APPLICATION AND METADATA FORMAT OF CRYOSPHERE DATA ARCHIVE PARTNERSHIP (CrDAP)

Hironori Yabuki [1]*

*[*1] Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokosuka, 237-0061, Japan*
Email: yabuki@jamstec.go.jp

ABSTRACT

An international data archive is critical for understanding the climate system dynamics of the cryosphere. Currently, no such system exists though data collection and integration efforts are ongoing. The Cryosphere Data Archive Partnership (CrDAP) is developing an open system for storing ctryospheric observation data and metadata. First stage data handling in CrDAP focused on integrating point observational and photographic data. The metadata structure of CrDAP was extended based on ISO 19115, which is a geographic information metadata standard of the International Organization for Standardization (ISO).

Keywords: Cryosphere, CrDAP, Metadata, ISO 19115

1. INTRODUCTION

The Eurasian cryosphere is an integral component of the earth's climate system, comprising frozen surfaces and structures such as glaciers, frozen ground and snow elements. Large fluctuations in the Eurasian cryosphere have been the focus of recent reports, including the IPCC AR-4 report (IPCC, 2007), which expresses great concern regarding the social impact of such fluctuations. Currently, worldwide snow and ice data collection methods are promoted by United States data centers including the National Snow and Ice Data Center (NSIDC) and the National Climate Data Center (NCDC). However, frozen ground and snow data are not stored in an international data organization center such as the World Meteorological Organization (WMO), at which the international and systematic data archive is poor. For a better understanding of the cold regions and wide fluctuations characterizing the Eurasian cryosphere, which encompasses several countries, a wide-reaching and improved coordination of cryospheric data is imperative. The IGOS-Cryosphere (IGOS, 2007) and IPY (Krupnik et al., 2011) have also pointed to the need for such a system and, specifically, one that includes Global Earth Observation System of Systems (GEOSS) data archiving functions.

2. PURPOSE

The purpose of this project is to showcase the reality of global environmental changes in the Eurasian cryosphere, by promoting data collection efforts and the cataloging of vital new and legacy cryospheric observations to the public via widespread data digitization. While this project operates in Japan, the data sources will come from all cold regions in the Eurasian countries and will be integrated into our system to promote their release and subsequent publication. Our database, called the Cryosphere Data Archive Partnership (CrDAP; accessible at: http://www.jamstec.go.jp/acdap/; Fig. 1) is intended to disseminate observational information to advance scientific understanding of the global climate system.

3. THE DEVELOPMENT OF METADATA

3.1 Types of data related to cryosphere research

In the framework of cryosphere research, a variety of data types exist in various formats. Designing a metadata structure that incorporates and integrates all data types is a challenging endeavor. Table 1 describes the data

classifications and forms used for planning CrDAP user registration formats. In the first phase of database development, the ground station meteorological data and photographic data were selected, and the structure of the metadata encompassing these data sets was determined. It functioned to archive datasets and metadata sets.

Figure 1. The top image of Cryosphere Data Archive Partnership (CrDAP)

Table 1. Data classification and format form of planning registration to CrDAP. P is point observation, A is areal observation and G is grid data.

Data category	element	detail	Type
Ground station Observation	(1) Meteorological and climate data	air temperature, humidity, wind, pressure, radiation, precipitation, soil moisture	P
	(2) Snow	depth, density, coverage, etc.	P
	(3) Frozen ground	ground temperature, melting depth, ice volume	P
	(4) Glacier	mass balance, glacier type, velocity, ice depth, ice temperature	P
	(5) Hydrology	river discharge, river frozen condition (icing and melting date), water temperature	P
	(5) Lake	water level, lake area, lake frozen condition (icing and melting date), water temperature	P, A
Remote sensing product	(1) Snow	distribution	G
	(2) Glacier		
	(3) Vegetation		
Map	(1) Frozen ground map	distribution	A
	(2) Vegetation map		
	(3) Glacier map		
Photograph and picture	(1) Glacier photo	aerial photograph, ground-based photograph	P, A
	(2) Frozen ground photo		
	(3) Vegetation photo		

3.2 The metadata structure

Before designing the CrDAP metadata structure, we investigated the status of existing atmospheric database projects. Assuming the exchange of metadata to a variety of portals, metadata structure is a necessary condition that conforms to international standards. We found several such projects underway, including the Inter-university Upper atmosphere Global Observation NETwork project (IUGONET; Kouno et al., 2011), which targets data in a polar region archive format. This project serves as an important step toward developing a system for inter-agency metadata cooperation. The Data Integration and Analysis System (DIAS) project (Kinutani et al., 2010), aims to gather all information related to earth observations and others project, and manages unprecedented amounts and levels of data of varied quality. The DIAS project analyzes a data document, collects the metadata and changes the metadata items to correspond to those of ISO 19115 (ISO, 2003). DIAS metadata structure can be viewed as a framework that encompasses the lowest common denominator among common items related to earth observations.

The CrDAP will be compatible with, and potentially extend beyond, the DIAS metadata structure. We will include fine-grained information into the CrDAP metadata structure and into the data sets themselves so that a user can generate a single metadata data set that includes two or more observation points, and retrieve all requested information on an individual observation.

In CrDAP, the metadata item numbers reflect the data types, and the CrDAP metadata structure can be divided into two units, a core unit common among the datasets (e.g., data set name and data provider information) and a variable item unit (e.g., site information) depending on the data item type. The latter unit type enables replacement of metadata as needed. A conceptual diagram of the data and metadata registration flow is shown in Figure 2. A variable item unit can be increased based on international standards, as dictated by the data form type. We patterned our CrDAP metadata structure after these features of the DIAS database, and made it expandable to other metadata objects and data types as needed. We used Microsoft Excel (Microsoft corp.) for entry of CrDAP metadata as an input tool. Because CrDAP assumes that the metadata structure will be populated by users and data providers in the future. Other projects often adopt Web interface registration methods for metadata online; however, the CrDAP project utilized Excel because it is the most common format used to create metadata files and is utilized and stored in an off-line environment, residing securely on the user's desktop. For a metadata author and a site administrator to check the contents of registered work, it is useful to keep the description information associated with the metadata. Finally, the metadata is transformed to XML format using an Excel macro and is used for registration work to the CrDAP site.

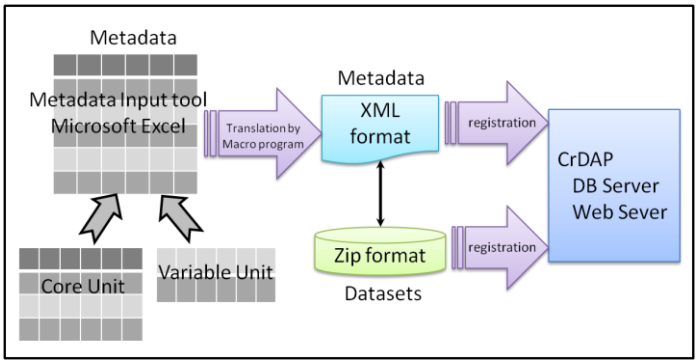

Figure 2. A conceptual diagram of data and metadata registration flow in CrDAP

4. APPLICATION OF CrDAP

4.1 Simple search and detailed search

In CrDAP, the data register a metadata XML file and pass the contained data set information, including the fine-grained information contained in the metadata itself, to a MySQL relational database management system for data storage and management. A data search can be conducted by querying the MySQL database in CrDAP. The system can distinguish between simple and detailed searches, and indicates whether the data set returned with the specified search conditions is partial or complete. Data search criteria can also request the fine-grained information embedded within the data sets.

4.2 Quick look

CrDAP does not implement a visualization function. However, it is vital to review the data summary before downloading the data set, which we achieved using Quick Look (Apple corp.). To create a Quick Look image, text and binary data are drawn and prepared independently. The resultant Quick Look image can be displayed on a browser presenting the search results. Photographic data is targeted data visualization, have implemented the Web user interface for CrDAP.

4.3 Download data

Restricted data set access files are displayed in different colors in the list of search results. A validated user ID and password are required to access the data download function. All accessible data sets are available for download in compressed zip formats.

5. CONCLUSION

The CrDAP team has constructed a highly flexible and powerful archive system for storing and managing cryospheric data and metadata. The first phase of development included determining the requirements for registration of the metadata structure for recognizing ground station meteorological and photographic data. The system is downloadable from a Web server, and the CrDAP metadata structure was made to be extended for recognizing and storing other data types, based on the ISO 19115 data standardization criteria. The CrDAP metadata structure performed the design which replaces metadata structure per unit of data type. This new metadata structure is mandatory for future data manipulation requirements, and we can respond to additional requirements by the simple replacement of a unit.

6. REFERENCES

IGOS (2007) Integrated Global Observing Strategy Cryosphere Theme Report - For the Monitoring of our Environment from Space and from Earth. Geneva: World Meteorological Organization. WMO/TD-No. 1405. 100 pp.

IPCC (2007) Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K and Reisinger, A. (eds.)]. IPCC, Geneva, Switzerland, 104 pp.

ISO (2003) ISO 19115:2003 Geographic information –Metadata. Geneva, International Organization for Standardization (ISO).

Kinutani, H., Shimizu, T., Yoshikawa, M., and Kitsuregawa, M. (2010) A Multidisciplinary Researchers' Collaboration for Disclosure of Earth Science Data in DIAS. The Institute of Electronics, Information and Communication Engineers. Technical Report of IEICE., 110, 328, DE2010–33, 45–50, (in Japanese)

Kouno, T., Koyama, Y., Hori, T., Abe, S., Yoshida, D., Hayashi, T., Shinbori, A., Tanaka, Y., Kagitani, M., Ueno, S., Kaneda, N., and Tadokoro, H. (2011) Development of Metadata Database for Upper Atmosphere Observation by using DSpace. The Third Forum on Data Engineering and Information Management - DEIM 2011 - DEIM Forum 2011 C8-5 (in Japanese).

Krupnik, I., et al., Editors. (2011) Understanding Earth's Polar Challenges: International Polar Year 2007–2008. University of the Arctic, Rovaniemi, Finland /CCI Press (Printed Version), Edmonton, Alberta, Canada and ICSU/WMO Joint Committee for International Polar Year 2007–2008.