

The metadata collection of Earth observation project data products by DIAS

Hiroko Kinutani^{1*}, Toshiyuki Shimizu², Masatoshi Yoshikawa², Masaru Kitsuregawa¹, Toshio Koike¹

¹The University of Tokyo, ²Kyoto University

Our project, DIAS (Data Integration and Analysis System) started in 2006, has a purpose of constructing data infrastructure that can integrate earth observation data, numerical model outputs, and socio-economic data effectively. DIAS also has a purpose to create knowledge enabling us to solve earth environment problems and to generate socio-economic benefits. From October 2010, we have released data of DIAS with Document-metadata, describing about dataset in English and Japanese. Anyone can use the DIAS data discovery system by accessing <http://dias-dss.tkl.iis.u-tokyo.ac.jp/ddc/finder>, and can download data files of 187 datasets through the system.

Through Earth Observation implementation plan we are aiming centralized registration of metadata about Earth observation data, the published data for the purpose of sharing and of observation towards the integrated use of the Global Earth Observation System of Systems in Japan. As a first step, we are supporting the creation of metadata for projects as listed on the "Implementation Plan for Earth Observation in Japan's FY 2011". We have conducted a survey of metadata about the contents of the data that has been collected in the projects to 155 projects for Earth Observation implementation plan of FY 2011 of the Ministry of Education from June 2012. Answered 95 projects, the number of respondents were 130. The main agencies who have responded are JAXA, NICT, JAMSTEC, NIPR, FFPRI, NIES, GSI, JMA, JCG, NIAES etc.. About data release, all data is published from 55 respondents, a part of data is published from 29 respondents and 46 respondents are answered as unpublished. We found that 65% of organizations have the means to publish data products because they are consist of research institutions and public service institutions. 90% of the data that has been published is already implemented quality control. Many climate information, weather, disaster, environment, earth science, ocean ecosystems are areas where the product is the appropriate data. In-situ, marine, earth observation satellite, simulation models are platforms where the product is the appropriate data. Survey results are published in <http://dias-d.tkl.iis.u-tokyo.ac.jp/dias-report/enquete/>. In February 2013, 292 data products have been collected from 155 projects and about 2,200 metadata have been collected.

Based on the preliminary survey, we are identifying the data products in each project, and are creating or collecting their metadata. Against the 45% respondents not creating metadata, we ask to create metadata using the DIAS metadata management system. Using the DIAS metadata management system, everybody can create the ISO 19139 format metadata (XML) and PDF and HTML dataset documentation. 55% respondents are answered have already created the metadata. The format of the metadata that is already created are various; DIF, JMP2.0, EML, for example. When respondents have metadata complying with the standard specifications for metadata related to Earth observation data, we collect metadata from them. But, when respondents don't have metadata complying with the standard specifications for metadata related to Earth observation data, we ask to create their metadata using the DIAS metadata management system.

We will plan the design and development of an integration and mediation metadata system in order to extract the most characteristics of various fields of data from metadata written by several kinds of metadata formats, and to provide basic information for integrated use of data. We will also provide search services that can have access to the data products from Earth Observation projects.

Keywords: DIAS, Earth Observation project, Satellite data, Model output data, In-situ data, metadata