



A review of data management for polar sciences in Japan

Masaki Kanao

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

October 24, 2019

A review of data management for polar sciences in Japan

Masaki Kanao^{1*}

^{1*} *Polar Environment Data Science Center, Joint Support-Center for Data Science Research,
Research Organization of Information and Systems,
10-3, Midori-cho, Tachikawa-shi, Tokyo 190-8518, Japan
Email: kanao@nipr.ac.jp*

Summary. Diverse data accumulated by many science disciplines in polar regions make up the most significant legacy of the International Polar Year (IPY2007-2008). The Polar Data Center (PDC) of the National Institute of Polar Research (NIPR), followed by the Polar Environment Data Science Center (PEDSC) of the Joint Support-Center for Data Science Research (DS) since 2016 have responsibility to manage these multi-disciplinary polar data in Japan as a National Antarctic Data Center (NADC). The data policy of PDC/PEDSC was established in February 2007, based on the requirements of the Standing Committee on Antarctic Data Management (SCADM) of the Scientific Committee on Antarctic Research (SCAR). During the IPY, a significant number of metadata records were compiled from IPY- endorsed projects. A tight collaboration has been established between the Global Change Master Directory (GCMD), the Polar Information Commons (PIC), and the newly established World Data System (WDS). In this presentation, a decade of history of polar data management is demonstrated, in particular focusing on several database of PEDSC, international collaboration among global data bodies and initiatives, data publication and citation, as well as data journal issues (Polar Data Journal). Linkages of data sharing among Asian Forum for Polar Sciences (AFoPS) countries, moreover, should be promoted by the involved countries in near future.

Keywords. Data Management, Data Publication, Database, Global Linkage, Polar Sciences

1. Introduction

At the 22nd Antarctic Treaty Consultative Meeting (ATCM) in 1998, affiliate countries were obliged to ensure that scientific data collected from Antarctic programs could be freely exchanged and used. Following Article No.III.1.c of the Antarctic Treaty, each country is required to establish NADC and to properly disclose the data collected from involved scientists. The PDC and PEDSC have performed the function of a NADC for Japan and established a data policy in February 2007, based on the requirements of the Standing Committee on Antarctic Data Management (SCADM) of the Scientific Committee on Antarctic Research (SCAR). This contributed to the subsequent SCAR Data and Information Management Strategy [1].

Dedicated data services have been conducted by PDC and PEDSC as a member of NADC under SCAR. Several different disciplines of

scientific data collected in polar region have great significance for global environmental research. To construct an effective framework for long-term strategy of the polar data, data must be made available promptly by Internet technologies such a repository network service. In addition to activities in polar science communities of SCAR and the International Arctic Science Committee (IASC), tighter linkages expected to be established with other cross-cutting science bodies under ICSU, such as CODATA, and WDS. Linkages among these data-management bodies need to be strengthened in the post IPY era.

2. Data Management and Database

The International Polar Year (IPY 2007-2008) was the world's most diverse science program. It was conducted during the 50th anniversary of the International Geophysical Year (IGY 1957-1958). The IPY greatly enhanced the exchange of ideas

across nations and scientific disciplines to unveil the status and changes of planet Earth as viewed from polar region. The interdisciplinary exchange helped us understand and address grand challenges such as rapid environmental change and its impact on human society. Eventually, Japanese researchers participated to 63 projects endorsed by the IPY Joint Committee. The huge amount of data accumulating during IPY should be the most important legacy if it is well preserved and utilized [2].

The science database provided by PDC and PEDSC has a tight connection with AMD in GCMD. In addition to the IPY-related data, data from Japanese national and other international projects had been compiled. In total, 300 metadata were compiled in Japanese Antarctic portal in GCMD. PDC and PEDSC store its metadata in their original format, but this includes the main items listed in GCMD Directory Interchange Format (DIF). There are tight cross-linkages in corresponding metadata in AMD. Metadata collected by IPY projects for Japan have also been accumulated in an IPY portal of GCMD. More than 250 metadata from Japan were stored in the IPY portal [3, 4]. This constitutes a significant proportion of all IPY metadata to GCMD.

3. Data Publication and Legacy

SCADM has been strongly connected with IPY data-management activities (IPY Data and Information Service: IPY-DIS). IPY data policy emphasized a need to make data available on the “shortest feasible timescale.” In accordance with IPY data policy, IPY-DIS recommended that data be formally cited when used, and the IPY Data Committee has developed initial guidelines for how data should be cited. The guidelines harmonized different approaches, and they adopted by many data centers relating polar area. After the end of IPY, a new initiative, the Polar Information Commons (PIC), began as a framework for open and long-term stewardship of polar data and information [2]. The PIC serves as an open, virtual repository for vital scientific data and information and provides a shared,

community-based cyber-infrastructure fostering innovation and improved scientific understanding while encouraging participation in research, education, planning, and management in polar region. PIC developed specialized tools that produce a small, machine-readable “badge” that is attached to the data. However, the badge requested data users to adhere to basic ethical norms of data use including proper citation. This service was coupled with a cloud-based repository that may not have a suitable archive elsewhere. NIPR/DS have made contributions to PIC, both by attaching badges and registration in the repository. As of 2019, Japan had contributed more than 50 data sets to the PIC.

Polar data have great relevance for modern, global environmental research well beyond the polar region. It is critical to explore the cloud approaches such as the PIC to develop an effective framework for open and long-term stewardship of polar data. The status of data-management before and after the IPY in Japan was introduced in this report. Several different aspects of the scientific data collected in the polar region have great advantage for global environmental research as well as in future. Linkages of data sharing among global community, in particular with Asian countries, should be promoted.

4. Global Linkage in Polar Sciences

The Asian Forum for Polar Sciences (AFoPS) is a non-governmental organization established in 2004 to encourage and facilitate cooperation for the advance of polar sciences among countries in the Asian region. The Forum consists of its six members, i.e., the national polar research institutions representing China, Japan, South Korea, India, Malaysia and Thailand. AFoPS also has four observer countries: Indonesia, Philippines, Sri Lanka and Vietnam, respectively.

The objectives of AFoPS has been recognized as the value of scientific research in bi-polar regions for the benefit of human activities, recognizing the importance of international cooperation in the polar regions and the need to work closely with other national operators, as well as aiming to serve the common interests in both polar sciences and logistics. Member countries will work together for the tasks as follows; provide a

foundation for cooperative research activities; present Asian achievements toward international polar communities; encourage more Asian countries' involvements in polar sciences.

Major activities of the AFoPS include; provide a forum to seek a common view on polar affairs among member countries; develop and support cooperative programs on polar activities (i.e., joint science projects, logistic cooperation and personnel exchange program between polar expeditions and institutes, etc.); convene joint symposia and workshops for sharing scientific results, information and experience joint symposium; conference activities within AFoPS working groups (WGs); support non member countries to develop their national polar programs; invite scientists to field expeditions and institutes; invite scientists to AFoPS meetings; provide personnel training and cooperation in outreach activities; produce joint publications on polar sciences.

Among data activities in polar communities of SCAR and IASC, moreover, tighter linkages of data sharing (data inter-operability) among AFoPS countries should be promoted in near future by the involved members of Asian countries. Establishing a new portal inside GCMD, for example, the simple and first step for data sharing and inter-operability.

Acknowledgments. The authors would like to express their appreciation to many collaborators involving polar data communities, in particular to the member of PDC and PEDSC, data committee members on SCAR, IASC, WDS, CODATA, AFoPS and IPY.

References

1. Finney, K., SCAR Data and Information Management Strategy 2009-2013. *SCAR Ad-hoc Group on Data Management*, 34, 2009
2. Parsons, M. A., Godoy, Ø., LeDrew, E., de Bruin, T., Danis, B., Tomlinson, S., Carlson, D., A Conceptual Framework for Managing Very Diverse Data for Complex, Interdisciplinary Science. *Journal of Information Science*, 1-21, 2011
3. 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, *Data Science Journal*, 12, WDS124-WDS128, 2013
4. Kanao, M., Okada, M., Friddell, J. and Kadokura, A., Science Metadata Management, Interoperability and Data Citations of the National Institute of Polar Research, Japan. *Data Science Journal*, 17, 1-6, 2018