Don't leave your Data unattended at any time!

Dirk Fleischer, Andreas Czerniak, Pina Springer, Carsten Schirnick, Hela Mehrtens
GEOMAR Helmholtz Centre for Ocean Research Kiel
datamanagement@geomar.de

GEOMAR

IN53C-1581

File based Import

Autonomous measurements have become an important means in science. Better technical equipment and enhanced telecommunication enable more and more distant device control.
The result of autonomous measurements is a better resolution, i.e. glider can measure physical parameter in a larger region and during a longer time than single stations on a cruise.
Data from such instruments are transfered by satellite and reach a server without human interaction. Here the workflow from measurement to digital preservation is uninterrupted.

Handwriting Recognition

Using the method definition from as basis for lab or field notes enables the application of handwriting recognition technology for data capturing. Handwriting recognition can become handy direct in the field while sampling or already in the lab with measuring protocolls at the table to document the measurements. This technology reduces the digitalization

Web Interface

In the lab or on board a research vessel the use of an webinterface may be sufficiant for data input. The definition is used to render the web forms designed by the scientist or lab staff itself.

IGSN
GND
OrcID

Data Analysis and Interpretation

The KDMI serves as a master database for analysis-optimised data warehouses.
The scientific interpretation includes the comparison with other data. Therefore the analysis data warehouses are read-only databases for the provenance data system. This dependency enables the combination of unpublished and OpenAccess data from the public domain or collaborating scientists.
Those analysis data warehouses can be optimised in structure and type (Geographical database, GraphDB, etc.) to the special tasks necessary for the data analysis. The scientists can do extractions, charts and statistical analysis they need for their science while the original data stay untouched and secure. If changes to the original data become necessary these changes can be done and will be brought through to the slave systems as well.
Accordingly the quality control of the data is the actual scientific analysis process and ensures high quality data for long term use after the initial interpretation is done.

International Geo Sample Number (IGSN) - AuthorID - OrcID - PersonID

Data Provenance

DOI - Dataset Archiving

+ DOI

Source: www.techweez.com

Data Provenance

The basic method of sample analysis is the provenance of scientific data. The consistent information about this method and its application is the most powerful quality information for a resulting dataset.
Therefore the architecture of the Kiel Data Management Infrastructure (KDMI) is a two level data capturing approach.
The first level is the method definition provided by a data creator before any data is actually available. This definition is the general information on the measuring process and it enables the second level.
The second data capurting level can be regarded as the realisation of the method definition. The innumerable repetitions of the provided definition and all its sequenciell data are connected at all times to their method definition, and all the provenance information is available for later quality assurance of scientific use.

Data Publication

Data publication and open access to scientific results has to become a simplified and easy process for each scientist, that can be done without any additional effort.
Data warehouse technologies are already well known in computer science and provide scientific enhancements and new algorithms. This is a field in computer science which is evolving all the time and improvments become available on a regular basis.
The technology behind the Extract Transformation and Load (ETL) processes is already well known. Using this knowledge and combining it with the application of scientific workflows is a new concept for scientific data management and provides a one click solution to data publication processes. We strongly encourage scientists to publish data in a data archive with citable DOI.

future ocean
KIEL MARINE SCIENCES
www.futureocean.org

T,S,v
N,P,Fe
O₂
www.sfb754.de

C o n t a c t
Dirk Fleischer
Data- and Computing Center | Data Management
GEOMAR Helmholtz Centre for Ocean Research Kiel
Wischhofstr. 1-3, 24148 Kiel, Germany
Phone: +49 431 600 2296
E-Mail: dfleischer@geomar.de

ECO₂
Sub-seabed CO₂ Storage:
Impact on Marine Ecosystems
www.eco2-project.eu

HELMHOLTZ | ASSOCIATION