

Japan Geoscience Union Meeting 2011

(May 22-27 2011 at Makuhari, Chiba, Japan)

©2011. Japan Geoscience Union. All Rights Reserved.



U004-P01

Room:Convention Hall

Time:May 27 10:30-13:00

Data Catalog Publication System in JAMSTEC

Yasunori Hanafusa^{1*}, Hideaki Saito¹, Mika Ichino¹, Satomi Minamizawa¹

¹JAMSTEC/DrC

Japan Agency for Marine-Earth Science and Technology (JAMSTEC) has developed the "JAMSTEC Data Catalog", which helps users to find comprehensively databases or datasets which publicize data obtained from JAMSTEC research activities.

JAMSTEC has explored global oceans especially deep sea, atmosphere, solid earth and biosphere and has also performed many kinds of simulations using High Performance Computers such as the Earth Simulator. JAMSTEC publicizes such data via databases specified by data types or data sites of research projects. For marine-earth observational data JAMSTEC has developed a web GIS based "JAMSTEC Data Search Portal" (J234-002, JpGU Meeting 2009), which enables users to search data by observational points or cruise tracks at single site. However there was no adequate search service to find a database or a dataset.

To handle metadata of databases and datasets in single manner the Data Research Center for Marine-Earth Sciences (DrC) adopted Directory Interchange Format (DIF) in the Global Change Master Directory (GCMD) of NASA as a metadata standard for JAMSTEC's databases and datasets which extend over various scientific categories. Designed to cover wide range of category DIF is suitable for JAMSTEC's metadata standard and also fit to be handled in information systems due to its XML format. DrC has compiled information on databases and datasets in JAMSTEC in DIF format and registered them to GCMD. At the same time DrC has also prepared Japanese version of those metadata in DIF format.

JAMSTEC Data Catalog is a web system which enables users to find and browse metadata in DIF format.

Data Catalog initializes a database system using pre-defined XML schema as a XSD file. When a metadata is registered, Data Catalog digests it into database and puts it to a full text search engine. Users can select metadata from hierarchical categories shown in a tree or a list style or search and narrow metadata by simple keyword search or complex search. Metadata have URL links leading users to databases or datasets and cross links to change the page languages. A data manager can select one or more terms for category classification from metadata schema.

Data Catalog is able to handle different metadata sets with other metadata schema than DIF. Registering metadata schema additionally DrC has developed on the same system the "Document Catalog" publishing cruise reports, technical reports and public relationship documents etc. and the "Geophysical Exploration Data Catalog" publishing single/multi-channel seismography data.

As of now Data Catalog disseminates mainly metadata of databases and data publishing sites in JAMSTEC. DrC is going to extend handling metadata to datasets of various research projects. Although a data manager registers metadata now, DrC is developing a metadata input tool which helps researchers to make metadata in DIF format both in Japanese and English by inputting information on a web interface by themselves.

Keywords: metadata, XML, schema, DIF