data, polar data, open access, Global Biodiversity Information Facility.

## 1. Introduction

The Global Biodiversity Information Facility (GBIF) is an intergovernmental organization aimed at becoming the largest biodiversity data network in the world. The Scientific Committee on Antarctic Research (SCAR) is a member of GBIF and has amongst its goals facilitating new interdisciplinary science.

## 2. Objectives

Making polar biodiversity data available through GBIF seemed a logical and effective way to establish proven methods for managing biodiversity polar data.

In fact, this interaction is a two-way avenue, as long as the polar biodiversity data gathered and published through GBIF can be discovered and incorporated into the National Polar Data Center.

## 3. Description of Work

Once the objectives and purposes of collaboration were identified, a few datasets were

chosen to explore the different options and to establish a methodology that could be later applied to other datasets.

Thus, two datasets—compiled in the context of the National Polar Data Center—were identified as suitable testbeds to standardize data and procedures according to the GBIF protocols, and subsequently, to make them publishable through the GBIF network:

- Antarctic lower plants
- Antarctic non-marine aquatic ecosystems

An additional dataset containing polar biodiversity data was then selected for incorporation into the National Polar Data Center:

- Antarctic Porifera database from the Spanish benthic expeditions

Details of these three datasets are given in Table 1.

For transformation of the abovementioned GBIF dataset, two procedures were established, one for data, and another for metadata.



**Table 1.** Details of the datasets used in this work.

Name	Origin	Number of records	Estimated number of species	Taxonomic Group
Antarctic lower plants	Spanish Polar Research Program	20	–	Lichens and mosses
Antarctic non-marine aquatic ecosystems	Spanish Polar Research Program	45	–	Various taxonomic groups
Antarctic Porifera database from the Spanish benthic expeditions	Spanish Antarctic expeditions	593	74	Antarctic Porifera

Data transformation implied mapping the available relevant data into the data profile used by GBIF, namely, Darwin Core [1]. This was achieved using specific programmatic functions to transform and standardize scientific names, geographic coordinates, and other details.

Regarding metadata, SCAR and GBIF use two well-established metadata profiles; Directory Interchange Format [2] and Ecological Metadata Language [3], respectively. Both metadata standards can be expressed as Extensible Markup Language (XML) and constructed according to rules determined by their respective XML Schema Definition files. In this case, the transformations involved use of the appropriated Extensible Stylesheet Language Transformations. To adapt the polar dataset already published within GBIF for the SCAR databank, only metadata was a concern, since a standardized profile for biodiversity data is not used within the Spanish Polar Data Center.

#### 4. Conclusions

A procedure for connecting two large data avenues such as SCAR and GBIF has been defined in a way that can be easily expanded and replicated.

The work carried out also demonstrates to projects dealing with polar biodiversity data how the impact of their research can be multiplied.

**Acknowledgments.** GBIF Spain and NPDC are supported by the Spanish Ministry of Economy and Competitiveness through agreements with the Spanish High Council of Scientific Research and the Spanish Geological Survey, respectively.

#### References

1. Global Biodiversity Information Facility, <http://www.gbif.org/informatics/standards-and-tools/publishing-data/data-standards/darwin-core-archives/> [accessed on: July 2013]
2. Centro Nacional de Datos Polares, <http://hielo.igme.es/index.php/es/site-map> [accessed on: July 2013]  
Global Biodiversity Information Facility, <http://www.gbif.org/informatics/discoverymetadata/ipt-and-metadata/> [accessed on: July 2013]

# 'Quantarctica': New Standalone GIS Package for Antarctic Research, Operation, and Education using Open-source Software

**Kenichi Matsuoka<sup>1</sup>, Anders Skoglund<sup>1</sup>, Angela Von Deschanden<sup>1</sup>, Stein Tronstad<sup>1\*</sup>**

<sup>1\*</sup> Norwegian Polar Institute, Tromsø, 9296, Norway  
Email: Tronstad@npolar.no

**Summary.** Antarctic geographic data was gathered from data centres worldwide and a Geographic Information System (GIS) project file—Quantarctica—was developed that works on Quantum GIS, a free, open-source software (GNU licensed). Quantarctica is standalone, thus Internet access is not required for it to run. Over the past two years, previous versions of Quantarctica have been used to examine geographical data at a range of scales (from continental to local views), prepare maps for publications, examine various project data alongside continent-wide datasets included in Quantarctica, and plot Global-Positioning-system-provided positions together with satellite images in real time during field campaigns. The package should also be useful for teaching about Antarctica in classrooms. It is hoped that researchers will contribute their field- and remote-sensing data, as well as model outputs, to improve *Quantarctica*.

**Keywords.** Geographic Information System (GIS), Antarctica, data sharing, Quantum GIS.

## 1. Introduction

Recently, submission of collected data is mandatory for many researchers working in the polar regions. As a consequence, it is necessary to develop straightforward ways to deliver these data to users so that the community will take full advantage of the deposited datasets. Some datasets, such as those containing ice thickness and coastline data, constitute a 'knowledge base' for numerous projects in different disciplines. It is beneficial to the entire polar community if such knowledge bases are provided in combination with a data visualization tool optimized to examine such datasets. Indeed, integrating different datasets into a useful form is time consuming and requires careful supervision.

There are several web-based Geographic Information System (GIS) tools with which people can visualize different datasets in tandem. Since these tools do not require software installation at the user side, they are accessible and convenient for visualization. However, such online systems are not useful at a remote camp at which Internet access is impractical. Moreover, they are inadequate as a research tool in many cases, as the settings cannot be changed nor fixed, and

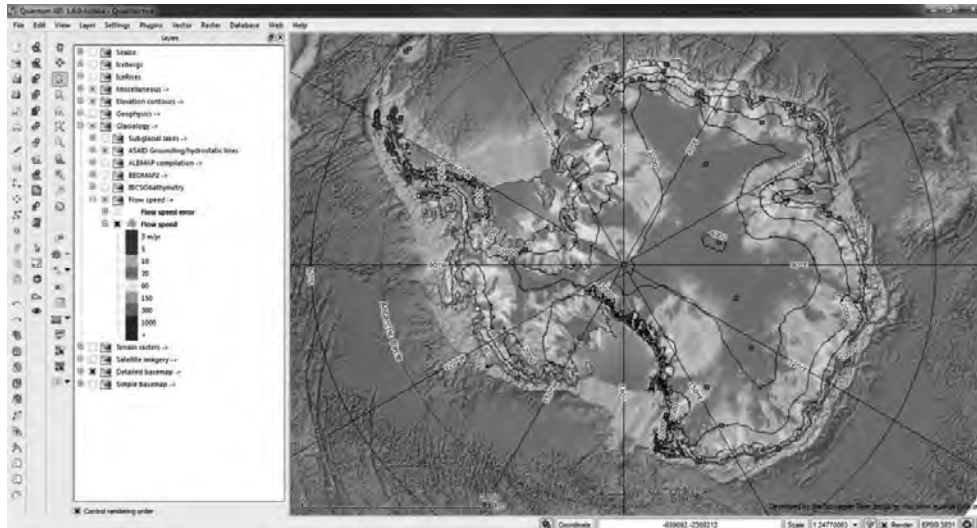
users cannot view their own data alongside the online-provided datasets.

Here, a new GIS package—*Quantarctica*—is presented that has been developed by the Norwegian Polar Institute (Figure 1). *Quantarctica* was originally developed as an in-house tool and has been tested extensively over the past two years. In response to the numerous requests for sharing the package, development of a public version commenced in autumn 2012, and the first version was released in July 2013. Hence, the package and associated relevant information are freely available ([www.quantarctica.org](http://www.quantarctica.org)), and *Quantarctica* is modifiable so that people can use it as a part of their own GIS project.

## 2. Platform software

Quantum GIS ([www.qgis.org](http://www.qgis.org)) is employed as the project platform. It is a GNU-licensed freeware that has been developed by the community since 2003. This community has also developed numerous plug-ins and maintained active and helpful interactions with users.

The data package is named *Quantarctica* (Quantum GIS + Antarctica) with a playful mind.



**Figure 1.** Quantartica screenshot. The background image shows ocean-bed topography (shaded). Over the Antarctic ice sheet and ice shelves, ice-flow speeds are shown in the package in red–blue colours, so several fast-flowing glaciers are highlighted. Contours show the ice surface elevation (1000 m intervals). The (red) squares indicate COMNAP-listed facilities. Users can select layers from the left column and overlay many (semi-transparent) data layers together.

### 3. Included datasets

Several of the datasets included in the current version (1.0) of Quantartica are listed below. When existing, we include the uncertainty of data (e.g., ice-flow speed) and availability of raw data used for the data compilations (BEDMAP2). These datasets are mostly peer-reviewed; and non-peer-reviewed data are those well-accepted and widely used within the community. These datasets are distributed as a part of Quantartica with the data provider's approval.

- **Geography data:** ADD Antarctic digital database, SCAR place names, COMNAP-listed facilities, terrain models.
- **Satellite imagery:** LIMA, MODIS, Radarsat. Continent-wide mosaics, high-resolution images, and tile indices of these high-resolution images.
- **Glaciology:** Subglacial lakes, ALBMAP, ASAIID grounding line, BEDMAP2, ice-flow speeds.
- **Geophysics:** ADMAP magnetic anomaly, GOCE gravity anomaly, GOCE geoid height.

### 4. Metadata and data visualization

Metadata are provided with each dataset. These clarify the data contributors, locations of the

original data, and references to where users can find full descriptions of the datasets.

Data are visualized using universal colour sets such that those who are colour-blind can see more information than they might using other colour sets.

### 5. Conclusions

The first version of Quantartica does not include data from, for example, the oceanography, biology, geology, and atmospheric sciences. This is primarily because the current development team does not have the expertise to judge the quality and usefulness of data from these disciplines. Since *Quantartica* is for the entire polar community, such quality control is important. It is expected that the community will assist this effort by suggesting useful datasets that should be included in future versions of Quantartica.

**Acknowledgments.** We acknowledge those providers who allowed us to distribute their datasets and helped us to annotate their datasets with metadata. We also acknowledge those users who have provided feedback.

# Polar Data Tools at Integrated Earth Data Applications (IEDA)

**Frank O. Nitsche<sup>1</sup>, Robert Arko<sup>1\*</sup>, Suzanne Carbotte<sup>1</sup>, Vicki Ferrini<sup>1</sup>, Andrew Goodwillie<sup>1</sup>,  
Suzanne O'Hara<sup>1</sup>, Rose Anne Weisse<sup>1</sup>, Kevin McLain<sup>1</sup>, Samanta Chan<sup>1</sup>, John Morton<sup>1</sup>,  
William B.F. Ryan<sup>1</sup>**

<sup>1\*</sup> Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W, Palisades, NY, 10964, USA  
Email: arko.ldeo.columbia.edu

**Summary.** Accessibility and preservation of data is needed to support multidisciplinary research in the key environmentally sensitive polar regions. Integrated Earth Data Applications (IEDA) is a community-based data facility funded by the US National Science Foundation to support, sustain, and advance the geosciences by providing data services for observational solid earth data from the Ocean, Earth, and Polar Sciences. IEDA services include a variety of databases and tools to help researchers archive and reference their data, as well as discover and analyze a wide range of data.

**Keywords.** Database, bathymetry, data references, Antarctic, Arctic.

## 1. Introduction

To maximize the utility of existing data, preserve them for future generations, and enable analysis of larger datasets and cross-disciplinary research, it is essential that data are archived and easily accessible. This is especially the case in the polar environment, where logistical efforts and data acquisition costs are extremely high. However, ensuring appropriate archiving, and maintaining data collections that are useful for other users, can be a large burden for individual investigators; one that takes time away from actual research.

In 2010, the Integrated Earth Data Applications (IEDA) facility was initiated with funding from the US National Science Foundation (NSF) to support ongoing maintenance and operations of a suite of data collections and services for the Solid-earth and Polar Science communities that is hosted by the Marine- Geoscience Science System and EarthChem. IEDA tools and services relevant to the polar research community include the Antarctic and Southern Ocean Data System (ASODS) [1], the US Antarctic Program Data Coordination Center (USAP-DCC) [2], GeoMapApp [3], the Global Multi-Resolution Topography (GMRT) Synthesis [4], as well as a number of services for sample-based data (System for Earth Sample Registration and EarthChem).

Recently, IEDA has continued to develop various system components to improve data access and visibility, and has developed tools to assist researchers and users throughout the entire data lifecycle—from generating a data management plan (required by NSF for new proposals), to the documentation and archiving of data, as well as data publishing. Here, IEDA tools and developments relevant to the polar community are described.

## 2. Data Repositories

One of the core components of IEDA for polar research is the ASODS ship-based data collection. These data are stored in an expedition-based relational database backend having a variety of data access and discovery tools. ASODS handles shipboard underway data acquired by US Antarctic Program vessels *NB Palmer* and *LM Gould*. ASODS also welcomes the contribution of derived datasets processed by the science party post-cruise.

ASODS includes a Data Portal and Search Interface that enables browsing via a map view or by expedition [1]. A sophisticated search interface facilitates searching under various criteria including cruise, ship, data type, and geographic location. This interface provides access to data

from over 2200 cruises throughout the global oceans, including both original field data and post-cruise products.

Whilst submission of data to discipline-specific data repositories is encouraged, datasets are collected in Antarctica without appropriate existing repositories. Such datasets acquired using US funding can be deposited at USAP-DCC [2], which has a web interface for data contribution that collects the key metadata and automatically generates Directory Interchange Format records for the Antarctic Master Directory (AMD). The portal enables searching of US datasets through a map browser or through a link to the AMD itself, and maintains close links with AMD. The search function also includes searching for non-US AMD records.

### 3. Data Referencing Tools

Researchers can be cautious about making data available due to acknowledgement concerns. Data publication and citation can address these concerns. IEDA offers a data publication service to register datasets in the Digital Object Identifier system as part of the DataCite consortium [5]. Moreover, IEDA is working with various publishers (e.g., Elsevier) to provide direct links from scientific papers to data in IEDA databases.

### 4. Visualization and Data Interaction

In addition to data archiving and search tools, IEDA provide tools for data visualization and analysis, including GeoMapApp—a sophisticated map-based client application, and Virtual Ocean—a globe-based three-dimensional interface. GeoMapApp, in particular, includes specific polar view interfaces for both the Antarctic and Arctic [3]. These tools support the analysis of different types of spatially co-located data, including gridded, profile, and point-based datasets. The base map for the tools is the GMRT dataset; a global compilation of high-resolution bathymetry and topography data [6]. Besides IEDA databases, these interfaces also provide access to a diverse range of other geoscience-

related datasets (e.g., earthquake catalogues, trackline gravity and magnetics, seismic reflection, satellite imagery, and environmental parameters) and enable loading of external datasets available through Open Geospatial Consortium services. Users can also explore and visualize—in the context of the built-in datasets—gridded, spreadsheet and shapefile-based datasets imported directly from their desktop.

### 5. Conclusions

Preservation of, and access to, scientific data are increasingly important components of the scientific research process. IEDA supports, sustains, and advances the Geosciences by providing data services and tools for observational solid-earth data from the Ocean, Earth, and Polar Sciences. IEDA systems enable these data to be discovered and reused by a diverse community both now and in the future. Comparable efforts are ongoing in many other institutions.

**Acknowledgments.** IEDA is funded by the US NSF through a Cooperative Agreement.

### References

1. Marine Geoscience Data System, <http://www.marine-geo.org/portals/antarctic/> [accessed on: September 2013]
2. US Antarctic Program Data Coordination Center, <http://www.usap-data.org/> [accessed on: September 2013]
3. GeoMapApp, <http://www.geomapapp.org/> [accessed on: September 2013]
4. Marine Geoscience Data System, <http://www.marine-geo.org/portals/gmrt/> [accessed on: September 2013]
5. DataCite, <http://www.datacite.org/> [accessed on: September 2013]
6. Ryan, W. B. F., Carbotte, S. M., Coplan, J. O., O'Hara, S., Melkonian, A., Arko, R., Weisse, R. A., Ferrini, V., Goodwillie, A., Nitsche, F., Bonczkowski, J., Zemsky, R., Global Multi-Resolution Topography synthesis. *Geochem. Geophys. Geosyst.*, 10, Q03014, 2009

# A Global Environmental Database Project at the National Institute for Environmental Studies and its Contribution to DIAS/GRENE

**Hideaki Nakajima<sup>1\*</sup>, Tomoko Shirai<sup>1</sup>, Nobuko Saigusa<sup>1</sup>, Yukihiro Nojiri<sup>1</sup>, Toshinobu Machida<sup>1</sup>, Hitoshi Mukai<sup>1</sup>**

<sup>1\*</sup> National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki, 305-8506, Japan  
Email: nakajima@nies.go.jp

**Summary.** The National Institute for Environmental Studies (NIES) has been collecting a wide variety of datasets related to the global environment. These datasets vary to those from the atmospheric environment, covering topics such as climate change, ozone depletion, acid rain, water environment, biodiversity, ecosystem, remote sensing, and Geographic Information Systems. Since the early 2000s, the Center for Global Environmental Research (CGER) within NIES has been running a database project that constructs and operates databases and data servers to distribute these global environmental datasets. Furthermore, in 2011, CGER started to cooperate with the Data Integration and Analysis System (DIAS) and Green Network of Excellence–environmental information (GRENE-ei) projects; providing its data related to global warming and biodiversity to them.

**Keywords.** Database, global environment, greenhouse gas, Data Integration and Analysis System, Green Network of Excellence.

## 1. Introduction

Since the early 2000s, the Office for Global Environmental Database at the Center for Global Environmental Research (CGER) within the National Institute for Environmental Studies (NIES) has been constructing and providing databases on topics related to Earth's global environment. The tasks of this office are divided into five major categories:

- Construct, maintain, and renew the database servers that supply the databases.
- Construct databases and provide the public/related researchers with data gathered by the Earth environmental monitoring project in NIES.
- Construct databases on social environmental sciences related to global warming.
- Develop convenient tools to analyze Earth environmental datasets.
- Achieve international cooperation on database-related issues.

## 2. Databases Provided by NIES

The databases provided by CGER from its server in NIES can be found on the Internet (see Figure 1 for the top page) [1]. These databases are categorized into the following seven types, where the number in parentheses represents the number of databases:

1. Climate Change (16)
2. Atmospheric Environment (12)
3. Oceanic and Inland Water (4)
4. Hydrology and Water Resources (1)
5. Biology and Ecosystems (2)
6. Remote Sensing and Geographic Information Systems (6)
7. International Cooperation (4).

The Climate Change category is further divided into four sub-categories: greenhouse gases observations (7), carbon sources/sinks (5), material flow (3), and effects and measures of global warming (1).

The Atmospheric Environment category is likewise divided into three sub-categories:

stratospheric ozone layer/UV (7), air pollution/acid rain (4), and trajectory analysis (1).

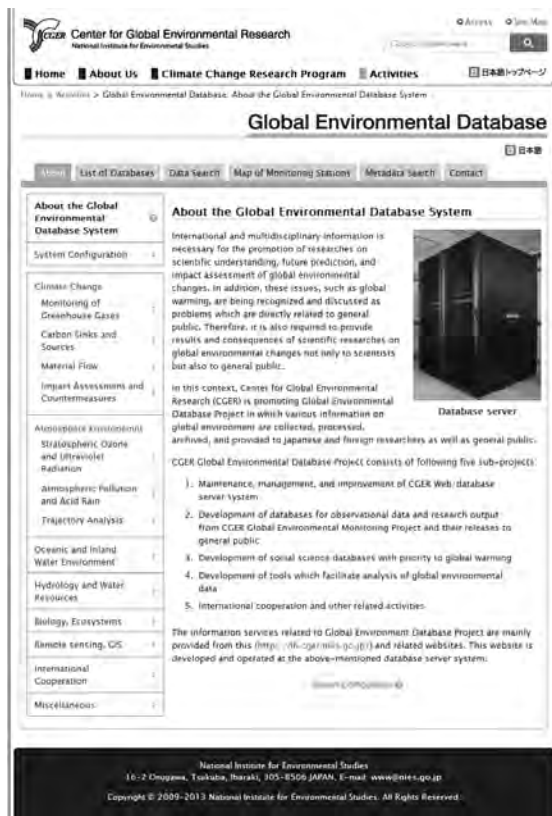


Figure 1. Top page of global environmental database.

### 3. NIES Contribution to DIAS Database Project

The Data Integration and Analysis System (DIAS) project aims to create knowledge that can be shared worldwide [2]. Specifically, DIAS will provide access to global and regional sensing data, and has developed a pilot system of an information storage infrastructure for public benefit applications and deepening of scientific knowledge in the areas of climate, water cycles, fisheries, agriculture, biodiversity, and their interdisciplinary fields. The DIAS design strategy is an operational framework that can provide public benefit in the form of policy-directed data delivery.

CGER has been participating in the DIAS project since April 2011; contributing to three areas of DIAS project: (1) DIAS upgrade and enlargement, (2) workbench prototyping, and (3) operational system planning. Consequently,

CGER has been submitting its observational data to DIAS on greenhouse gas measurements from ground- and ship-based platforms, as well as from flux measurement sites [1–2]. Figure 2 shows cruise trails of sea-surface carbon dioxide (CO<sub>2</sub>) partial pressure measurements by a voluntary ship Ms. Skaugran in 1995–1999. The data taken by these cruises will be stored and utilized in DIAS.

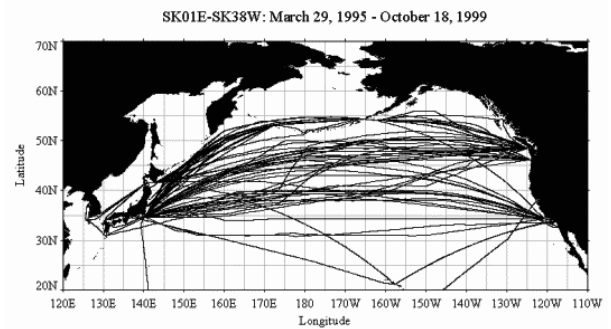


Figure 2. Cruise trails of sea-surface CO<sub>2</sub> partial pressure measurements by Skaugran in 1995–1999.

### 4. NIES Contribution to GRENE-ei Biodiversity Project

Green Network of Excellence–environmental information (GRENE-ei) is one the four GRENE categories, and consists of Health, Carbon Cycle, Agriculture, Biodiversity, Water, and City. This project has started to utilize the environmental information accumulated by DIAS, and enhance green innovation and networking in the field of earth environmental science and information engineering. NIES has participated in the biodiversity field of the GRENE-ei project since April 2011.

### 5. Conclusions

NIES has a wide variety of data related to Earth's environment. The CGER database project has been constructing databases from these data. Furthermore, CGER has started to cooperate with the DIAS and GRENE-ei projects.

**Acknowledgments.** We acknowledge Dr. J. Zeng, Dr. E. Maita, and Mr. T. Sato for their efforts in data management and data server maintenance.

## References

1. National Institute for Environmental Studies, <http://db.cger.nies.go.jp/> [accessed on: September 2013]
2. Data Integration & Analysis System, <http://www.editoria.u-tokyo.ac.jp/projects/dias> [accessed on: September 2013]
3. Zeng J., Nakajima, H., et al., Linking carbon dioxide variability at Hateruma Station to east Asia emissions by Bayesian inversion. *Lagrangian Modeling of the Atmosphere, Geophysical Monograph Series*, American Geophysical Union, 200, 349, 2012
4. Saigusa, N., et al., CO<sub>2</sub> flux measurement and the dataset in CarvoEastAsia. *J. of Forest Res.*, 18, 41–48, 2013



# Web-based Technologies in Data/Information Management for Polar Data in Russian World Data Centres

**Evgeny P. Kharin<sup>1</sup>, Michael V. Nisilevich<sup>1\*</sup>, Natalia Sergeyeva<sup>1</sup>, Ludmila P. Zabarinskaya<sup>1</sup>**

<sup>1\*</sup> *Geophysical Center, Russian Academy of Sciences, Molodezhnaya St. 3, Moscow, 719296, Russia.*  
Email: m.nisilevich@gcras.ru

**Summary.** The Russian World Data Centres (WDCs) for Solar-Terrestrial Physics and Solid Earth Physics played a significant role in providing prompt and free access to polar data during the International Polar Year 2007–2008 (IPY), and took part in two programmes: ‘IPY data and information service for distributed data management—IPYDIS’ and ‘Dataware for geophysical research during the IPY’. Moreover a website dedicated to IPY, ‘International Polar Year 2007–2008’, was created by the Geophysical Center, Russian Academy of Sciences. Both WDCs participated in the creation of the Russian national Portal, ‘IPY-Info’, which is an integrated high-quality multidisciplinary information system including metadatabases; databases; and systems for data collection, storage, and communication.

**Keywords.** Data, metadata, free access, International Polar Year, World Data Centre.

## 1. Introduction

The major objectives of the International Polar Year 2007–2008 (IPY) were conducting new international and multidisciplinary research in the polar regions of the Earth, and creating a system for the accumulation, long-term storage, management, and distribution of IPY data. This system was to ensure prompt and free access to the polar data and was to be based on new Internet technologies. A significant role in solving the latter was played by the World Data Centres (WDCs).

The Russian WDC for Solar-Terrestrial Physics (WDC–STP) and WDC for Solid Earth Physics (WDC–SEP) were active participants of the Russian Federation National Scientific Programme of IPY, and took part in two programmes: ‘IPY data and information service for distributed data management—IPYDIS’ and ‘Dataware for geophysical research during the IPY’ (a project of the Russian Academy of Sciences).

## 2. Main Activities

The main output from the implementation of the two aforementioned programmes was the creation of a dedicated website ‘International

Polar Year 2007–2008’ [1] that provides users with the results of various geophysical measurements from observatories and stations, scientific expeditions, experiments, and so on, in the Arctic and Antarctic regions carried out by the Former Soviet Union and then Russia from 1957 to the present day and stored in archives of both Russian WDCs. The site offers geomagnetic, seismological, heat flow, and gravity data; observations of cosmic rays, ionospheric phenomena, and volcanoes; and also contains some information about the history of exploration and study of the Arctic and Antarctic. Moreover, there are north and south polar regions maps showing the locations of Russian geomagnetic observatories, cosmic-ray stations, trajectories of the ‘North Pole’ drifting stations, spatial distributions of earthquake epicenters, points of heat flow measurements, and volcanoes. Convenient user access to IPY data is organized on the site.

Some historical data, for example, geomagnetic measurements at ‘North Pole’ drifting stations for the period 1956–1974, hourly and minute values of the geomagnetic field components, and analogue magnetograms from the Russian Arctic and Antarctic observatories,

were converted into digital form especially for this site. The earthquake catalogue for the Arctic Basin for 1962–2006 was digitized according to the annual data issues: 'Earthquakes in the USSR' and 'Earthquakes of the Northern Eurasia' [2–3]. The results of ionospheric observations (parameter foF2) at the drifting station 'North Pole-6' for 1958 and 1959 were also digitized. The WDC IPY site is permanently supplemented by new data from IPY and post-IPY periods. It contains links to the webpages of other Russian holders of Arctic and Antarctic data from IPY. Conversion of the analogue data into digital form will enable scientists to use longer and more complete time series of observations in the analysis and synthesis of research results from the IPY programme.

### 3. IPY-Info Portal

The Geophysical Center of the Russian Academy of Sciences, which includes both WDC–STP and WDC–SEP, was defined as the Disciplinary Centre for geophysical IPY data collection in the Russian scientific programme of IPY. In this regard, both WDCs participated in the creation of the Russian national 'IPY-Info' Portal [4], which is an integrated high-quality multidisciplinary information system that includes metadatabases; databases; and systems for data collection, storage, and communication.

Many Russian participants of IPY—institutes, scientific organizations, observing networks, expeditions, and so on—conducted investigations and observations in the polar regions. They provided the results of these observations and research to the appropriate Disciplinary Centres. The activities of the geophysics Disciplinary Centre include: accumulation of IPY geophysical datasets, preparation of metadata descriptions, registration of the datasets on the IPY-Info Portal, creation of a data catalogue on geophysics, development of data management technologies, provision of long-term storage of IPY data and metadata, and provision of free access to IPY information resources online.

The IPY-Info Portal was created at the All-Russian Research Institute of Hydrometeorological Information – WDC (RIHMI–WDC), Obninsk, Russia. The Portal provides a unified access point for the entirety of the comprehensive information on the polar regions obtained before, during, and after IPY. The Russian IPY-Info Portal serves as a component of the international Portal 'IPY Data and Information Service—IPYDIS'. Metadata circulate in the system of data gathering, storage, exchange, and processing at international and national levels.

### 4. Conclusions

The experiences of the Russian WDC–STP and WDC–SEP of managing the IPY website and storing IPY data could be used in future projects concerning the polar regions of the Earth. IPY data stored in the WDCs can be used freely by any member of the scientific community for future research activities.

### References

1. International Polar Year 2007–2008, <http://www.wdcb.ru/WDCB/IPY/IPY.html> [accessed on: September 2013]
2. Kondorskaya, N., *Earthquakes in the USSR in 1962–1991* (annual books). 'Nauka', Moscow, 1964–1997.
3. Starovoit O., *Earthquakes of the Northern Eurasia, 1992–2006* (annual books). Geophysical Survey RAS, Obninsk, 1997–2012.
4. Russian national IPY-Info Portal <http://www.mpg-info.ru/> [accessed on September 2013] (in Russian)

# Polar Magnetic Data at WDC Kyoto—Services under International Collaborations

**Toshihiko Iyemori<sup>1\*</sup>, Masahito Nosé<sup>1</sup>, Hiroaki Toh<sup>1</sup>, Masahiko Takeda<sup>1</sup>, Yukinobu Koyama<sup>1</sup>, Yoko Odagi<sup>1</sup>, Oleg Troshichev<sup>2</sup>, Alexander Yanzhura<sup>2</sup>**

<sup>1\*</sup> Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

<sup>2</sup> Arctic and Antarctic Research Institute, St. Petersburg, 199397, Russia

Email: iyemori@kugi.kyoto-u.ac.jp

**Summary.** Geomagnetic data from the polar regions are essential in Solar-Terrestrial Physics, and numerous observation sites have been positioned in these regions since the 19th century. The World Data Center for Geomagnetism, Kyoto (WDC Kyoto) collects geomagnetic data worldwide; some from the first Polar Year (1882–1883), but most have been collected since the International Geophysical Year (1957–1958) through international collaborations. In particular, there has been a strong collaboration with polar geomagnetic observatories since 1978—when WDC Kyoto started derivation of auroral electrojet (AE) indices—to improve the quality and shorten the time lag of AE index derivation. As a result, the situation has been much improved; however, continual efforts are still required. In this Forum, the role of international collaboration in data collection and services are presented and discussed, as well as recent attempts at interdisciplinary use of the data.

**Keywords.** Geomagnetic data, data transfer, service, quality, international collaboration.

## 1. Introduction

Many phenomena caused by solar wind-magnetosphere interaction—such as auroral electrojet (AE), DP-2 current system, Pc5 pulsations, and substorms—appear in the polar ionosphere because of magnetic connections with the outer part of the magnetosphere. Consequently, geomagnetic data have been widely used in Solar-Terrestrial Physics for many years (e.g., see [1]) and various observations, in particular geomagnetic observations, have been conducted at distant locations in the polar region under the international collaboration framework since the late 19th century.

At the forefront of these international project was the International Geophysical Year (IGY;1957–1958), and the World Data Centre (WDC) system was created for collection, exchange, and dissemination data obtained during (and after) IGY. WDC for Geomagnetism, Kyoto (WDC Kyoto) was founded at IGY, and has been collecting geomagnetic data worldwide since that time.

In this paper, the collection and service status of polar geomagnetic data at WDC Kyoto is presented, as well as the importance of international collaboration and current efforts to make data available across disciplines.

## 2. Data Collection and International Collaborations

The oldest data held at WDC Kyoto are in a data book published in 1882, the first Polar Year. Many data before IGY are in printed form. In fact, the annual reports of each observatory were typically published in printed form until recently, and are an important source of metadata for our recent effort of metadata database construction.

From IGY to the early 1990s, many of the data are on microfilms. After the mid-1990s, digital magnetometers became standard and, in the very early 2000s, the majority of data are sent in digital form to WDC Kyoto through the Internet.

The WDC system and the International Association of Geomagnetism and Aeronomy

played essential roles in collecting the data, and in recent years, the International Real-time Magnetic Observatory Network organizes data acquisition at greater than 130 geomagnetic observatories around the world.

WDC Kyoto started derivation of AE indices in 1978; when many of the data were obtained as analogue magnetograms and the polar geomagnetic data necessary for AE index derivation were given as analogue magnetograms on microfilm.

For derivation of AE indices, data need to be in digital form, and the digitizing of the analogue magnetograms on the microfilms proved challenging. In particular, Russian magnetograms were difficult to digitize due to the quality of the magnetograms or microfilms. To improve this quality, collaboration was initiated with the WDC in Moscow and the Arctic and Antarctic Research Institute (AARI).

### 3. RapidMag Project

During the period in which WDC Kyoto digitized analogue magnetograms, collaboration for digitization from Russian magnetograms was mainly with WDC Moscow. However, after digital magnetometers were installed, obtaining Russian data was mainly through cooperation with AARI, which operates the magnetic observatories in the polar region.

The 'Project for Upgrading Russian AE Stations' (PUREAS) was initiated in 1999 to support and develop a system for acquiring and transferring near real-time magnetometer data from six Russian stations, and so enable derivation of quick-look AE indices. The real-time data started to become available in 2001 from Pebek and in 2006 from Amderma. The current project 'RapidMag' is an extension of PUREAS, with the aim of supporting the Russian auroral zone magnetometer network.

Both projects were promoted through strong collaboration among the institutes in Japan (WDC Kyoto and National Institute of Information and Communications Technology), Russia (AARI and Institute of Geosphere Dynamics), and USA (Johns

Hopkins University Applied Physics Laboratory and University of Alaska) and have been very successful.

Currently, data from Russian AE stations are transferred to Kyoto continuously in near real-time.

### 4. Metadata Database

As written in the Introduction, polar geomagnetic data are essential in the study of various Solar-Terrestrial phenomena, and not only the resulting AE indices but also the original data have many users. The main applications for polar geomagnetic data are in space weather research and forecasts (or nowcasts). However, they are also used in examining long-term variation of the Earth's environment, and in medical research, education, and so on.

To make polar geomagnetic data more useful for other disciplines, WDC Kyoto is working as part of the Inter-university Upper atmosphere Global Observation NETWORK project, which aims to construct a metadata database for upper-atmospheric research.

### 5. Conclusions

Geomagnetic data have been obtained in the polar region through strong and wide-ranging international collaborations for more than 100 years. The RapidMag project has been one of the most successful cases. Metadata exchanges among data centres of various disciplines under an international framework such as the World Data System shall be essential for the promotion of interdisciplinary use of polar geomagnetic data.

**Acknowledgments.** Results presented in this paper are due to the efforts and achievements of numerous people and institutions over many years.

### References

1. Nishida, A., *Geomagnetic Diagnosis of the Magnetosphere*. Springer-Verlag, 1978

# Cosmic-ray Neutron Data held by WDC for Cosmic Rays

**Takashi Watanabe<sup>1\*,2</sup>, Masafumi Hirahara<sup>1</sup>, Fumio Abe<sup>1</sup>, Yuka Kadowaki<sup>1</sup>**

<sup>1\*</sup> Solar-Terrestrial Environment Laboratory, Nagoya University, Nagoya, 464-8601, Japan

<sup>2</sup> National Institute of Information and Communications Technology, Koganei, 184-8795

Email: takashi.watanabe@icsu-wds.org

**Summary.** The WDC for Cosmic Rays, Nagoya University, is a unique data centre providing with long-term and quality-controlled cosmic-ray intensities observed by a world-wide network of neutron monitors. The principal data held by the WDC are pressure-corrected and scale-adjusted one-hour counts of cosmic-ray neutrons. Cosmic-ray data obtained in the polar region are very important in study cosmic-ray related researches because the lowest portion ( $< 1$  GeV) of the energy spectrum of galactic cosmic rays, accessible by ground-based observations, is covered by monitors in this region. The database can be obtained via Web page, <http://center.stelab.nagoya-u.ac.jp/WDC/CR/>.

**Keywords.** World Data Centre, cosmic rays, neutron flux.

## 1. Introduction

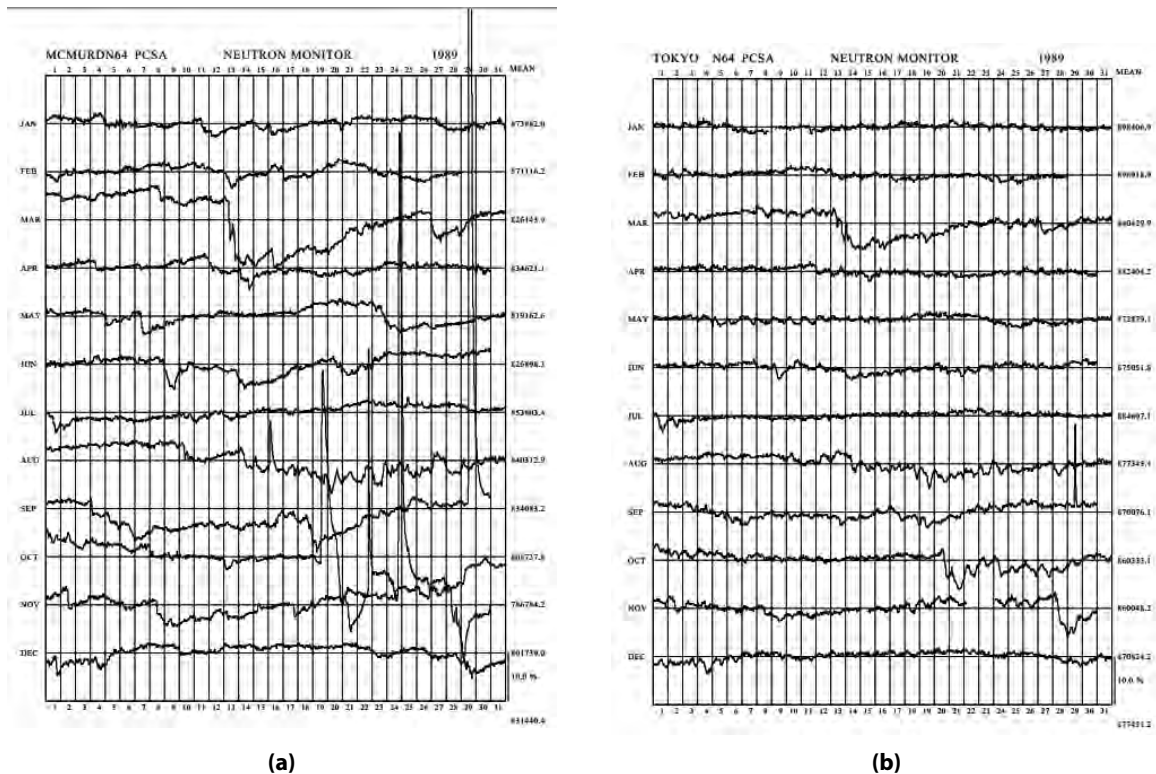
The World Data Center (WDC) for Cosmic Rays was established in 1957, under the former ICSU WDC System, at the Institute of Physical and Chemical Research (RIKEN) as a WDC-C2 to provide with database of cosmic-ray neutron observations in unified formats. In 1991, The WDC was moved to the Solar-Terrestrial Environment Laboratory, Nagoya University. This WDC is a unique data centre providing with quality-controlled and uniformly formatted cosmic-ray neutron data. Our database has been opened via Web page [1]. The database includes worldwide cosmic-ray neutron observations (pressure-corrected and scale-adjusted 1 hour counts) since 1953. In 1960s, more than 60 stations were operating in the world, and about 40 stations have been under operation in 2013. In total, 24 neutron monitors were operated in polar region ( $|\text{lat.}| > 65$  deg. in our case) since 1957, and 13 of them are under operation in 2013 (Apatity, Barentsburg, Cape Schmidt, Inuvik, Mawson, McMurd, Mirny, Norilsk, Sanae, South Pole, Terre Adelie, Thule, and Tixie Bay).

## 2. Cosmic Ray Observations

The energy range of cosmic rays covered by ground-based neutron monitors is controlled by the cut-off rigidity which is determined by the geomagnetic field, ranging from  $< 1$  GV in the polar region to 20 GV in the equatorial region. This means that cosmic rays with the energy higher than 20 GeV are observed in the low-latitude region and those higher than 0.5 GeV are observed in the polar region. In this reason, we need global network of neutron monitors to cover the wide energy range of cosmic rays (we need good local-time coverage also to observe transient phenomena).

## 3. Characteristics of Polar Cosmic-ray Data

Typical examples of neutron observations in high-latitude and low-latitude regions are shown in Figures 1(a) and (b), respectively. Figure 1(a) is monthly plots of neutron data taken at McMurd ( $R = 0.6$  GV), Antarctica, in 1989. Figure 1(b) is the same as 1(a), but for Tokyo ( $R = 11.6$  GV), Japan. Prominent upward spikes seen in these figures are Ground Level Enhancements (GLEs), which indicated that energetic solar particles were observed by ground-based monitors [2]. These events were observed when very strong solar



**Figure 1.** Two examples of monthly plots for 1989; (a) McMurd and (b) Tokyo. The vertical axis of each figure represents the rate of changes (%) with respect to the monthly average of hourly counts. Each horizontal line represents the 0% level with respect to the monthly average. The space between two horizontal lines shows the level of 10%.

flares took place mainly in the western solar hemisphere. Very energetic GLEs can be observed by low-latitude neutron monitors (e.g. 9 September 1989, in Figure 1(b)). The energy spectrum of each GLE is determined by assembling high-latitude and low-latitude neutron observations.

Important cosmic-ray phenomena also seen in Figure 1 are those known as Forbush decreases [3]. They are short-term decreases of the neutron flux enduring several days (or more). As shown in Figure 1(a), Forbush decreases are clearly seen in neutron data obtained in the polar region. A correlation between Forbush decreases and changes of atmospheric circulation has been suggested [4].

Concerning long-term (decades) variations of the cosmic-ray flux, an 11-year variation of neutron flux is the most prominent one, appearing in an opposite sense to the 11-year sunspot cycle. Correlation between the 11-year cosmic-ray flux and the global cloud cover has been a subject of extensive discussions [e.g. 5]. To study long-term variations of cosmic ray flux in

several decades, neutron data obtained in the polar region are very important because the modulation of the cosmic-ray flux caused mainly by changes of heliospheric magnetic environment will clearly appear in the low-energy part of the cosmic-ray energy spectrum. A provisional analysis of McMurd neutron data shows a general increasing tendency of the cosmic-ray flux in 1964–2012, reflecting low solar activity in current solar cycles.

#### 4. Concluding Remarks

Cosmic ray neutron observations in the polar region are very important to study cosmic-ray phenomena because the lowest-energy part of the cosmic-ray energy spectrum is only accessible in the high-latitude region. These data will be important not only to study cosmic-ray physics but also for researches of climate change. Due to difficulties to operate polar stations under very severe environmental conditions, strong support from scientific community, including data centres, will be important to help their long-term operation.

## References

1. WDC for Cosmic Rays, Nagoya University, <http://center.stelab.nagoya-u.ac.jp/WDCCR/> [accessed on August 2013]
2. M. A. Shea, Smart, D. F., Space Weather and the Ground-Level Solar Proton Events of the 23rd Solar Cycle, *Space Sci. Rev.*, 17, 161-188, 2012
3. S. E. Forbush, Time-Variations of Cosmic Rays, *Encyclopedia of Physics*, XLIX, Geophysics III, Part 1, 159-247, 1966
4. B. A. Tinsley, Geoffrey, M. Brown, Philip H. Scherrer, Solar variability influences on weather and climate: Possible connections through cosmic ray fluxes and storm intensification, *J. Geophys. Res.*, 94, 14783-14792, 1989
5. E. M. Agee, Kiefer, K., Corneti. E., Relationship of Lower-Troposphere Cloud Cover and Cosmic Rays: An Updated Perspective, *J. Climate*, 25, 1057-1060, 2012

# Observations and Data Handling at International Center for Space Weather Science and Education, Kyushu University, Japan

**Hideaki Kawano<sup>1\*</sup>, Shuji Abe<sup>1</sup>, Teiji Uozumi<sup>1</sup>, Dmitry G. Baishev<sup>2</sup>, Akimasa Yoshikawa<sup>1</sup>,  
MAGDAS/CPMN group**

<sup>1</sup> International Center for Space Weather Science and Education, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

<sup>2</sup> Yu.G.Shafer Institute of Cosmophysical Research and Aeronomy, Siberian Branch, Russian Academy of Sciences, 31 Lenin ave., Yakutsk 677980, Russia  
Email: hkawano@geo.kyushu-u.ac.jp

**Summary.** MAGnetic Data Acquisition System/Circum-pan Pacific Magnetometer Network (MAGDAS/CPMN) is currently the largest magnetometer network in the world. It includes eight magnetometers in Siberia and two in North America. These magnetometers are usually located in the subauroral region, but enter into the auroral region during active substorms. Publication and conservation of the obtained MAGDAS data are affected by the current situation regarding the research-funding system.

**Keywords.** MAGnetic Data Acquisition System/Circum-pan Pacific Magnetometer Network, global ground magnetometer network, time calibration using Global Positioning System, auroral-zone observations, subauroral-zone observations.

## 1. Introduction

The MAGnetic Data Acquisition System/Circum-pan Pacific Magnetometer Network (MAGDAS/CPMN; see Figure 1) [1] commenced ground magnetic field observations more than 10 years ago under the leadership of its Principle Investigator (PI), Kiyohumi Yumoto. In 2013, Yumoto resigned

from this position, and Akimasa Yoshikawa became the new PI. MAGDAS currently has 72 ground magnetometers—each with Global-Positioning-System-based time calibration—located mainly along three lines on the Earth: the meridian running through Japan, the meridian running through Egypt, and the magnetic

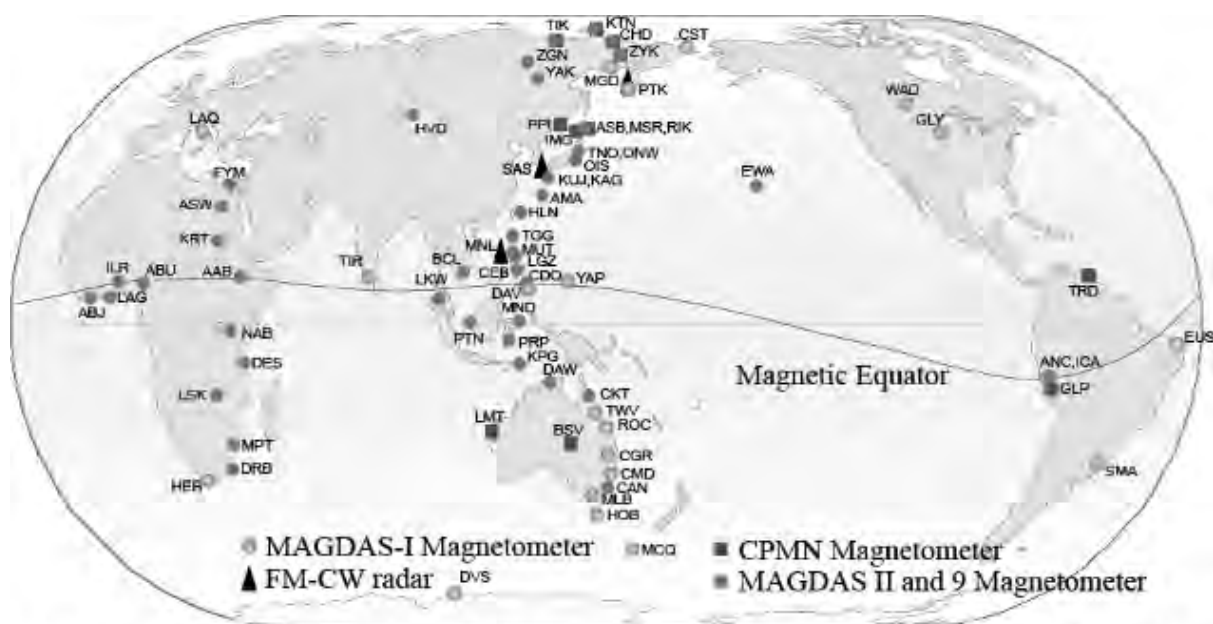


Figure 1. MAGDAS/CPMN (MAGnetic Data Acquisition System/Circum-pan Pacific Magnetometer Network)



equator. MAGDAS also has three frequency-modulated continuous-wave radars to measure the height of the ionosphere. These instruments are maintained mainly by the *International Center for Space Weather Science and Education (ICSWSE)* in collaboration with local host institutes. Many MAGDAS magnetometers send their one minute-resolution ('1-min' below) real-time data to ICSWSE via the Internet. In addition, the magnetometers can record one second-resolution ('1-sec' below) data to memory cards onsite, which the local hosts send to ICSWSE a few times a year via postal mail. These MAGDAS data have been used in greater than 330 scientific journal papers; suggesting the importance of MAGDAS within the research community.

This paper discusses measurements by MAGDAS of magnetic field variations in the auroral and subauroral regions; as well as publication, conservation, and sustainable use of the obtained data.

## **2. MAGDAS Observations in the Polar Regions**

As shown in Figure 1, MAGDAS has eight ground magnetometers located in Siberia and two in North America. These magnetometers are usually located in the subauroral region, but during active substorm intervals, the auroral zone moves to lower latitudes and they enter that zone. Thus, the magnetometers can be used to monitor substorm activities. In addition, they can be used to monitor magnetic storms since they have larger spatial scales than substorms.

The magnetometers in Siberia are especially important because substorm observations on this meridian have been less frequent than on the European and American meridians. It is furthermore important to compare ground magnetic field observations in the Siberian region with satellite observations on the same meridian. Hence, upgrading of the magnetometers in Siberia is ongoing, and it is planned that the number of magnetometers within this region will be increased.

## **3. Publication, Conservation, and Sustainable Use of Obtained Data**

MAGDAS data users are classified into three groups: the First, Second, and Third Parties. The First Party refers to ICSWSE members, who can use all MAGDAS 1-sec data; the Second Party refers to the local hosts of the MAGDAS stations, who can use all 1-min data and their own 1-sec data; and the Third Party refers to general users worldwide, who need agreement from the MAGDAS PI to use any MAGDAS data.

Calibration, conservation, and publication of the vast amounts of obtained MAGDAS data require considerable financial backing and human resources, and are supported mainly by Japanese governmental research funding. To maximize the chance of receiving these funds, it is important that a large number of research papers are published with first authors belonging to ICSWSE. The situation is similar for the local host institutes, and is why the First and Second Parties are given higher priority regarding data usage.

For sustained conservation and more open publication of MAGDAS data, it is vital to have a confirmed source of long-term funding. Towards that end, it would be desirable for a funding system to exist that appreciates *data publishing* and *data citation* as goals in themselves and supports them in the long term.

## **4. Conclusions**

For data publication, in general, and for space science in the polar regions using this published data, in particular, to become even more active than ever, it is imperative that the concepts of data publishing and data citation become popular and appreciated within the research community.

## **References**

1. Yumoto K., the MAGDAS Group, Space weather activities at SERC for IHY: MAGDAS. *Bull. Astr. Soc. India*, 35, 511–522, 2007

# Data and Science Lessons Learned from Atmospheric/Environmental Observations in Japan and Alaska

**Yasuhiro Murayama<sup>1\*</sup>, Takenari Kinoshita<sup>1</sup>, Seiji Kawamura<sup>1</sup>,  
Hironori Iwai<sup>1</sup>, Mamoru Ishii<sup>1</sup>, Minoru Kubota<sup>1</sup>,  
SALMON/Okinawa/Urban atmosphere data system development group, Fujitsu FIP  
corporation<sup>2</sup>**

<sup>1\*</sup> National Institute of Information and Communications Technology, 4-2-1 Nukui-kita, Koganei, Tokyo 184-8795 Japan

<sup>2</sup> 2-4-32 Aomi, Koto-ku, Tokyo 135-8686, Japan

Email: murayama@nict.go.jp

**Summary.** Data taken from various ground-based observations of lower, middle, and upper atmospheres and others in Alaska and Japan since 1993 have been collected, analyzed and archived for near real-time displays and usage, which employed high-speed experimental networks. This attempt started in 1998, as one of early efforts to develop a scientific data network system in Japan, and has been used to support various earth scientific research activities from Okinawa (24–27N); Koganei, Tokyo (36N); Wakkanai (45N); and Yamagawa (31N), as well as Alaska (59–65N).

**Keywords.** Alaska, Japan, ground-based remote-sensing, scientific data network system.

## 1. Introduction

Various types of ground-based experiments of radio and optical remote-sensing and related measurements for environmental and atmospheric observations have been carried out in the past few decades at Koganei, Tokyo (35.7N)—the headquarters of the National Institute of Information and Communications Technology—observatories in the Okinawa region (24.2–26.7N), and sites in south Kyushu (Yamagawa; 31.2N) and northern Hokkaido (Wakkanai; 45.4N). In addition, sites at King Salmon and Poker Flat in Alaska (58.7–65.4N) have been studied in cooperation with the Geophysical Institute, University of Alaska Fairbanks [1]. Targets were ionospheric, atmospheric, and oceanic phenomena from ground and ocean surface to several hundred kilometres in altitude. Some of these experiments are still continuing, while others have been stopped or are to be renewed.

A systematic effort started in 1998 to develop a *data network system* called 'SALMON' (originally an acronym of System for Alaska Middle and

upper atmosphere Observation data Network) [2], with the initial intention of handling the Alaskan observational datasets on a real-time basis with the then limited Internet bandwidth and available computer technology (this was approximately ten years after NASA's Planetary Data System operation had commenced; e.g., [3–4]).

## 2. SALMON Data Network System

The mechanism known as SALMON, included one of the early attempts in Japan to build a scientific data-handling system. From its conception in the late 1990s, SALMON has enabled near real-time data collection, archiving, interactive data browsing, and dissemination—especially in upper atmosphere science fields; bridging Alaskan experimental sites and the Japanese data processing facility across the Pacific Ocean using a number of high-speed Internet experiments: APAN (the Asia Pacific Advanced Network) [5], vBNS (very high-speed Backbone Network Service) (e.g., [6]), and Abilene/Internet2. Since the network capacity out of Alaska was limited to the order of 10<sup>1</sup> Mbps in the project's first years

(enhanced to  $10^{2-3}$  Mbps later), the initial design included avoiding interference to the commodity Internet of all Alaskan people.

### 3. Database and Scientific Potential

Although the US–Japan joint Alaskan program (1993–2006), which supported the databases of 11 experiments, has already ended, some of the databases are still being continued. As of 2005, 116 refereed journal articles had been published by international collaborators from more than ten institutes using data from these databases. In course of its development, SALMON has merged with a sister project—the Okinawa Subtropical Observation Data Network (which includes data from advanced polarimetric weather radar, 400-MHz wind profiler, long-range ocean high-frequency radar, etc.)—to share data transfer processes and computer resources. Moreover, data have been added to the database since 2006 from urban atmosphere observations at Koganei, Tokyo; including 1.3-GHz wind profiler, coherent carbon dioxide differential absorption/Doppler lidar, ceilometer, tower and surface meteorological measurements.

Such a variety of observational space–physics, aeronomical, meteorological, and oceanographic databases has led to new multidisciplinary and data-combination studies for cases that had not been initially envisaged. For example, a Fourier Transformed Infrared Spectrometer deployed in Alaska was used to study pollution gases (carbon monoxide (CO), hydrogen fluoride, methane, etc.) mainly in the lower atmosphere; however, its spectroscopic data was found useful to estimate the total amount of ‘stratomesospheric CO’ at heights around 30 km and above. Combining mesospheric wind velocity data observed with medium-frequency radar in Alaska, which is used mainly by the upper atmosphere dynamics community, furthermore enabled a study coupling stratomesospheric wind circulation and the trace gas transport [7].

### 4. Conclusions

In the context of the ICSU World Data System and international scientific data management, it is thought that new insights will be gained through a metadata handling experiment (predominantly descriptive metadata; i.e., data of data content) recently started by some of the authors in collaboration with the Japanese Inter-university Upper atmosphere Global Observation NETWORK (IUGONET) initiative. Efforts have also just commenced towards integrating and coordinating the archiving, publication, searching, reuse, analysis, and visualization of the datasets.

### References

1. Murayama, Y., Ishii, M., Kubota, M., Mori, H., Mizutani, K., Ochiai, S., Kasai, Y., Kawamura, S., Tanaka, Y., Masuko, H., Iguchi, T., Kumagai, H., Kikuchi, T., Sato, K., Collins, R. L., Watkins, B., Mark, C., Bristow, W. B., and Smith R. W., Comprehensive Arctic Atmosphere Observing System and Observed Results for “System Performance Demonstration”, *J. NICT*, 54 (1/2), 5–16, 2007 (<http://www.nict.go.jp/publication/shuppan/kihou-journal/journal-vol54no1.2/01.pdf> [accessed on: September 2013])
2. Oyama, S., Murayama, Y., Ishii, M., and Kubota, M., Development of Data Network Systems for Global Environmental Measurements, *J. NICT*, 49 (2), 253–257, 2002 (<http://www.nict.go.jp/publication/shuppan/kihou-journal/journal-vol49no2/5-01.pdf> [accessed on: September 2013])
3. McMahon, S. K., Overview of the Planetary Data System, *Planetary and Space Science*, 44(1), 3–12, 1996
4. Russell, C. T., The Planetary Data System, *Planet. Space Sci.*, 44(1), 1–2, 1996
5. Asia-Pacific Advanced Network, <http://www.apan.net/> [accessed on: August 2013]
6. Jamison, J., Nicklas, R., Miller, G., Thompson, K., Wilder, R., Cunningham, L., and Song, C., vBNS: not your father’s Internet, *IEEE Spectrum*, 35 (7), 38–46, 1998
7. Jones, N. B., Kasai, Y., Dupuy, E., Murayama, Y., Urban, J., Barret, B., Sinnhuber, M., Kagawa, A., Koshiro, T., Ricaud, P., and Murtagh, D., Stratomesospheric CO measured by a ground-based Fourier Transform Spectrometer over Poker Flat, Alaska: Comparisons with Odin/SMR and a 2-D model, *J. Geophys. Res.*, 112(D20), D20303, 2007

# Japanese Contribution to Super Dual Auroral Radar Network (SuperDARN)

**Nozomu Nishitani<sup>1\*</sup>, Akira S. Yukimatu<sup>2</sup>, Tsutomu Nagatsuma<sup>3</sup>, Tomoaki Hori<sup>1</sup>,  
Yoshizumi Miyoshi<sup>1</sup>**

<sup>1\*</sup> *Solar-Terrestrial Environment Laboratory, Nagoya University, Nagoya, Aichi, 464-8601, Japan*

<sup>2</sup> *National Institute of Polar Research, Tachikawa, Tokyo, 190-8518, Japan*

<sup>3</sup> *National Institute of Information and Communications Technology, Koganei, Tokyo, 184-8795, Japan*

Email: nisitani@stelab.nagoya-u.ac.jp

**Summary.** The Super Dual Auroral Radar Network (SuperDARN) is an international research collaboration to develop a global network of ground-based coherent-scatter radars operating in the high-frequency band with the aim of monitoring plasma convection mainly in the polar ionosphere. The Japanese community has been making a substantial contribution to SuperDARN in terms of both radar deployment/operation and the release of a useful data analysis platform to researchers worldwide. As an international joint project, SuperDARN continues to provide large volumes of data for exploring the dynamics of the Earth's magnetosphere, ionosphere, thermosphere, and mesosphere.

**Keywords.** Super Dual Auroral Radar Network, Ionosphere, Japanese contribution, Database, Integrated data analysis.

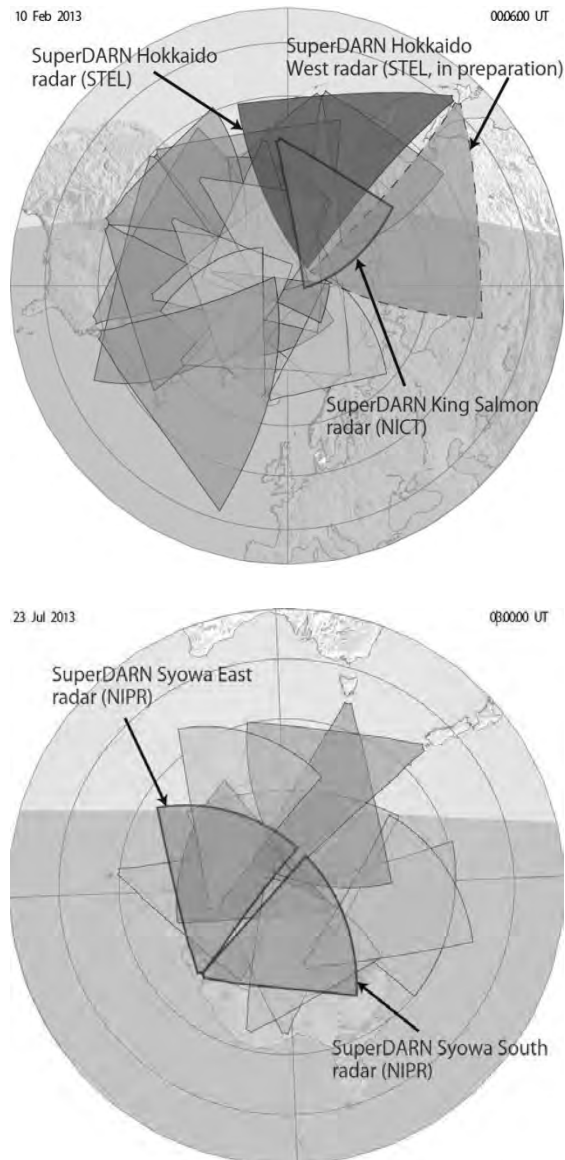
## 1. SuperDARN

The Super Dual Auroral Radar Network (SuperDARN) [1] is an international research collaboration to develop a global network of ground-based coherent-scatter radars operating in the high-frequency (HF) band. Since HF waves are scattered by field-aligned irregularities of electron density in the ionosphere, this radar system can observe horizontal motions of ionospheric plasma as Doppler shifts of backscattered signals. The SuperDARN radars are a suitable and very powerful tool for diagnosing bulk motions of plasma mainly in the E- and F-layer ionospheres over a range of a few thousand kilometres. The continuous operation and large-scale coverage in latitude and local time realized by the radar network (currently, 22 radars are operational in the northern hemisphere and 11 in the southern hemisphere) provide detailed profiles of evolution of ionospheric plasma motions, as well as global characteristics of ionospheric convection—which is linked to plasma convection in the magnetosphere. The radars can also be utilized as a network for studying the thermosphere and upper mesosphere.

## 2. Japanese Contribution to SuperDARN

The Japanese solar-terrestrial/space physics community has contributed considerably to SuperDARN since the very early stages of the project. SuperDARN officially began in 1995, with several radars constructed around magnetic latitudes of approximately 60° in both hemispheres, and looking towards the auroral zone and polar caps. The National Institute of Polar Research was one of the first institutions to join SuperDARN that had radars in the Antarctic region; that is, the Syowa South and Syowa East radars installed in 1995 and 1997, respectively, at the Syowa Station (a Japanese Antarctic base)—as shown on the world map in Figure 1. The National Institute of Information and Communications Technology then added the King Salmon radar in Alaska to the network in 2001, which expanded SuperDARN towards western Alaska, and even Far East Russia. More recently, the Solar-Terrestrial Environment Laboratory constructed the Hokkaido HF radar on Hokkaido Island, Japan in 2006. Among all of

those in SuperDARN, this radar still covers the region of lowest magnetic latitudes.



**Figure 1** Fields of view of SuperDARN radars in the (upper panel) northern and (bottom panel) southern hemispheres.

A recent Japanese effort is the new SuperDARN database and data analysis tool developed for the Energization and Radiation in Geospace (ERG) project [2]. The ERG spacecraft will provide data on the inner magnetosphere, which is strongly linked to plasma dynamics in the mid-latitude, subauroral, and auroral ionosphere. The science centre of the ERG project has released to the international community a database of archived SuperDARN data in Common Data Format, as well as a data analysis tool [3] based on the THEMIS Data Analysis

Software suite developed for the THEMIS mission [4]. This data analysis platform enables integrated data analysis; seamlessly combining various types of satellite and ground data.

### 3. Future perspectives

SuperDARN is still continuing to grow at this time. Several new radars are to be constructed in both hemispheres during upcoming years. In particular, the Hokkaido West radar is going to be constructed near the original Hokkaido HF radar and thereby this pair of radars will cover a greater area, from middle Siberia through Alaska. In addition, the Svalbard radar, a pair of Azores radars, the Dome C North radar, the Buckland radar, and several Russian radars will join SuperDARN; forming a vast field of view that covers mid-high latitudes of both hemispheres. As it has been for last decade, SuperDARN will remain a unique and outstanding observation resource; providing large volumes of data concerning the dynamics of the polar ionosphere/thermosphere/mesosphere.

### 4. Conclusions

The Japanese solar-terrestrial/space physics community has been playing an active role in SuperDARN by deploying/operating radar, as well as developing a useful data analysis platform. The international efforts of SuperDARN promote various types of studies on the polar ionosphere/thermosphere/mesosphere.

### References

1. Greenwald, R. A., et al., Darn/Superdarn: A Global View of the Dynamics of High-Latitude Convection/ *Space Sci. Rev.*, 71, 761–796, 1995
2. Miyoshi, Y., et al., The Energization and Radiation in Geospace (ERG) Project. In Summers, D., et al. (Eds.), *Dynamics of the Earth's Radiation Belts and Inner Magnetosphere*, *Geophys. Monogr. Ser. 199*, AGU, Washington D. C., 103–116, 2012
3. Hori, T., et al., An integrated analysis platform merging the SuperDARN data within the THEMIS tool developed by ERG-Science Center (ERG-SC). *Adv. Polar Sci.*, 24, 69–77, 2013
4. Angelopoulos V., The THEMIS mission. *Space Sci. Rev.*, 2008.

# Continuous Broadband Seismic Observation on the Greenland Ice Sheet under Greenland Ice Sheet Monitoring Network

**Seiji Tsuboi<sup>1\*</sup>, Masaki Kanao<sup>2</sup>, Yoko Tono<sup>1</sup>, Tetsuto Himeno<sup>3</sup>, Genti Toyokuni<sup>4</sup>, Dean Childs<sup>5</sup>, Trine Dahl-Jensen<sup>6</sup>, Kent R. Anderson<sup>7</sup>**

<sup>1\*</sup> Japan Agency for Marine-Earth Science and Technology, 3173-25 Showa-machi, Kanazawa-ku, Yokohama-shi 236-0001, Japan

<sup>2</sup> National Institute for Polar Research, 10-3 Midori-cho, Tachikawa-shi, Tokyo 190-8518, Japan

<sup>3</sup> Seikei University, 3-3-1 Kichijoji-kitamachi, Musashino-shi, Tokyo 180-8633, Japan

<sup>4</sup> International Research Institute of Disaster Science, Tohoku University, 6-6 Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-8578, Japan

<sup>5</sup> IRIS PASSCAL Instrument Center, 801 Leroy Place, New Mexico Tech., Socorro, NM 87801, USA

<sup>6</sup> Geological Survey of Denmark and Greenland, Øster Voldgade 10, 1350 København K - Tlf.: 38142000, Denmark

<sup>7</sup> Incorporated Research Institutions for Seismology, 1200 New York Avenue NW, Suite 400, Washington DC 20005, USA

Email: tsuboi@jamstec.go.jp

**Summary.** The Greenland Ice Sheet monitoring Network (GLISN) is a new, international, broadband seismic capability for Greenland that is being implemented through collaboration between Denmark, Canada, France, Germany, Italy, Japan, Norway, Poland, Switzerland, and the USA. It is designed to monitor glacial earthquakes, the occurrences of which are reported to have increased this century. Currently, about 30 seismic observatories, designated as GLISN contributing stations, are located in and around Greenland. Broadband seismograms obtained from these stations are archived and opened to the research community through the Data Management Centre of the Incorporated Research Institutions for Seismology.

**Keywords.** Glacial earthquakes, Greenland ice sheet, broadband seismograph.

## Introduction

Glacial earthquakes along the edges of Greenland have been observed with strong seasonality and increasing frequency since 2002 by continuously monitoring data from the Global Seismographic Network [1-2]. Such glacial earthquakes, which are of magnitude 4.6–5.1, may be modelled a large glacial ice mass sliding several metres downhill on its basal surface for a duration of 30–60 s. Glacial earthquakes have been observed at seismic stations within Greenland, but currently the coverage is very sparse. To define the fine structure and detailed mechanisms of glacial earthquakes within the Greenland ice sheet, a broadband, real-time seismic network needs to be installed throughout the ice sheet and its perimeter. The National Institute for Polar Research and the Japan Agency for Marine-Earth Science and Technology are members of The

Greenland Ice Sheet monitoring Network (GLISN), and since 2011, have jointly collaborated with the Polar Services of the Incorporated Research Institutions for Seismology (IRIS) to install and maintain a broadband seismograph station on the ice sheet.

## Greenland Ice Sheet Station

ICE-S (DK.ICESG) was installed in June 2011, in collaboration with IRIS Polar Services (Figure 1). The station is equipped with a CMG-3T broadband seismometer and a Quanterra Q330 data logger. The station was visited in May 2012 and 2013, and one year of continuous records was successfully retrieved from the broadband seismometer. Moreover, the telemetry system was updated to enable real-time monitoring of the station. Figure 2 shows observed seismograms for the West of Greenland

earthquake that occurred on 11 April 2013 (Mw4.6), and demonstrates the high quality of ICESG seismograms. Continuous broadband seismograms from ICESG station are also available via the IRIS Data Management Centre as soon as they are retrieved from the observatory.

### Conclusions

GLISN was initiated in 2010 and will continue for another 5 years. Data from this observation network are freely available to the research

community.

**Acknowledgments.** This work is supported by JSPS KAKENHI 24403006.

### References

1. Ekström, G., Nettles, M., Abers, G. A., Glacial earthquakes. *Science*, 302, 622–624, 2003
2. Elström, G., Nettles, M., Tsai, V. C., Seasonality and increasing frequency of Greenland glacial earthquakes. *Science*, 311, 1756–1758, 2006.

The Greenland Ice Sheet Monitoring Network (GLISN) - 2013

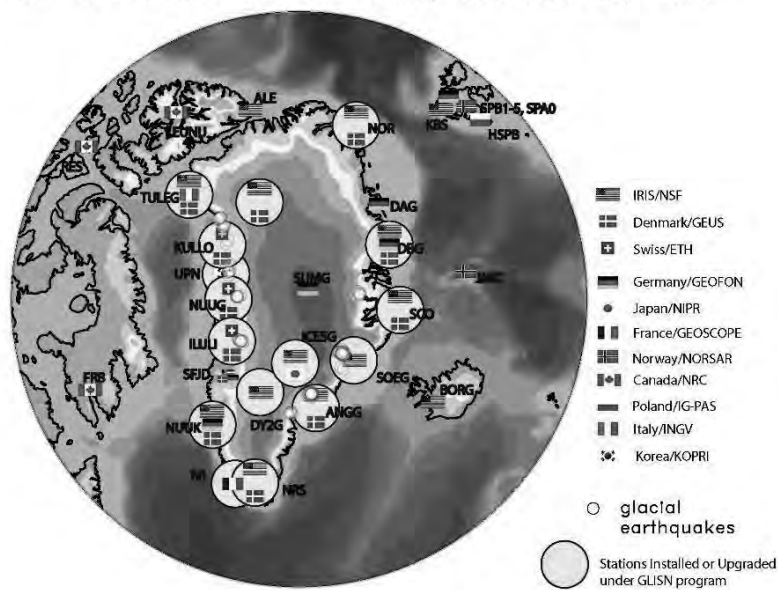


Figure 1. GLISN station map.

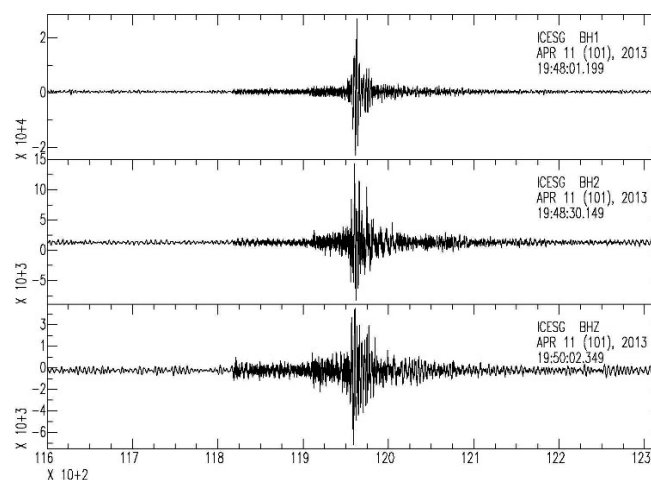


Figure 2. Three component seismograms of ICESG station for 11 April 2013, West of Greenland earthquake.

# Current Status of Science Data Archives for the Data Obtained by the Japanese Antarctic Research Expedition

**Akira Kadokura<sup>1\*</sup>, Masaki Kanao<sup>1</sup>, Masaki Okada<sup>1</sup>**

<sup>1\*</sup> National Institute of Polar Research, Research Organization of Information and Systems, 10-3, Midori-cho, Tachikawa-shi, Tokyo 190-8518, Japan  
Email: kadokura@nipr.ac.jp

**Summary.** The current status of the data archives for the scientific data obtained by the Japanese Antarctic Research Expedition is briefly explained. Data in various scientific fields have been archived at the National Institute of Polar Research and other responsible institutions. Those data are mainly categorized into two types: (1) steady long-term observation data and (2) specific purpose research observation data. The data policies and statuses of the archives and publications for these two types are different from each other. In this presentation, the present situation regarding the archived data will be shown and explained, as well as some examples of the scientific results. Future directions will also be discussed.

**Keywords.** Antarctic, science data, data archive, publication, Japanese Antarctic Research Expedition.

## 1. Introduction

Scientific observations in various scientific fields have been performed by the Japanese Antarctic Research Expedition (JARE) at Antarctic stations and out in the field since 1957. Scientific data obtained in JARE activities have been archived at the National Institute of Polar Research (NIPR) and other responsible institutions in Japan. In this presentation, the current status of the archived data will be shown and explained, as well as some examples of the scientific results.

## 2. Observations by JARE

Observations by JARE are divided mainly into two categories: steady (long-term, continuous) observations and specific purpose research observations. In the first category, long-term continuous observations have been carried out of meteorological parameters, atmospheric minor constituents (e.g., carbon dioxide (CO<sub>2</sub>)), snow accumulation, auroral and geomagnetic activities, ionosphere, seismic activity, gravity, penguin populations, ocean chemical analysis, and marine biology. In the second category, a large number of short and middle-term projects have been carried out thus far.

Archiving and publication of the steady observation data have been continuously maintained under a well-defined data policy by both an appropriately appointed research group at NIPR and departments in other responsible institutions. Conversely, responsibility for the specific observation data lies with each project team or individual scientist, each under an independent data policy. As a result, some data are fittingly archived and can be accessed from well-maintained webpages, but others are very difficult to access.

## 3. Data Examples

### 3.1 Steady observation data

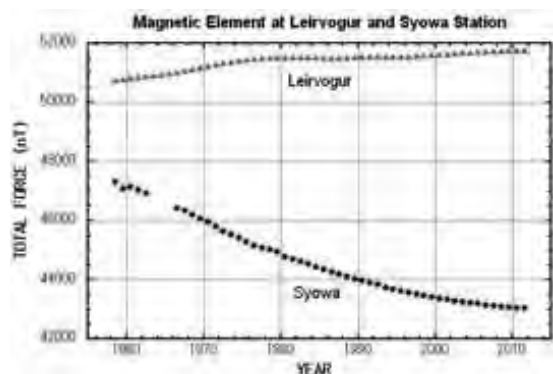
Observation of meteorological parameters at Syowa Station has been performed by JARE members from the Japan Meteorological Agency (JMA). The resulting data are archived at the office of Antarctic observations in JMA and can be accessed from their well-maintained webpage (in Japanese).

Ionospheric observations at Syowa Station have been conducted by the National Institute of Information and Communications Technology (NICT). These data are archived by the World Data Centre (WDC) for Ionosphere and Space Weather



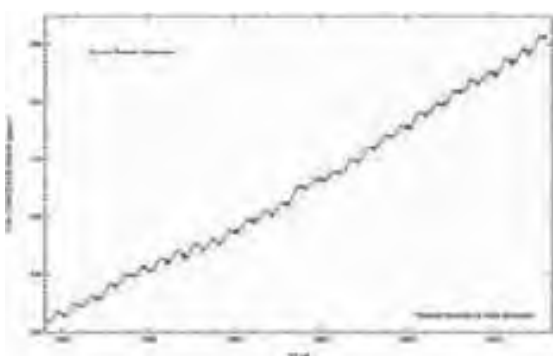
at NICT and can be accessed through their website [1].

Ocean chemical analysis data were originally collected by the Hydrographic and Oceanographic Department of the Japan Coast Guard and archived in the Japan Oceanographic Data Center [2]. However, these tasks have been performed by NIPR since 2012 and data are available through a new webpage [3].



**Figure 1.** Geomagnetic total field observation.

Other steady observations have also been carried out by NIPR, with the data archived by the respective research group within NIPR. Auroral and geomagnetic observations have been maintained by the Space and Upper Atmospheric Sciences (SUAS) group, and the data archived at the Auroral Data Center (formerly, WDC for Aurora) in the Polar Data Center [4]. Figure 1 shows an example of long-term observations of geomagnetic total field at Syowa Station compared with the same observations at a geomagnetic conjugate station in Iceland.



**Figure 2.** Concentration of CO<sub>2</sub>

As regards other observations: the

Meteorology and Glaciology group are responsible for atmospheric minor constituents and snow accumulation (see Figure 2 for an example of long-term variation of CO<sub>2</sub> concentrations at Syowa Station), the Geoscience group for seismic activity and gravity, and the Bioscience group for penguin populations and marine biology.

### 3.2 Specific research observation data

Metadata for specific research observations carried out to date can be seen via the website of the Polar Data Center in NIPR [5]. Details of the meta-database are described in [6]. A vast database for Antarctic meteorite, animal and plant specimens, and Super Dual Auroral Radar Network data are maintained by the Geoscience, Bioscience, and SUAS groups, respectively.

## 4. Conclusions

The current status of the data archives for the scientific data obtained by JARE will be briefly explained in the presentation.

**Acknowledgments.** All archived data were obtained due to the great efforts of the many past JARE members.

## References

1. National Institute of Information and Communications Technology, [http://wdc.nict.go.jp/wdc\\_e.html](http://wdc.nict.go.jp/wdc_e.html) [accessed on: September 2013]
2. Japan Oceanographic Data Center, <http://www.jodc.go.jp/index.html> [accessed on: September 2013]
3. National Institute of Polar Research, <http://biows.nipr.ac.jp/JARE/> [accessed on: September 2013]
4. Datacenter for Aurora in NIPR, <http://polaris.nipr.ac.jp/~aurora/> [accessed on: September 2013]
5. National Institute of Polar Research, <http://scidbase.nipr.ac.jp/> [accessed on: September 2013]
6. Kanao, M., Okada, M., Kadokura, A., Metadata Management at the Polar Data Center (PDC) of the National Institute of Polar Research (NIPR). In: *Proceedings of the International Forum on 'Polar Data Activities in Global Data Systems'*, Tokyo, 2013

# Metadata Management at the Polar Data Center (PDC) of the National Institute of Polar Research (NIPR), Japan

**Masaki Kanao<sup>1\*</sup>, Masaki Okada<sup>1</sup>, Akira Kadokura<sup>1</sup>**

<sup>1\*</sup> National Institute of Polar Research, Research Organization of Information and Systems, 10-3, Midori-cho, Tachikawa-shi, Tokyo 190-8518, Japan  
Email: kanao@nipr.ac.jp

**Summary.** The Polar Data Center of the National Institute of Polar Research has had a responsibility to manage Japanese data as a National Antarctic Data Centre for the past two decades. During the International Polar Year 2007–2008, a significant number of multidisciplinary metadata were compiled mainly from IPY-endorsed projects involving Japanese activities. The amalgamated metadata share a tight collaboration with the Global Change Master Directory, the Polar Information Commons, as well as the newly established World Data System.

**Keywords.** International Polar Year, Polar Data Center, metadata management, Global Change Master Directory, World Data System.

## 1. Introduction

The Polar Data Center (PDC) of the National Institute of Polar Research (NIPR) is the Japanese National Antarctic Data Centre (NADC), and has a strong relationship with the Scientific Committee on Antarctic Research (SCAR) under the International Council for Science (ICSU). In terms of Japanese activities during the International Polar Year 2007–2008 (IPY), PDC compiled many of the polar data from IPY-endorsed projects [1].

In this presentation, the state of metadata management involving Japan—particularly that concerned with the tasks of PDC—are demonstrated. Moreover, tight links have been forged with other science bodies of ICSU, such as the Committee on Data for Science and Technology and the new World Data System.

## 2. Polar Data Center

At the 22nd Antarctic Treaty Consultative Meeting in 1998, affiliate countries were obliged to ensure that scientific data collected from Antarctic programmes could be freely exchanged and utilized. Following the Articles of the 1998 Antarctic Treaty, each country is required to establish a NADC and to appropriately provide

data collected by those scientists involved. PDC at NIPR performs this NADC-function for Japan.

PDC established a data policy in February 2007, based on the requirements of the Standing Committee on Antarctic Data Management (SCADM) of SCAR. This policy then contributed to the subsequent SCAR Data and Information Management Strategy [2–3].

## 3. Metadata Management

PDC has the significant task of archiving and delivering digital data obtained from the polar regions. Summary information of all the archived data (i.e., metadata) is available to the polar science community, together with more general information of interest. The compiled metadata describe all types of scientific disciplines (space and upper-atmospheric science, meteorology and glaciology, geoscience and bioscience) from both long- and short-term projects in the Arctic and Antarctic, in particular, data collected by the Japanese Antarctic Research Expedition [4].

As of June 2013, a total of 255 records have been accumulated in the scientific meta-database provided by PDC including metadata from IPY-endorsed projects [5]. A new content management system for providing these

metadata has been in place since April 2011.

The scientific database provided by PDC is closely connected to the Antarctic and Arctic Master Directories (AMDs) in the Global Change Master Directory (GCMD) of the National Aeronautics and Space Administration. In addition to data from IPY, those from Japanese national and other international projects have been compiled in GCMD. Moreover, 279 metadata records have been amalgamated (up to June 2013) within the Japanese Antarctic portal of GCMD.

Although PDC stores its metadata in their original format, this includes the main items listed in the GCMD Directory Interchange Format. Moreover, there are tight cross-linkages between corresponding metadata in the AMDs and PDC. Metadata collected by Japanese IPY projects have also been compiled in an IPY portal within GCMD. To date, a total of 250 metadata records contributed by Japan are in the IPY portal, and constitute a significant proportion of all IPY metadata in GCMD.

The scientific results from IPY have emerged, but it is clear that deeper understanding will require the creative use of the myriad data from various science disciplines. Many of these projects provided well-coordinated observation platforms, and a number of these continue in the post-IPY era. The vast amount of data accumulated during and after IPY could be its most important legacy if these are well preserved and utilized [6].

#### 4. Conclusions

The status of metadata management in PDC of NIPR is summarized in this short presentation. Many dedicated data service tasks have been conducted by PDC staff as the Japanese NADC. Several aspects of the scientific data collected in the polar regions have great significance for global environmental research in this century. To construct an effective framework for a long-term polar data strategy, data must be made available promptly and new Internet technologies—such a repository network service similar to the Polar Information Commons—must be employed.

**Acknowledgments.** The authors would like to express their appreciation to a significant number of collaborators associated with IPY activities both in national and international projects. They also acknowledge the members of SCADM and the IPY Data Committee under the IPY Joint Committee for their great efforts to adhere to the management issues of polar data.

#### References

1. Sato, N., Ito, H., Kanao, M., Kanda, H., Naganuma, T., Ohata, T., Watanabe, K., Yamanouchi, Y., Engaging Asian Nations in IPY: Asian Forum for Polar Sciences (Japanese Section). In Krupnik, I., Allison, I., Bell, R., Cutler, P., Hik, D., Lopez-Martinez, J., Rachold, V., Sarukhanian, E., Summerhayes, C. (Eds.), *Understanding Earth's Polar Challenges: International Polar Year 2007–2008—Summary by the IPY Joint Committee*, Art Design Printing Inc., Edmonton, Alberta, 5.3, 555–574, 2011
2. Finney, K., SCAR Data and Information Management Strategy (DIMS) 2009–2013. In Summerhayes, C., Kennicutt, C. (Eds.), *SCAR Ad hoc Group on Data Management*, Scott Polar Research Institute, Cambridge, 1–34, 2009
3. de Bruin, T., Finney, K., The SCAR Data Policy. *SCAR newsletter*, 27, 3, 2011
4. Kanao, M., Kadokura, A., Okada, M., Yamnouchi, T., Shiraishi, K., Sato, N., Parsons, M. A., The State of IPY Data Management: The Japanese Contribution And Legacy. *CODATA Data Science Journal*, 12, WDS124–WDS128, 2013
5. <http://scidbase.nipr.ac.jp/> [accessed on: September 2013]
6. Parsons, M.A., de Bruin, T., Tomlinson, S., Campbell, H., Godoy, Ø., LeClert, J., the IPY Data Policy and Management Subcommittee, The State of Polar Data—the IPY Experience. In Krupnik, I., Allison, I., Bell, R., Cutler, P., Hik, D., Lopez-Martinez, J., Rachold, V., Sarukhanian, E., Summerhayes, C. (Eds.), *Understanding Earth's Polar Challenges: International Polar Year 2007–2008—Summary by the IPY Joint Committee*, Art Design Printing Inc., Edmonton, Alberta, 3.11, 457–476, 2011

# Data Management Plans for the Southern Ocean Observing System (SOOS)

**Louise Newman<sup>1\*</sup>, Kim Finney<sup>2</sup>, Michael Meredith<sup>3</sup>, Oscar Schofield<sup>4</sup>**

<sup>1\*</sup> SOOS International Project Office, Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, 7000, Australia

<sup>2</sup> Australian Antarctic Division, Channel Highway, Kingston, Tasmania, 7000, Australia

<sup>3</sup> British Antarctic Survey, High Cross, Madingley Rd, Cambridge, United Kingdom

<sup>4</sup> Rutgers University, 100 George Street, New Brunswick, New Jersey, 08901, United States  
Email: newman@soos.aq

**Summary.** The Southern Ocean is fundamental to the operation of the Earth system; playing a central role in global climate and planetary-scale biogeochemical cycles. The Southern Ocean is changing rapidly, and the critical need to observe and understand the Southern Ocean is well established; however, the harsh conditions and remote location have led to it being the most under-sampled region of the world. Sustained observations are required to detect, interpret, and respond to the physical, chemical, and biological changes that are, and will continue to be, measured. The Southern Ocean Observing System (SOOS) was developed to integrate the global assets of the international community in order to enhance data collection, provide access to datasets, and guide the development of strategic-sustained-multidisciplinary science in the Southern Ocean. For SOOS to succeed, it is critical that a data system be established that ensures both past and future datasets are accessible and of known quality. This presentation will outline the SOOS data management policy and strategic plans.

**Keywords.** Southern Ocean, observations, strategy, policy, data sharing.

## 1. Introduction

The Southern Ocean is fundamentally important to the Earth system, influencing climate, biogeochemical and ecological cycles over very large scales. Society faces many difficult and pressing issues, including mitigation and adaptation to climate change and sea-level rise, managing the effects of ocean acidification, and conservation of marine resources and biodiversity. These cannot be addressed without an improved understanding of Southern Ocean processes and feedbacks, and the sensitivity of these to changes.

Limited observations suggest the Southern Ocean is changing. Yet chronic under-sampling makes the causes and consequences of such changes difficult to assess, and limits the effectiveness of any response. The Southern Ocean Observing System (SOOS) has thus been created to facilitate integration of resources,

enhance data collection and access, and guide the sustained development of strategic, multidisciplinary science in the Southern Ocean.

SOOS is an initiative of the Scientific Committee on Antarctic Research and the Scientific Committee on Oceanic Research. Opening officially at the end of 2011, SOOS is coordinated from an International Project Office hosted by the Institute for Marine and Antarctic Studies at the University of Tasmania (Hobart, Australia).

The mission of SOOS is to coordinate and expand the efforts of all nations that gather data from the Southern Ocean. SOOS is developing a coherent and efficient observing system to deliver the observations required to address key scientific and societal challenges.

The objectives of SOOS are to:

- Design and implement a comprehensive and multidisciplinary observing system for the

Southern Ocean.

- Advocate and guide the development of new observation technologies.
- Unify current observation efforts and leverage further resources.
- Effectively integrate and coordinate national and international projects and programmes, across traditional disciplinary boundaries and between nations.
- Facilitate and develop a data system that provides seamless access to essential data products for the Southern Ocean.

For SOOS to succeed, it is critical that a data system be established that ensures both past and future datasets are accessible and of known quality.

The SOOS Data Management strategy is based on the following principles:

1. Open access to data.
2. Use of existing resources and data centres where possible.
3. Improvement of access to and quality of historical data.
4. Good data management practices.
5. Protocols for data collection, quality control, and archiving.

SOOS has developed a Data Management Subcommittee tasked with identifying and coordinating cost-effective, collaborative mechanisms for managing and publishing observational data considered to be within the

purview of the SOOS initiative. The goal of this Subcommittee is that SOOS end-users have ready and free access to relevant data and products.

A key deliverable of SOOS is the development of the SOOS Data Portal, which will provide one-stop access to distributed data archives holding all SOOS-related data. The majority of physical and biological datasets are handled by separate data systems, making interdisciplinary research difficult. SOOS is working with the Australian Integrated Marine Observing System e-Marine Information Infrastructure and the Australian Ocean Data Network to develop the SOOS Data Portal; an operational prototype of which was released in January.

## 2. Conclusions

This presentation will provide a short overview of SOOS and its scientific objectives, before outlining the SOOS Data Management policy, the data requirements, and the strategic plans currently being developed to help meet those requirements.

**Acknowledgments.** The Southern Ocean Observing System acknowledges the support of the host institution, the Institute for Marine and Antarctic Studies, University of Tasmania; and the office sponsors: the Australian Antarctic Division, Antarctica New Zealand, and the New Zealand Antarctic Research Institute.

# NICT Science Cloud: Distributed Storage System and Parallel Data Processing Applicable for Polar Research Data

**Ken T. Murata<sup>1\*</sup>, Hidenobu Watanabe<sup>1</sup>, Kazunori Yamamoto<sup>1</sup>, Fusako Isoda<sup>2</sup>, Eizen Kimura<sup>3</sup>, Osamu Tatebe<sup>4</sup>, Masahiro Tanaka<sup>4</sup>, Kentaro Ukawa<sup>5</sup>, Kazuya Muranaga<sup>5</sup>, Yutaka Suzuki<sup>5</sup>, Takamichi Mizuhara<sup>6</sup>**

<sup>1\*</sup> National Institute of Information and Communications Technology, 4-2-1 Nukui-Kitamachi, Koganei, Tokyo, 184-8795, Japan

<sup>2</sup> Sciece Service Co., LTD., 1-3-15 Nihonbashi-Horidomecho, Chuo-ku, Tokyo, 103-0012, Japan

<sup>3</sup> Ehime University, Situkawa, Toon City, Ehime, 791-0295, Japan

<sup>4</sup> University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577, Japan

<sup>5</sup> Systems Engineering Consultants Co., LTD., Setagaya Business Square, 4-10-1 Yoga, Setagaya-ku, Tokyo, 158-0097, Japan

<sup>6</sup> CLEALINKTECHNOLOGY Co.,Ltd., 1-7 Hikaridai, Seika-cho, Souraku-gun, Kyoto, 619-0237, Japan  
Email: ken.murata@nict.go.jp

**Summary.** Data-oriented science is considered as the forth paradigm for upcoming Big Data science, including polar research. The science cloud at the National Institute of Information and Communications Technology (NICT) was designed and constructed as a basis for data-oriented science in 2010. In the present paper, the basic concepts behind the NICT Science Cloud are discussed, especially focussing on Big-data processing. Information and communication technologies are demonstrated that are designed to transfer, steward, process, and publicize large-scale data acquired for analysis in Solar-Terrestrial Physics.

**Keywords.** Distributed storage system, parallel data processing, big data, National Institute of Information and Communications Technology Science Cloud.

## 1. Introduction

During the past 50 years, along with the appearance and development of high-performance computers and supercomputers, numerical simulation is considered to be a third methodology for science, following theoretical (first) and experimental and/or observational (second) approaches. The variety of data yielded by the second approach has been steadily increasing. This is due to the progress of technologies, experiments, and observations. The amount of data generated by the third methodology (i.e., numerical simulation) has also been rapidly increasing. This is due to the tremendous developments in, and programming techniques of, supercomputers.

The majority of data files created by both experiments/observations and numerical simulations are saved in digital formats and analyzed on computers. The researchers (domain

experts) using these methods are interested in not only how to perform experiments/observations and numerical simulations but also what information (new findings) can be extracted from data. However, data does not usually tell anything about the science; scientific information is implicitly hidden in the data. Researchers have to 'mine' data to extract information and archive new sciences; this is a basic concept of data-intensive (data-oriented) science for 'Big Data'.

As the scales of experiments/observations and numerical simulations grow larger, new techniques and facilities are required to extract information from the increasing amount of data files. The techniques are grouped under the topic of informatics, which is a fourth methodology and is relatively new to science. All methodologies must be supported by appropriate infrastructure and computational facilities; for example, polar

ionosphere and atmosphere are studied in polar science via ground-based observatories and also numerical simulations, respectively. The facility that often underpins informatics, and which deals with large-scale data from observatories and supercomputers, is a computational cloud system for science.

## 2. NICT Science Cloud

This paper proposes a cloud system for informatics that has been developed at the National Institute of Information and Communications Technology (NICT), Japan. The NICT Science Cloud [1] is the first open cloud system for scientists who are going to carry out the informatics for their own science.

The NICT Science Cloud is not for simple uses. Many functions are expected of the Science Cloud, such as data standardization, data collection and crawling, large and distributed data storage, security and reliability, databasing and meta-databasing tasks, data stewardship, long-term data preservation, data rescue,, data mining, parallel processing, data publication and provision, semantic web, three- and four-dimensional visualization, out- and in-reach, and capacity building.

Figure 1 is a schematic of the NICT Science Cloud. Data from both observations and simulations are stored in the storage system of

the Science Cloud. It should be noted that there are two types of observational data. The first type is from archive sites, which are outside of the NICT Science Cloud: these are data that can be downloaded through the Internet and channelled to the Science Cloud. The other data type is generated by equipment connected directly to the Science Cloud. These equipment-centric configurations are often called sensor clouds.

## 3. Conclusions

In the present talk, the NICT Science Cloud will first be introduced. Next, the efficiency of the Science Cloud will be demonstrated through several scientific results achieved using this Cloud System. Through these discussions and demonstrations, the performance potential of the Science Cloud for many types of research fields will be revealed.

**Acknowledgments.** The present work is based on the NICT Science Cloud project.

## References

1. Murata, K. T., Watari, S., Nagatsuma, T., Kunitake, M., Watanabe, H., Yamamoto, K., Kubota, Y., Kato, H., Tsugawa, T., Ukawa, K., Muranaga, K., Kimura, E., Tatebe, O., Fukazawa, K. and Murayama, Y., A Science Cloud for Data Intensive Sciences, *Data Science Journal*, 12, WDS139–WDS146, 2013.

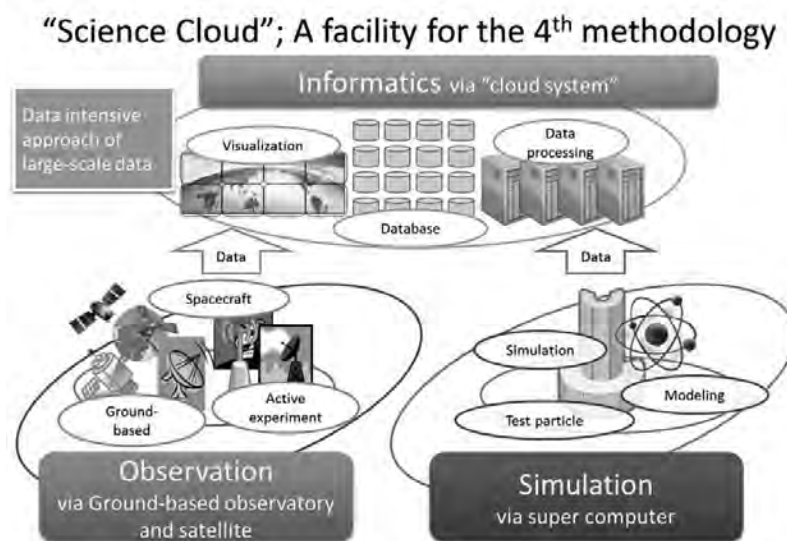


Figure 1. Schematic of NICT Science Cloud.

# Operation of Data Acquisition, Transfer, and Storage System for Worldwide Observation Networks

**Tsutomu Nagatsuma<sup>1\*</sup>, Ken T. Murata<sup>1</sup>, Kazunori Yamamoto<sup>1</sup>, Takuya Tsugawa<sup>1</sup>, Hideaki Kitauchi<sup>1</sup>, Takumi Kondo<sup>1</sup>, Hiromitsu Ishibashi<sup>1</sup>, Michi Nishioka<sup>1</sup>, Masaki Okada<sup>2</sup>**

<sup>1\*</sup> National Institute of Information and Communications Technology, 4-2-1 Nukui-kita, Koganei, Tokyo 184-8795, Japan

<sup>2</sup> National Institute of Polar Research, 10-3 Midorichou, Tachikawa, Tokyo 190-8518, Japan  
Email: tnagatsu@nict.go.jp

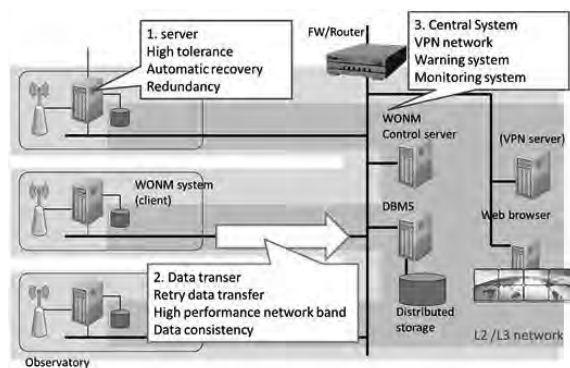
**Summary.** This paper presents operation of a system to acquire, transfer and store data from worldwide observation networks named the Wide-area Observation Network Monitoring system and developed at the National Institute of Information and Communications Technology (NICT). This system provides easier management of data collection than legacy systems by means of autonomous system recovery, periodical state monitoring, and dynamic warning procedures. Observatories for space weather prediction and research work located worldwide have been equipped with this system, which is connected to the NICT Science Cloud. A demonstration and discussion will be presented concerning this challenging system, especially from the viewpoint that worldwide observatories are easily operated through a web application.

**Keywords.** Wide-area Observation Network Monitoring system, worldwide observatories, Showa station, ionosonde, National Institute of Information and Communications Technology Science Cloud.

## 1. Introduction

The National Institute of Information and Communications Technology (NICT) has been establishing a global Space Weather Monitoring Network for space weather forecast and research [1]. The basic concept of this project is to improve the reliability and probability of space weather forecast by introducing real-time data obtained from a global network of space-weather-related observational facilities, for example, ionosondes, magnetometers, high-frequency radars, and Global Position System receivers [2]. Data archives and analyses of archived data are also important for development of space weather forecasting. In this project, NICT operates about 30 observatories covering a wide-area of the Earth. All data will be transferred to NICT and stored in a large-scale storage system in the NICT Science Cloud [3] on a real-time basis. However, managing the entire systems operation has become increasingly challenging because the system contains a large number of observational instruments each having its own characteristics.

The probability of a problem occurring is also increased when a data transfer network connects many observatories. A shortage of human resources to maintain the system will be a further difficulty to face. For these reasons, an integrated management system of global multipoint observations has been developed.



**Figure 1.** A schematic picture of the WONM.

This paper thus presents an integrated system—the Wide-area Observation Network Monitoring (WONM) system—to acquire, transfer, and store worldwide Space and Earth environment



observation data.

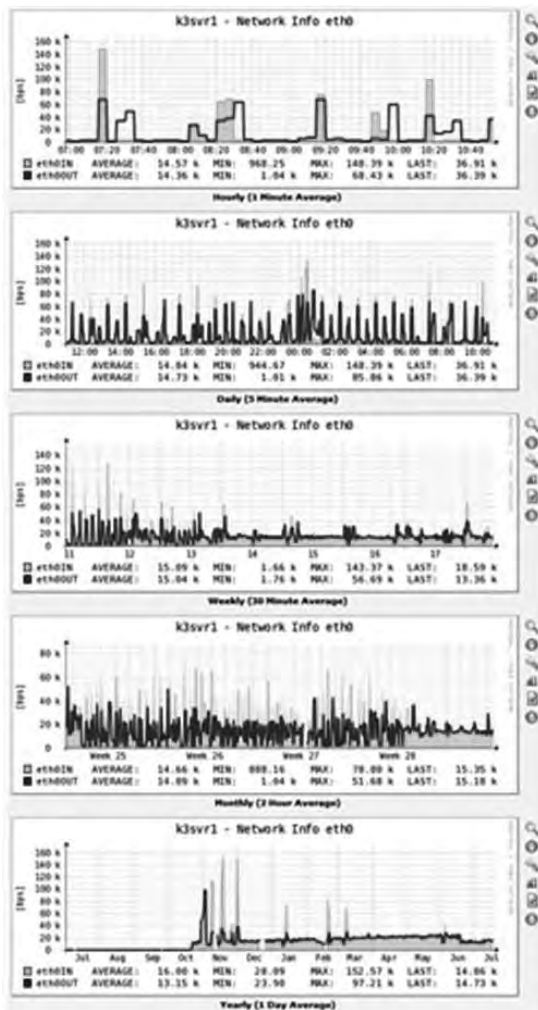


Figure 2. An example of WONM status monitoring (dump image): network traffic monitoring

## 2. WONM system

Figure 1 shows a schematic of the WONM system. A small, low-power personal computer (PC) equipped with the WONM software and an uninterruptible power supply (UPS) is placed at each observatory. Once the PC is set-up, observation data file transfer to the NICT Science Cloud commences. Status/housekeeping information on the PC, network, UPS, and so on, required for data transfer are collected by the system and also transferred to NICT. This status information is monitored by an operator of the system via a website (see Figure 2). The system has already been set-up at eight observatories, including the Showa station in Antarctica (Figure 3).

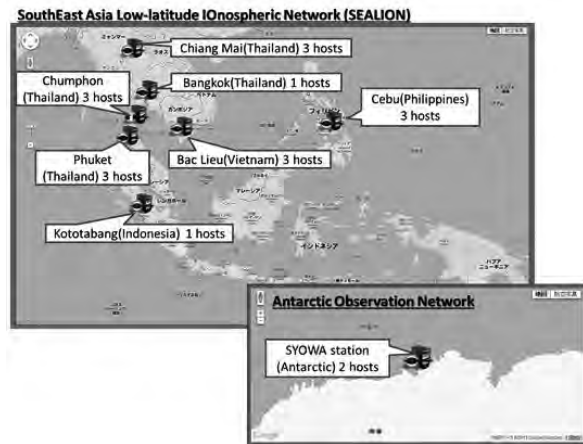


Figure 3. Observatories of WONM targets.

## 3. Conclusions

Operation of the WONM system has proved that it can be made available and is practical for users worldwide (Figure 3). Since the basic design of the system is not limited for space weather, there are a variety of applications for the system. In many worldwide observation projects, difficulties are found concerning routine operation, data transfer, status monitoring, databases, data sharing, and data stewardship. By connecting the WONM system to the NICT Science Cloud, an integrated system of wide-area observation networks is achieved that can be easily managed as though all observatories were located in the same room.

**Acknowledgments.** The present work is based on the NICT Science Cloud project.

## References

1. Nagatsuma, T., New Ages of Operational Space Weather Forecast in Japan. *Space Weather*, 11, 207–210, 2013
2. Watanabe, H., Yamamoto, K., Tsugawa, T., Nagatsuma, T., Watari, S., Murayama, Y., Murata, K. T., An Integrated Management System of Multipoint Space Weather Observation. *Data Science Journal*, 12, WDS175–WDS178, 2013
3. Murata, K. T., Watari, S., Nagatsuma, T., Kunitake, M., Watanabe, H., Yamamoto, K., Kubota, Y., Kato, H., Tsugawa, T., Ukawa, K., Muranaga, K., Kimura, E., Tatebe, O., Fukazawa, K., Murayama, Y., A Science Cloud for Data Intensive Sciences. *Data Science Journal*, 12, WDS139–WDS146, 2013

# IUGONET Data Analysis Software (UDAS) for Upper Atmosphere Study

**Yoshimasa Tanaka<sup>1\*</sup>, Atsuki Shinbori<sup>2</sup>, Tomoaki Hori<sup>3</sup>, Yukinobu Koyama<sup>4</sup>, Shuji Abe<sup>5</sup>, Norio Umemura<sup>3</sup>, Yuka Sato<sup>1</sup>, Manabu Yagi<sup>6</sup>, Satoru UeNo<sup>7</sup>, Akiyo Yatagai<sup>2</sup>**

<sup>1\*</sup> National Institute of Polar Research, 10-3 Midori-cho, Tachikawa-shi, Tokyo 190-8518, Japan

<sup>2</sup> Research Institute for Sustainable Humanosphere, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

<sup>3</sup> Solar-Terrestrial Environment Laboratory, Nagoya University, Furou-cho, Chikusa, Nagoya, Aichi 464-8614, Japan

<sup>4</sup> Data Analysis Center for Geomagnetism and Space Magnetism, Graduate School of Science, Kyoto University, Kitashirakawa-Oiwake Cho, Sakyo-ku, Kyoto 606-8502, Japan

<sup>5</sup> International Center for Space Weather Science and Education, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

<sup>6</sup> Planetary Plasma and Atmospheric Research Center, Tohoku University, Aramaki-aza-aoba, Aoba, Sendai 980-8578, Japan

<sup>7</sup> Kwasan and Hida Observatories, Graduate School of Science, Kyoto University, Kurabashira, Kamitakara-cho, Takayama, Gifu 506-1314, Japan

Email: ytanaka@nipr.ac.jp

**Summary.** Data analysis software developed by the Inter-university Upper atmosphere Global Observation Network (IUGONET) project is introduced. Named iUgonet Data Analysis Software (UDAS), this software has been created to promote interdisciplinary studies on the complex atmosphere of the Earth. It is written in Interactive Data Language and is provided as a plug-in software for the THEMIS Data Analysis Software suite. UDAS enables users to easily download, visualize, and analyze various upper atmospheric data distributed across members of IUGONET and having different file formats. Example plots of upper atmosphere data are presented to verify the effectiveness of the analysis software for polar atmospheric research.

**Keywords.** Analysis software, upper atmosphere, ground-based observation, Inter-university Upper atmosphere Global Observation Network, interdisciplinary study.

## 1. Introduction

The Inter-university Upper atmosphere Global Observation Network (IUGONET) is a Japanese inter-university collaborative project consisting of five Japanese institutes/universities: Tohoku University, Nagoya University, Kyoto University, Kyushu University, and the National Institute of Polar Research [1]. For decades, these institutes/universities have been leading ground-based, global network observations of the upper atmosphere and maintaining their own databases.

The main scientific target of this project is to comprehensively understand the long-term variation in the upper atmosphere. To achieve that target, it is necessary to investigate the relationships among data measured worldwide

by various types of instruments (e.g., magnetometers, atmospheric and ionospheric radars, lidars, imagers, solar telescopes, planetary radio telescopes). Data analysis software called iUgonet Data Analysis Software (UDAS) [2] has thus been developed that enables users to plot and analyze many data types on the same platform.

## 2. Overview of UDAS

UDAS is written in Interactive Data Language (IDL) [3], which is widely used in the solar-terrestrial physics community. This language is based on the THEMIS Data Analysis Software suite (TDAS) [4]; an integrated analysis platform that was built to visualize and analyze satellite- and

ground-based data obtained during the THEMIS mission [5]. The remarkable characteristics of TDAS are summarized in Subsection 2.1.

UDAS is provided on the IUGONET website [6] as a plug-in software for TDAS—such that the advantages of TDAS can be utilized—and includes routines to load various types of data provided by IUGONET members. In addition to these loading routines, other new routines developed by IUGONET have been added to UDAS (see Subsection 2.2). Since TDAS has also been adopted as the primary data analysis platform for the Japanese satellite mission Energization and Radiation in Geospace (ERG) [7], UDAS routines have been developed in collaboration with the ERG Science Center.

### 2.1 Characteristics of TDAS

- TDAS consists of open-source IDL codes that users can modify freely.
- It can automatically download data files from remote web servers through the Internet, regardless of the location and file formats.
- It enables users to appropriately visualize unfamiliar data by using only a few basic commands.
- It includes many useful routines for time series analysis, namely, digital filters, Fourier Transforms, and Wavelet Transforms.
- A Graphical User Interface is available for IDL and TDAS novices.

### 2.2 New developments by IUGONET

- IDL routines have been developed to retrieve metadata from the IUGONET metadata database [1].
- IDL routines for statistical analysis have been developed to objectively judge analysis results, for example, trends of time-series data and cross-correlation between two different datasets.
- Three-dimensional simulation data are archived in Common Data Format such that they are easily handled by TDAS.
- TDAS has been successfully compiled to run in the IDL Virtual Machine environment, which is a freely-distributed cross-platform

utility for running complied IDL code without a paid license.

## 3. Conclusions

IUGONET has developed the data analysis software, UDAS, to visualize and analyze various ground-based observational data of the upper atmosphere. UDAS has been released as a plug-in software for TDAS at the IUGONET website, and is expected to promote interdisciplinary collaborative research on the upper atmosphere.

**Acknowledgments.** The IUGONET project was initially supported by the Special Budget (Project) for FY2010 and, more recently, by the Ministry of Education, Culture, Sports, Science and Technology, Japan. We would like to acknowledge the cooperation and generosity of the THEMIS Science Support Team for allowing us to use TDAS for our data analysis software and for assisting in our development of the IUGONET plug-in tool.

## References

1. Hayashi, H. et al., Inter-university Upper Atmosphere Global Observation Network (IUGONET). *Data Sci. J.*, 12, WDS179–184, 2013
2. Tanaka, Y. et al., Development of IUGONET data analysis software. *J. Space Sci. Info. Jpn.*, 91–98, 2012 (in Japanese)
3. Exelis Visual Information Solutions, <http://www.exelisvis.com/ProductsServices/IDL.aspx> [accessed on: September 2013]
4. THEMIS, <http://themis.ssl.berkeley.edu/software.shtml> [accessed on: September 2013]
5. Angelopoulos, V., The THEMIS mission. *Space Sci. Rev.*, 141, 5–34, 2008
6. Inter-university Upper atmosphere Global Observation Network, <http://www.iugonet.org/en/software.html> [accessed on: September 2013]
7. Miyoshi, Y. et al., The Energization and Radiation in Geospace (ERG) Project. In: Summers, D., Mann, I. R., Baker, D. N., and Schulz, M. (Eds), *Dynamics of the Earth's Radiation Belts and Inner Magnetosphere*, *Geophys. Monogr. Ser. 199*, AGU, Washington D.C., 103–116, 2012

# Development of a Sensor Observation Service (SOS) Javascript Library

**Alex J. Tate<sup>1\*</sup>, Paul M. Breen<sup>1</sup>**

<sup>1\*</sup> *British Antarctic Survey, High Cross, Madingley Road, Cambridge, CB3 0ET, UK*  
Email: ajtate@bas.ac.uk

**Summary.** This paper discusses the development of a client-side Javascript library to provide an Application Programming Interface to the Sensor Observation Service (SOS) of the Open Geospatial Consortium. The library is modular and layered; consisting of a Core module and a User Interface (UI) module. The Core module encapsulates key concepts of the SOS, such as the connection parameters, the service's Capabilities document, methods to access the service's Features of Interest, Offerings, and Observed Properties. It also provides various utility functions. The UI module provides a number of User Interface components, such as SOS.Plot and SOS.Table for presenting the data acquired from the service, and SOS.Menu and SOS.Map for enabling navigation of the available service Offerings.

**Keywords.** Open Geospatial Consortium, Sensor Observation Service, Javascript, client-side.

## 1. Introduction

In 2012, a project was raised to reorganize British Antarctic Survey (BAS) meteorological data holdings. A component of this project was to provide open and interoperable access to those data via standards-based web services. This component is named SOSMet; the frontend library of which is described in this paper.

## 2. The SOSMet Project

The SOSMet Project consists of the backend web-service implementation, and a frontend web application.

### 2.1 Backend

The backend consists of the 52 North implementation of the Sensor Observation Service (SOS). This is a Tomcat web application written in Java, which sits on top of a Postgres database, with the Postgres Geographical Information Systems extensions.

### 2.2 Frontend

A discovery, browse, and access web application was then developed to interface with the above web service. This frontend web application was built by utilising the SOS library.

## 3. The SOS Library

The SOS library is modular and layered; consisting of a Core module, and a User Interface (UI) module. The Core module encapsulates key concepts of the SOS, such as the connection parameters, the service's Capabilities document, methods to access the service's Features of Interest, Offerings, and Observed Properties. It also provides various utility functions. The UI module provides a number of User Interface components, such as SOS.Plot and SOS.Table for presenting the data acquired from the service, and SOS.Menu and SOS.Map for enabling navigation of the available service Offerings.

The library is implemented as purely client-side Javascript. It is built on top of

- OpenLayers—for low-level SOS request/response handling, as well as some simple mapping for discovery;
- jQuery—for the UI and general application plumbing; and
- flot—which is a jQuery plugin, for the plotting.

The UI components can be used standalone (e.g., a SOS.Plot can be embedded within a web page), or they can be combined into a web

application. Indeed, there is a `SOS.App` component within the UI module, which pulls together all UI components into a general-purpose SOS web application.

#### **4. Conclusions**

Although born out of a specific requirement to provide access to BAS meteorological data, the

SOS library as it stands is generic; that is, it 'speaks' SOS, rather than BAS meteorology. This is the first phase of developing the frontend for BAS meteorological data; a general purpose SOS library. As a result of its general purpose nature, this library has far broader applications than that for which it was originally developed.

# Outline of Arctic Data Archive System (ADS)

**Hironori Yabuki<sup>1,2\*</sup>, Takeshi Sugimura<sup>2</sup>**

<sup>1\*</sup> National Institute of Polar Research, 10-3 Midoricho, Tachikawa, Tokyo, 109-8518, Japan

<sup>2\*</sup> Japan Agency for Marine-Earth Science and Technology, 2-15 Natsushima-cho, Yokosuka, Kanagawa, 237-0061, Japan

Email: yabuki.hironori@nipr.ac.jp

**Summary.** Arctic research by Japanese researchers has been carried out since the last century. The results of their research include many irreplaceable data, such as observational time series, field samples, and their analyses. Since researchers and organizations have managed these data in an ad hoc way, many data have not been adequately managed and systematically stored. Recently, however, managed information systems have proliferated in various branches of earth science. One such a system is the developing Arctic Data Archive System which is an open system for the registration of observation data and its metadata covering the Arctic region.

**Keywords.** Arctic, data, metadata, observation, Arctic Data Archive System.

## 1. Introduction

In the Arctic region, various changes are being caused by global warming. In previous studies, decreases of Arctic sea ice extent, increases in soil temperature in the region of Siberia, permafrost melting, increases in Arctic river runoff, and reductions in snow cover are all evidence of such change. The impact of human activities on ecosystems due to these changes is of concern. The actual condition and mechanism of environmental change in the arctic has not been elucidated. Previous studies have been carried out that focus separately on the atmosphere, oceans, and land. However, the Arctic is a system consisting of the interaction of the atmosphere, ocean, land surface, and snow and ice. Moreover, these systems encompass phenomena operating at different spatial and time scales. To clarify the variability of the Arctic environment, through interdisciplinary research, more attention is needed in using databases that integrate the results of observations across multiple areas at different scales.

Fortunately arctic research by Japanese researchers has been carried out since last century. The results of their research include many irreplaceable data, such as observational

time series, field samples, and their analyses. Since researchers and organizations have managed these data in an ad hoc way, many data have not been adequately managed and systematically stored.

For this reason the Arctic Data archive System (ADS) was initiated to collect, manage, and make accessible a greater number of arctic data. This project is supported by the Green Network of Excellence (GRENE): Rapid Change of the Arctic Climate System and its Global Influences.

## 2. Objectives

The objectives of ADS are as follows:

1. Archive of observational data in the Arctic region.
2. To promote the mutual use of data between observational researchers and modellers.
3. Promotion of the use of satellite data.

Modellers require field work to obtain necessary data—for example, local observation data—to validate their models. Conversely, observation researchers must put effort into work that picks out trends/patterns from big simulation data to make better sense of observed data values through modelling. ADS functions as a repository, and it aims to help integrate the work

of model and observation researchers.

The use of satellite data is essential for wide-area analysis of the Arctic region. Until now, satellite observation data—which are published by various agencies—were difficult to discover and were often complex products; requiring specialized knowledge for interpretation and analysis. Such interpretation was generally performed by experts. In consideration of the above problems, ADS aims to promote the use of satellite data by a wide range of researchers. To build systems that satisfy these needs, inputs from both modellers and observation researchers are required, and a data working group is being formed of participating researchers to facilitate this in the GRENE Arctic Project.

### 3. ADS

Data stored in ADS has searchable metadata. The metadata schema used is appropriate for many types of formats typical of Earth environment data, and additional schema are planned for other relational data types. To enter the metadata, an Excel spreadsheet has been adopted as a general-purpose input tool so that data can easily be added even when using the system offline. ADS has adopted a system that can find data visually using a web-interface that harnesses a Google Map (see Figure 1).

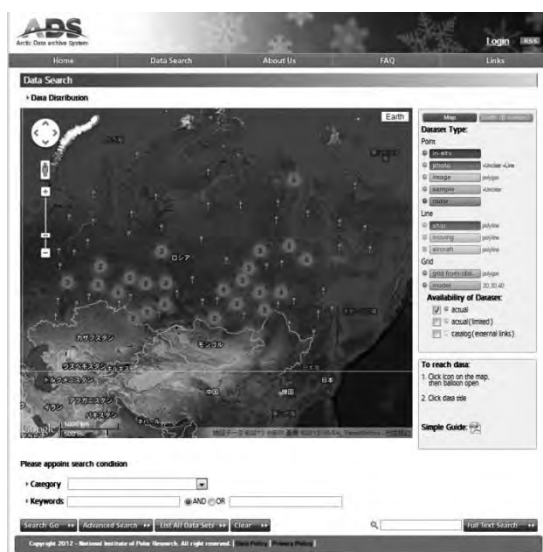


Figure 1. Screenshot of map search using ADS [1]

### 4. Conclusions

To promote the use of observational data, ADS has implemented a data policy and guidelines that clearly define the data ownership of observational researchers. In the future, we plan to implement capacity building for data providers, to promote the creation of metadata and dataset construction. It is planned to carry out not only archiving of data that are captured during the project's period but also collection and archival of historical observation data obtained in the Arctic. ADS is also expected to cooperate with another system—the 'Data Integration and Analysis System'—which will be developed to operate nationally within Japan. Furthermore, ADS will need the capability to accept data from international Arctic data providers.

**Acknowledgments.** This project is supported by Green Network of Excellence (GRENE): Rapid Change of the Arctic Climate System and its Global Influences.

### References

1. National Institute of Polar Research, <https://ads.nipr.ac.jp> [accessed on: September 2013]

# The ICSU World Data System: Trusted Data Services for Global Science

**Rorie Edmunds**<sup>1\*</sup>

<sup>1\*</sup> WDS International Programme Office, c/o NICT, 4-2-1 Nukui-kitamachi, Koganei, Tokyo 184-8795, Japan  
Email: rorie.edmunds@icsu-wds.org

**Summary.** The International Council for Science's World Data System (ICSU-WDS) is striving to build a worldwide *community of excellence* for scientific data by certifying Member Organizations—holders and providers of data or data products—from wide-ranging fields using internationally recognized standards. These 'WDS Members' are the building blocks of searchable common directories and catalogues with which to form a data system that is both interoperable and distributed.

**Keywords.** Scientific data, data centre, data service, accreditation, long-term preservation.

## 1. Introduction

Operating under the auspices of the International Council for Science (ICSU), the World Data System (WDS) is an interdisciplinary global federated system of long-term data archives and data services [1]. Roles in the system encompass the entire data management chain, from data production to publication, and WDS members can hold a number of these roles. Emphasis is also placed on linking with data production facilities, in particular, science projects and programmes.

## 2. History

The International Geophysical Year in 1957–1958 saw the creation by ICSU of the World Data Centres (WDCs) and Federation of Astronomical and Geophysical data analysis Services (FAGS). Charged with managing the data gathered by this global project and making those available to the scientific community, both the WDCs and FAGS remained largely unchanged for 50 years.

It was another major international scientific initiative, the International Polar Year (IPY; 2007–2008), that saw the disbanding of the WDCs and FAGS in 2009 by the ICSU General Assembly. The expectation was that the existing infrastructure could house the diverse IPY data; however, it quickly became clear that the predominantly

disciplinary centres and services were not able to respond to modern data needs.

## 3. The WDS Concept

ICSU-WDS was thus borne from the legacy of the WDCs and FAGS with the aim of transitioning from stand-alone components to a common globally coordinated and distributed data system. It ensures universal and equitable access to, and long-term stewardship of, *quality-assured scientific data and data services, products, and information* covering a broad range of natural sciences and expanding to social sciences. In doing so, WDS supports ICSU's mission of 'strengthening international science for the benefit of society' [2].

### 3.1 Data Policy

To contribute to international data-sharing efforts, ICSU-WDS has adopted the principles of the Global Earth Observation System of Systems [3]. Data, metadata, and products shared within WDS will be

- Exchanged fully and openly, recognizing relevant international instruments and national policies and legislation.
- Made available with minimum time delay and at minimum cost.
- Free of charge or at reproduction cost for research and education.



### 3.2 Governance

ICSU-WDS is governed by a Scientific Committee (WDS-SC) composed of prominent scientists actively involved in Data/Computer Sciences, and dealing with large dataset issues and long-term data stewardship. The WDS-SC includes directors of WDS Member Organizations and covers a broad range of disciplines and geographic areas.

The WDS-SC is supported by an International Programme Office (WDS-IPO), which under its supervision and guidance is responsible for the daily implementation of WDS-SC decisions, WDS activities, and WDS communication functions.

## 4. WDS Members: Constructing the System

WDS Members from wide-ranging fields are the fundamental components of a worldwide *community of excellence* for scientific data.

As of 14 August 2013, ICSU-WDS has 75 Member Organizations grouped into two types:

- 53 Regular and 7 Network Members, who hold, serve, or produce data; and include virtual data centres.
- 2 Partner and 13 Associate Members, who provide various backing to ICSU- WDS or are simply interested in the endeavour.

These organizations primarily participate in ICSU-WDS by contributing their data holdings, services, and products; which are the cornerstone of the federated data system. They also provide essential contributions to WDS Working Groups in which they have an interest, and are fostered to actively participate in common projects and activities that advance WDS aims.

## 5. Member Accreditation

Organizations voluntarily submit an Expression of Interest to join ICSU-WDS under one of the four WDS membership categories: Regular, Network, Partner, or Associate. Such candidate organizations are then formally accepted as WDS Members by the WDS-SC, who safeguard the trustworthiness of ICSU-WDS by certifying Regular and Network Members according to

criteria on scientific relevance, governance, data management, technical infrastructure, and security (Partners and Associates are co-opted).

Hence, when Regular and Network Members join ICSU-WDS, they become part of a community acknowledged to have met an accreditation based on internationally recognized standards.

In fact, Network Members—organizations representing groups of data centres/services—play an additional role in the WDS accreditation process, in that they are expected to promote long-term data stewardship amongst their network nodes. This can be achieved by either:

1. Encouraging data centres/services under their umbrella to become accredited WDS Regular Members.
2. Establishing a certification procedure aligned with that used for WDS Regular Members (e.g., the International Oceanographic Data and Information Exchange's Quality Management Framework [4]).

## 6. Conclusions

Even though the comprehensive accreditation procedure prescribed by ICSU-WDS is unique in the data management landscape, it is not considered to be definitive. ICSU-WDS is currently working with the Data Seal of Approval [5] under the context of a Research Data Alliance Working Group [6] in an attempt to implement a common certification service, and thus forge a global network of trusted data providers.

## References

1. ICSU World Data System, <http://www.icsu-wds.org> [accessed on: August 2013]
2. International Council for Science, <http://www.icsu.org> [accessed on: August 2013]
3. Group on Earth Observations, [http://www.earthobservations.org/geoss\\_dsp.shtml](http://www.earthobservations.org/geoss_dsp.shtml) [accessed on: August 2013]
4. International Oceanographic Data and Information Exchange, [http://www.iode.org/index.php?option=com\\_oe&task=viewDocumentRecord&docID=9958](http://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&docID=9958) [accessed on: August 2013]

5. Data Seal of Approval, <http://www.datasealofapproval.org/> [accessed on: August 2013]
6. Research Data Alliance, <https://rd-alliance.org/node> [access on: August 2013]



# Organizing Committees

## Scientific Organizing Committee

DANIS, Bruno – SCAR/SC-ADM/Royal Belgian Institute of Natural Sciences  
de BRUIN, Taco – SCAR/SC-ADM/Royal Netherlands Institute for Sea Research  
\*FINNEY, Kim – SCAR/SC-ADM/WDS-SC/Australia Antarctic Division  
HIK, David – President, IASC/University of Alberta  
KANAOK, Masaki – SCAR/SC-ADM/Polar Data Centre (PDC), National Institute of Polar Research (NIPR)  
KOPYLOV, Vasilii – WDS-SC/Director, All-Russian Research Institute of Hydrometeorological Information  
LARSEN, Jan René – Secretary, SAON/Arctic Monitoring and Assessment Program  
MINSTER, Jean-Bernard – Chair, WDS-SC/Scripps Institution of Oceanography, University of California, San Diego  
PARSONS, Mark – IPY-Data/Managing Director, Rensselaer Polytechnic Institute  
WATANABE, Takashi – Senior Advisor, WDS-IPO/Director, WDC–Cosmic Rays

## Local Advisory Committee

ENOMOTO, Hiroyuki – Director, Arctic Research Centre (ARC), NIPR  
\*IYEMORI, Toshihiko – Director, WDS for Geomagnetism, Kyoto; Kyoto University  
KITAMOTO, Asanobu – National Institute of Informatics  
NAKAJIMA, Hideaki – National Institute for Environmental Studies  
NAKAMURA, Takuji – Vice Director-General, NIPR  
NISHITANI, Nozomu – Solar-Terrestrial Environment Laboratory, Nagoya University  
OHATA, Tetsuo – National Representative, SAON/Japan Agency for Marine-Earth Science and Technology (JAMSTEC)  
OHISHI, Masatoshi – National Astronomical Observatory  
SHINOHARA, Iku – Japan Aerospace Exploration Agency  
SHIRAISHI, Kazuyuki – Chair, SCAR National Committee/Director-General, NIPR  
SUGIMOTO, Atsuko – Chair, IASC National Committee/Hokkaido University  
TSUBOI, Seiji – Director, JAMSTEC Data Centre  
WATANABE, Kentaro – Head of International Affairs Section, NIPR  
YAMANOUCHI, Takashi – Vice Director-General, NIPR  
YOKOYAMA, Kazumi – National Museum of Nature and Science (NMNS)

## Local Organizing Committee

EDMUNDS, Rorie – Programme Officer, WDS-IPO  
IIDA, Takahiro – Southern Ocean Observing System (SOOS) Data Committee/NIPR  
ISHII, Mamoru – Director, WDC–Ionosphere and Space Weather/NICT  
KADOKURA, Akira – Director, PDC, NIPR  
\*KANAOK, Masaki – SCAR/SC-ADM/PDC, NIPR  
MOKRANE, Mustapha – Executive Director, WDS-IPO  
MURAYAMA, Yasuhiro – NICT  
OKADA, Masaki – PDC, NIPR  
WATANABE, Takashi – Senior Advisor, WDS-IPO/Director, WDC–Cosmic Rays  
YABUKI, Hironori – SAON Data Committee/ARC, NIPR /JAMSTEC

---

\* Committee Chair

# Hosting, Sponsoring, and Supporting Organizations

## Host Organizations

SCAR's Standing Committee on Antarctic Data Management

WDS Scientific Committee

## Sponsor Organizations

National Institute of Polar Research

National Institute of Information and Communications Technology

National Museum of Nature and Science

International Programme Office of World Data System

Society of Geomagnetism and Earth, Planetary and Space Sciences

## Support Organizations

International Council for Science

Scientific Committee on Antarctic Research (SCAR)

International Arctic Scientific Committee

ICSU World Data System (WDS)

ICSU Committee on Data for Science and Technology

Sustaining Arctic Observing Network

Science Council of Japan

Graduate School of Science, Kyoto University

Solar-Terrestrial Environment Laboratory, Nagoya University

