

METADATA MANAGEMENT AT THE POLAR DATA CENTRE OF THE NATIONAL INSTITUTE OF POLAR RESEARCH, JAPAN

M Kanao^{1}, M Okada¹, and A Kadokura¹*

¹*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*

Emails: {[okada.masaki](mailto:okada.masaki@nipr.ac.jp), [kadokura](mailto:kadokura@nipr.ac.jp)}@nipr.ac.jp

ABSTRACT

The Polar Data Centre of the National Institute of Polar Research has had the responsibility to manage the data for Japan as a National Antarctic Data Centre for the last two decades. During the International Polar Year (IPY) 2007–2008, a considerable number of multidisciplinary metadata that mainly came from IPY-endorsed projects involving Japanese activities were compiled by the data centre. Although long-term stewardship of those amalgamated metadata falls to the data centre, the efforts are in collaboration with the Global Change Master Directory, the Polar Information Commons, and the newly established World Data System of the International Council for Science.

Keywords: International Polar Year, Polar Data Centre, Metadata management, Global Change Master Directory, World Data System

1 INTRODUCTION

The rapid technological development in Earth observations by both satellite- and ground-based networks in the polar region has led to a large number of observation data being collected every day. The processing and utilization of these data are important issues for the promotion of Polar Sciences. There have been several programs involving scientific data management and the provision of information infrastructure. The International Polar Year (IPY) 2007–2008 was the most diverse international science program in recent history. It was conducted during the 50th anniversary of the International Geophysical Year (IGY) 1957–1958. IPY 2007–2008 greatly enhanced the exchange of ideas across nations and scientific disciplines; unveiling the status of and changes to planet Earth as viewed from the polar regions. This interdisciplinary exchange has helped us to understand and address grand challenges, such as rapid environmental changes and their impact on society.

The Polar Data Centre (PDC) of the National Institute of Polar Research (NIPR) has served as 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). During IPY 2007–2008, PDC compiled many of the polar data emanating from projects involving Japanese activities (Sato, Ito, Kanao, Kanda, Naganuma, Ohata, et al., 2011). In this paper, the current status of metadata management in Japan, particularly that concerned with the tasks of PDC, is demonstrated. A tight linkage has been put in place with other scientific bodies of ICSU, such as the Committee on Data for Science and Technology (CODATA) and the new World Data System (WDS).

2 POLAR DATA CENTRE

At the 22nd Antarctic Treaty Consultative Meeting (ATCM) held in 1998, affiliated countries were obliged to ensure that scientific data collected by Antarctic programs can be freely exchanged and utilized. Following the Articles of the 1998 Antarctic Treaty, each country was required to establish a National Arctic Data Centre (NADC) within which scientists are expected to submit collected data appropriately. The PDC at NIPR has been performing the NADC function in Japan, and established a data policy in February 2007 based on the requirements of the Standing Committee on Antarctic Data Management (SC-ADM) of SCAR. This contributed to the establishment of the subsequent SCAR Data and Information Management Strategy (SCAR-DIMS; Finney, 2009; de Bruin & Finney, 2011).

Regarding auroral data, in particular, PDC has administered the World Data Centre (WDC) for Aurora since 1981, which is responsible for data archiving and dissemination of observational data relating to auroral activities—all-sky camera observations images, spectroscopic observations, satellite observations (auroral

images and energetic particle fluxes), geomagnetic observations, and observations of upper atmosphere phenomena.

PDC is also responsible for archiving and analysis of Earth observing satellite data (Polar Operational Environmental Satellite of the National Oceanographic and Atmospheric Administration), seismological data, and geodetic data in the locality of Syowa Station (SYO, 69S, 30E), Antarctica. In addition, PDC manages various information infrastructures such as: a mainframe computer and workstations, network systems of domestic and Antarctic facilities, and Earth observing satellite facilities.

3 METADATA MANAGEMENT

The principal task of PDC is to archive and make accessible digital data obtained from the polar regions. Summary information of all archived data (metadata) is available to the polar science community as well as data users having an interest in polar phenomena. The compiled metadata span a wide variety of science disciplines related to polar research (space and upper-atmospheric sciences, meteorology and glaciology, geoscience, and bioscience) from both long- and short-term research projects performed in the Arctic and Antarctic, particularly data collected by the Japanese Antarctic Research Expedition (Kanao, Kadokura, Yamanouchi, & Shiraishi, 2008; Kanao, Kadokura, Okada, Yamnouchi, Shiraishi, Sato, et al., 2013). As of June 2013, a total of 255 records had been archived in the amalgamated meta-database provided by PDC, including metadata from IPY-endorsed projects (<http://scibase.nipr.ac.jp/>; Figures 1(a), (b), and (c)).

NIPR Metadata Portal

Top Page

<http://scibase.nipr.ac.jp/>

NIPR_MetaDB; N=255

English and Japanese pages

The screenshot shows the NIPR Metadata Portal interface. The main content area displays a 'meta data list' with the following entries:

Date	Time	Title
2012/6/1	11:27	Ocean color remote sensing data
2012/6/1	11:23	Polar plant biodiversity database
2012/6/1	10:54	Database of animal specimens from polar regions
No image	2012/6/1 10:51	Blendon Samolins on Board the Squid/Fuji/Squid in 1992-1992
No image	2012/6/1 10:41	Element Content and Salinity of Sea Ice Cores collected near Syowa Station, Antarctica in 1992/92
No image	2012/6/1 10:38	Chlorophyll concentrations and environmental parameters recorded with a moored buoy system
No image	2012/9/30 16:02	Chlorophyll concentrations and environmental parameters recorded with a moored buoy system
No image	2012/9/30 15:57	Oceanographic and Marine Biological data near Syowa Station
No image	2012/9/30 15:51	Whale sighting efficiency of the crew on board ocean research vessels in SCOSAS/STBS
No image	2012/9/30 15:10	Zooplankton Data at Syowa Station

The 'Hot News' section includes:

- User Manual was renewed. (2nd version) (2012-04-18 10:30:51)
- Dr. Mytasha Mokrane, Science and Information Technology of ICBU visited NIPR. (2012-02-09 18:11:38)
- 2012 JASUJ Union Session (U-01) "Toward a New Framework of Global Data Activity" (2012-01-27 15:29:01)
- "The Access Checker" began at the metaDB portal. (2011-12-02 16:01:29)
- IPY Data Management page was created (2011-11-21 12:07:17)
- Standing Committee on Antarctic Data Management meeting (SCADM-15) (2011-11-07 17:37:39)
- The first ICBU World Conference - Global Data for Global Science - was held (2011-11-07 17:21:01)
- Introduction Lecture for "Metadata Portal Server" was held (2011-07-28 17:48:21)
- about user password (2011-07-01 09:26:28)
- The Polar Information Commons: PIC page was opened. (2011-06-29 14:54:09)

Figure 1(a). Top page of NIPR metadata portal (<http://scibase.nipr.ac.jp/>)

A new content management system enabling access to the metadata has been in place since April 2011. The index page can be switched instantaneously from English to Japanese so that it can be utilized by both international and domestic users. There are several sophisticated utilities for users such as a data search engine and a data input page for adding new metadata, which cover five major scientific disciplines. At the time of writing this article (November 2013), an increase in the number of scientific branches is planned (e.g., Project, Monitoring, IPY), so as to match the increase in polar projects with NIPR involvement. Moreover, to ensure interoperability between the NIPR database and metadata portals operated by other polar communities and countries, a database system using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH; <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>) should be developed in the near future.

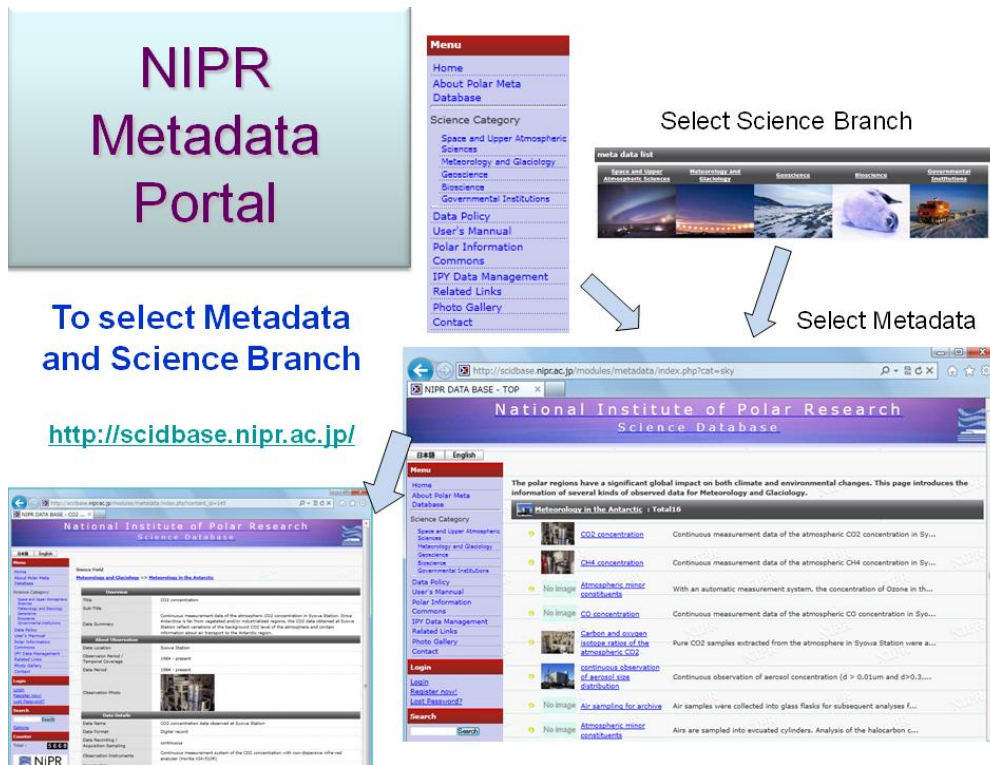


Figure 1(b). Flowchart for selecting data from the NIPR metadata portal

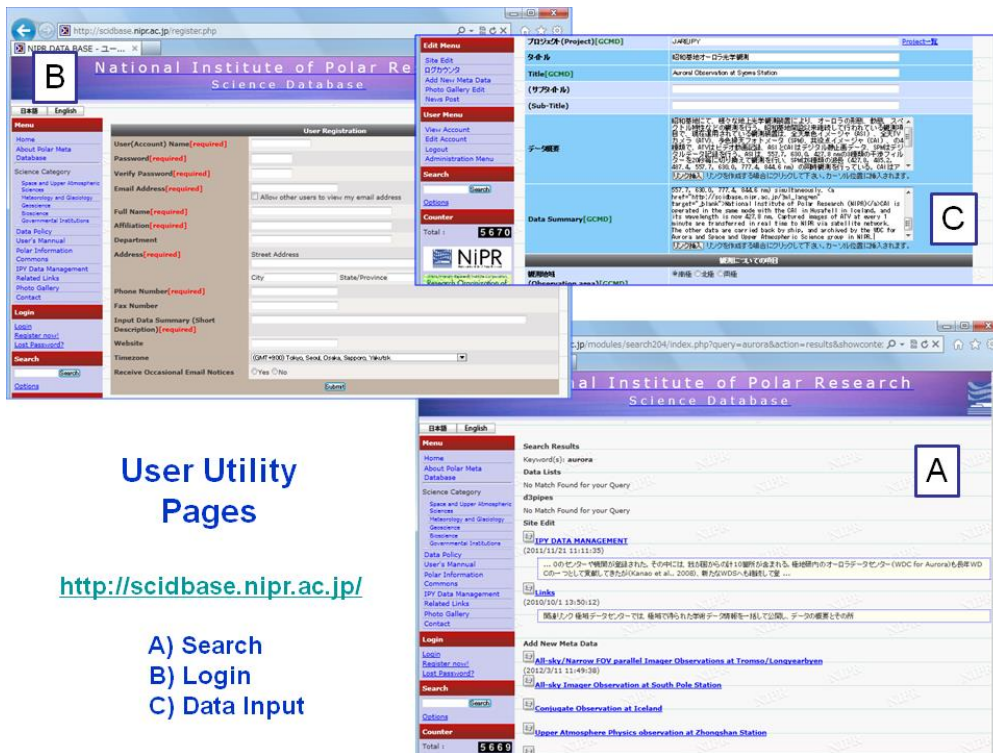


Figure 1(c). Flowchart for user utility pages of NIPR metadata portal

The polar database provided by PDC is also linked with those held by the Antarctic and Arctic Master Directories (AMDs) in the Global Change Master Directory (GCMD) of the National Aeronautics and Space Administration (NASA). In addition to IPY data, those from other national and international projects have been compiled, and 279 metadata records have been amalgamated (June 2013) in the Japanese Antarctic portal of

GCMD (http://gcmd.nasa.gov/KeywordSearch/Home.do?Portal=amd_jp&MetadataType=0; Figure 2). Hence, although PDC stores all metadata in their original format, the main items listed in the GCMD Directory Interchange Format (DIF) are also included, and metadata in both the AMDs and the PDC meta-database are closely linked.



Figure 2. Japanese Antarctic portal (AMD_JP) in GCMD

A total of 250 metadata records collected by Japanese IPY projects have also been compiled in an IPY portal within GCMD (<http://gcmd.gsfc.nasa.gov/KeywordSearch/Home.do?Portal=ipy&MetadataType=0>). These metadata constitute a significant portion of all IPY metadata held by GCMD. Although many scientific outcomes of IPY 2007–2008 have already emerged, deep understanding of polar phenomena will require creative use of the myriad data collected by the various scientific disciplines. The vast number of data accumulating during and after IPY 2007–2008 will be its most important legacy but only if they are well preserved and utilized (Parsons, de Bruin, Tomlinson, Campbell, Godoy, LeClert, et al., 2011).

4 INTERNATIONAL COLLABORATION

To construct an effective long-term strategy of polar (meta) data management, datasets must be made available promptly, and new Internet technologies must be employed for such a repository network service. The future plan should be to identify relevant developments, new directions, and emerging technologies within specific disciplines, and these should then be promoted among the polar data communities.

SCAR's SC-ADM under SCAR has been heavily involved in IPY data-management activities (e.g., the IPY Data and Information Service; IPY-DIS). The conclusion of IPY 2007–2008 saw the commencement of the Polar Information Commons (PIC), a new framework for long-term stewardship and provision of polar data and information. PIC's mandate is to (1) serve the polar community as an open, virtual repository for vital scientific data and information and (2) provide a shared, community-based cyber-infrastructure for fostering innovation and improved scientific understanding as well as encouraging participation in research, education, planning, and management in the polar regions.

PIC has developed specialized tools that produce small, machine-readable 'badges' that can be attached to metadata or data. These badges assert that data are openly available and enable generic search engines or customized portals to automatically identify and locate relevant data. This service is coupled with a cloud-based data repository for those data that may not have a suitable archive elsewhere. NIPR and other Japanese organizations have made considerable contributions to PIC through the attaching of data/metadata badges and the registration of datasets in the cloud-based repository (15 as of October 2013).

Through a decision of the 29th General Assembly of ICSU in 2008, a new World Data System (ICSU-WDS) was established based on the 50-year legacy of two ICSU science bodies—the World Data Centres (WDCs) and the Federation of Astronomical and Geophysical Data Analysis Services. ICSU-WDS aims at a transition from existing standalone WDCs and individual services to a common, globally interoperable, distributed data system that incorporates emerging technologies and new scientific data activities, including polar data as a legacy of IPY. The new system will build on the potential offered by advanced interconnections between data-management components for disciplinary and multidisciplinary applications. ICSU-WDS has also agreed to take the necessary steps to archive polar data in order to preserve the legacy of IPY 2007–2008.

5 CONCLUSION

The status of metadata management in PDC of NIPR has been summarized in this report. Many dedicated data-service tasks have been conducted by the staff of PDC as a member NADC of SCAR. Scientific data collected in the polar region have showed already their great significance for global environmental research in this century. To construct an effective long-term strategy for polar data management, data must be made available promptly, and new Internet technologies such as a repository network service similar to PIC must be employed. Moreover, interoperability between metadata portals can be promoted via a system using the OAI-PMH protocol that will be developed in the near future. Alongside the data activities of SCAR and IASC polar communities, tighter linkages should also be established with other multi-disciplinary science bodies under ICSU, such as CODATA and WDS.

6 ACKNOWLEDGEMENTS

The authors would like to express their gratitude to their many collaborators associated with IPY activities, both in national and international projects. They also acknowledge the members of SC-ADM, the IPY Data committee under the IPY-JC, for their great efforts in advancing polar data management. We also extend our appreciation to the committed individuals of WDS and CODATA whose fruitful discussions and arrangements initiated PIC, as well as helping to formulate the new data strategy of ICSU. Finally, we thank Prof. T. Watanabe and members of the Scientific Organizing Committee for organizing this special issue in the CODATA Data Science Journal, as an outcome of the International Forum on ‘Polar Data Activities in Global Data Systems’; held in Tokyo in October 2013.

7 REFERENCES

- de Bruin, T. & Finney, K. (2011) The SCAR Data Policy. *SCAR Newsletter* 27, pp 3.
- Finney, K. (2009) SCAR Data and Information Management Strategy (DIMS) 2009–2013. In Summerhayes, C. & Kennicutt, C. (Eds.), *SCAR Ad-hoc Group on Data Management Report 34*, Cambridge: Scott Polar Research Institute.
- Kanao, M., Kadokura, A., Yamanouchi, T., & Shiraishi, K. (2008) The Japanese National Antarctic Data Centre and the Japanese Science Database. *JCADM Newsletter* 1, pp 10.
- Kanao, M., Kadokura, A., Okada, M., Yamnouchi, T., Shiraishi, K., Sato, N., et al. (2013) The State of IPY Data Management: The Japanese Contribution and Legacy. *CODATA Data Science Journal* 12, pp WDS124–WDS128.
- Parsons, M.A., de Bruin, T., Tomlinson, S., Campbell, H., Godoy, Ø., LeClert, J., et al. (2011) The State of Polar Data—The IPY Experience. In Krupnik, I., Allison, I., Bell, R., Cutler, P., Hik, D., Lopez-Martinez, J., et al. (Eds.), *Understanding Earth's Polar Challenges: International Polar Year 2007–2008—Summary by the IPY Joint Committee 3.11*, Edmonton, Alberta: Art Design Printing Inc, pp 457–476.
- Sato, N., Ito, H., Kanao, M., Kanda, H., Naganuma, T., Ohata, T., et al. (2011) Engaging Asian Nations in IPY: Asian Forum for Polar Sciences (AFoPS) (Japanese Section). In Krupnik, I., Allison, I., Bell, R., Cutler, P., Hik, D., Lopez-Martinez, J., et al. (Eds.), *Understanding Earth's Polar Challenges: International Polar Year 2007–2008—Summary by the IPY Joint Committee 5.3*, Edmonton, Alberta: Art Design Printing Inc, pp 555–574.

(Article history: Available online 30 September 2014)