Kiel Data Management Infrastructure for Marine Sciences

Pina Springer, Hela Mehrtens, Dirk Fleischer, Carsten Schirnick & Andreas Czerniak

GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany – http://portal.geomar.de – datamanagement@geomar.de

GEOMAR

Motivation

The aim of the Kiel Data Management Infrastructure (KDMI) is to provide ONE central place for description, storage and archival of research data for all marine research in Kiel, independent of project status but with particular access restrictions for each project. This virtual research environment facilitates preparation of data for journal publication, data exchange within the project and data publication to World Data Centers. This approach ensures data sharing during the project's runtime and open access plus long term availability of data beyond the project's end.

State of the Art

Kiel's scientific sites are using a virtual research environment. Each project is a community on its own and can use optional software components (Fig. 1).

Fig. 1
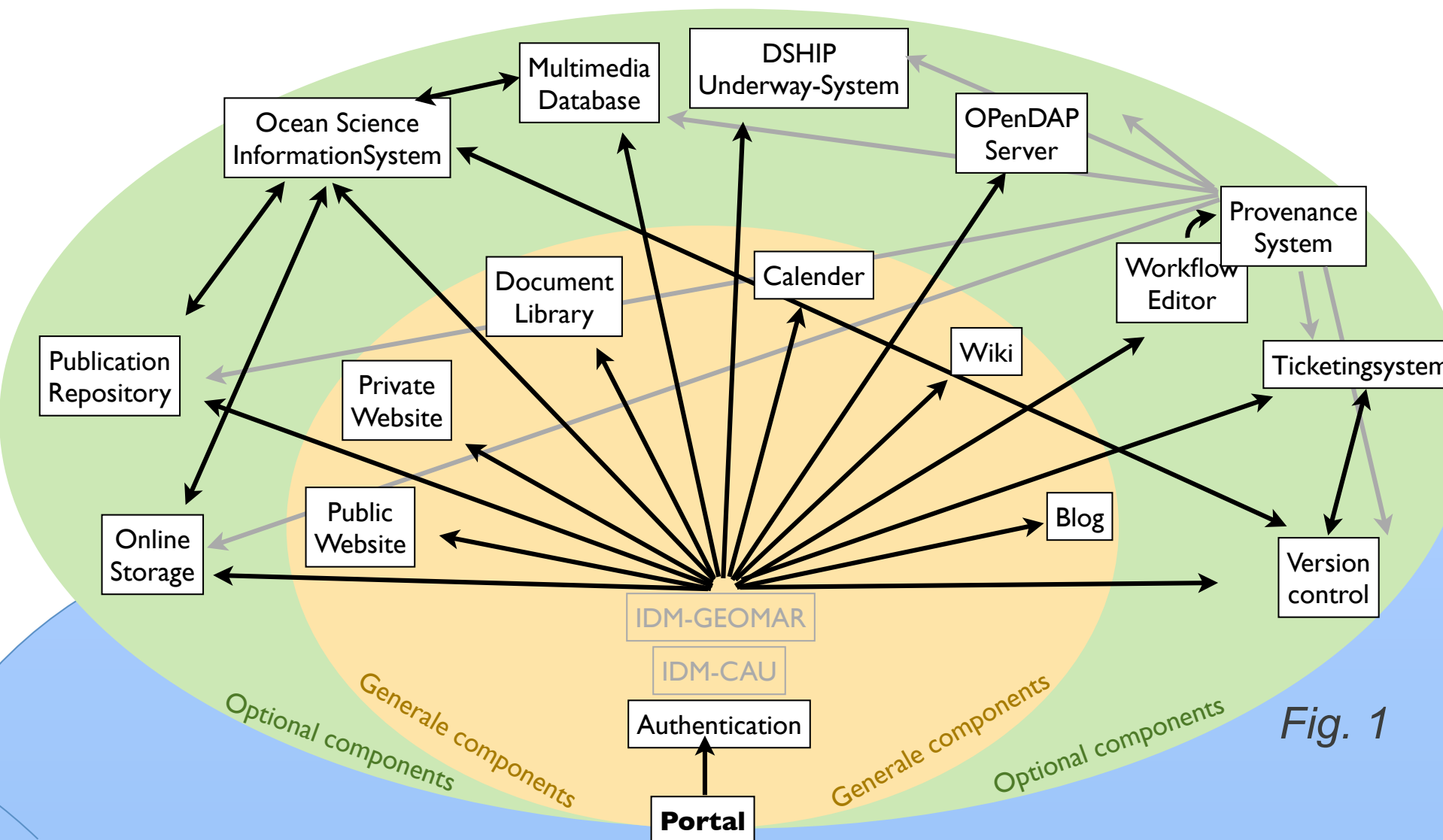
The Team

The Kiel Data Management Team (KDMT) is a joint group of large-scale projects at GEOMAR and Kiel University: the Cluster of Excellence 'The Future Ocean', the Collaborative Research Center 'SFB 754', the EU FP7 project 'ECO2', the Helmholtz Initiative and Networking Fund Project 'MaNIDA' and the BMBF funded projects 'BIOACID II' and 'SOPRAN III'. This nation-wide unique approach enables a coordinated development process for non redundant data management activities and a team with broad knowledge and acceptance among the scientists.

Fig. 2: Data capturing by handwriting recognition just in time. Smart forms and pens provide new solutions for data management and field work.

VIRTUAL RESEARCH ENVIRONMENT

Fig. 6: Ocean Science Information System (OSIS).

The Portal

At portal.geomar.de every project has its own private pages with access to document library, publication lists partly with full text access and the central information hub for expeditions and numerical model experiments (OSIS, Fig. 6).
The information hub now contains more than 1100 expeditions and 3300 files. More than 8000 links relate to print publications, internal (e.g. THREDDS) or open data repositories (e.g. PANGAEA, GEOROC) or the GEOMAR sediment core and sample repository. The portal is part of the first KDMI level to gather the scientific data for sharing and further internal usage within KDMI.

Fig. 5: Citable data sets related to print publications (Fig. 3) in the data library PANGAEA.

Data Capturing

The usage of smart pens and integrated software components allows lossless transfer of data and information from paper into digital content (Fig. 2). Thus, scientists can simultaneously pursue and access handwritten documents while having a digital realization of their information and data at hand.
The seamless integration between data or sample creation and data system requires an unique identification of data and sample. Therefore GEOMAR is a funding member of the IGSN e.V. to implement world wide unique sample references, locations and citations.
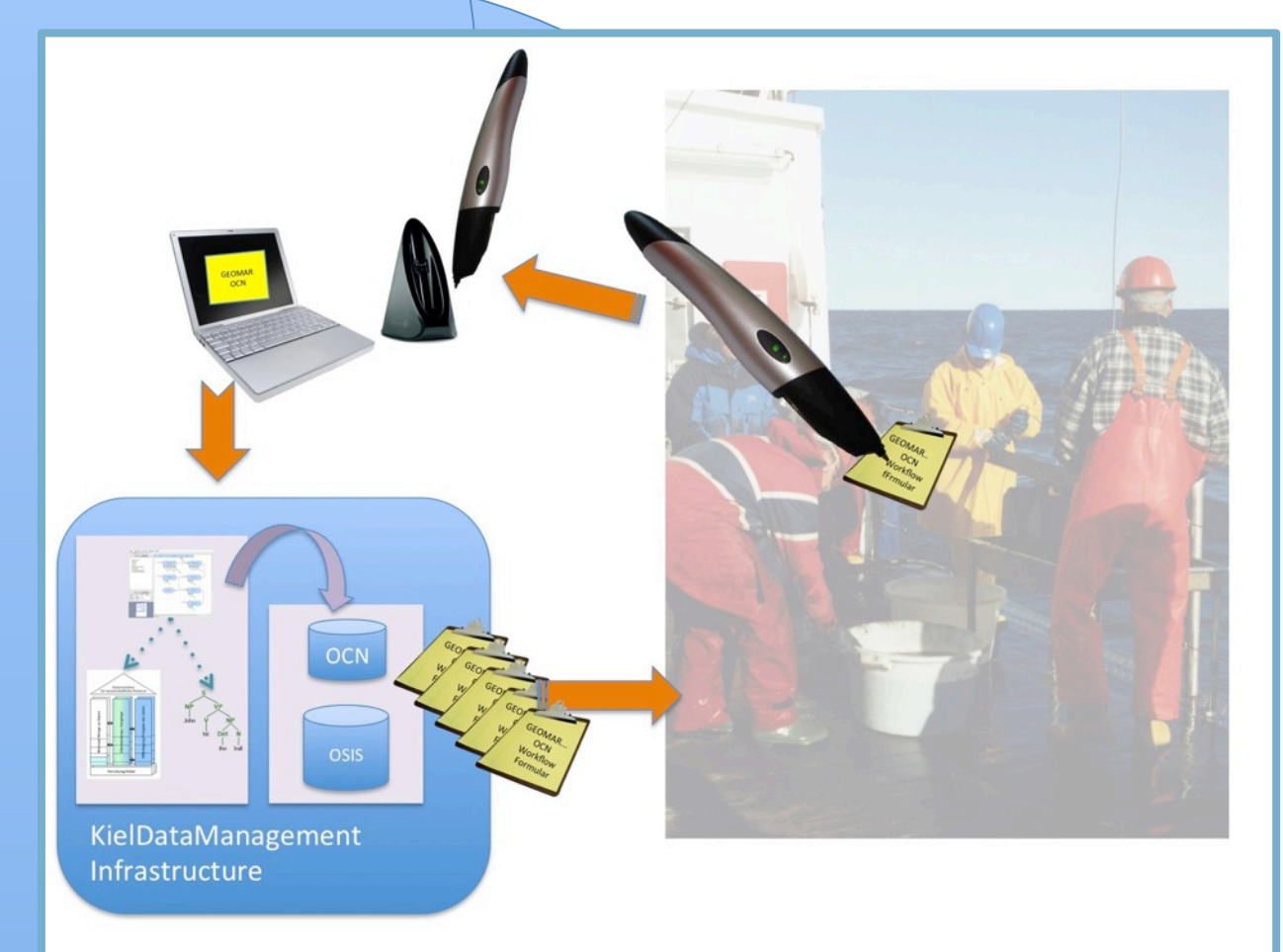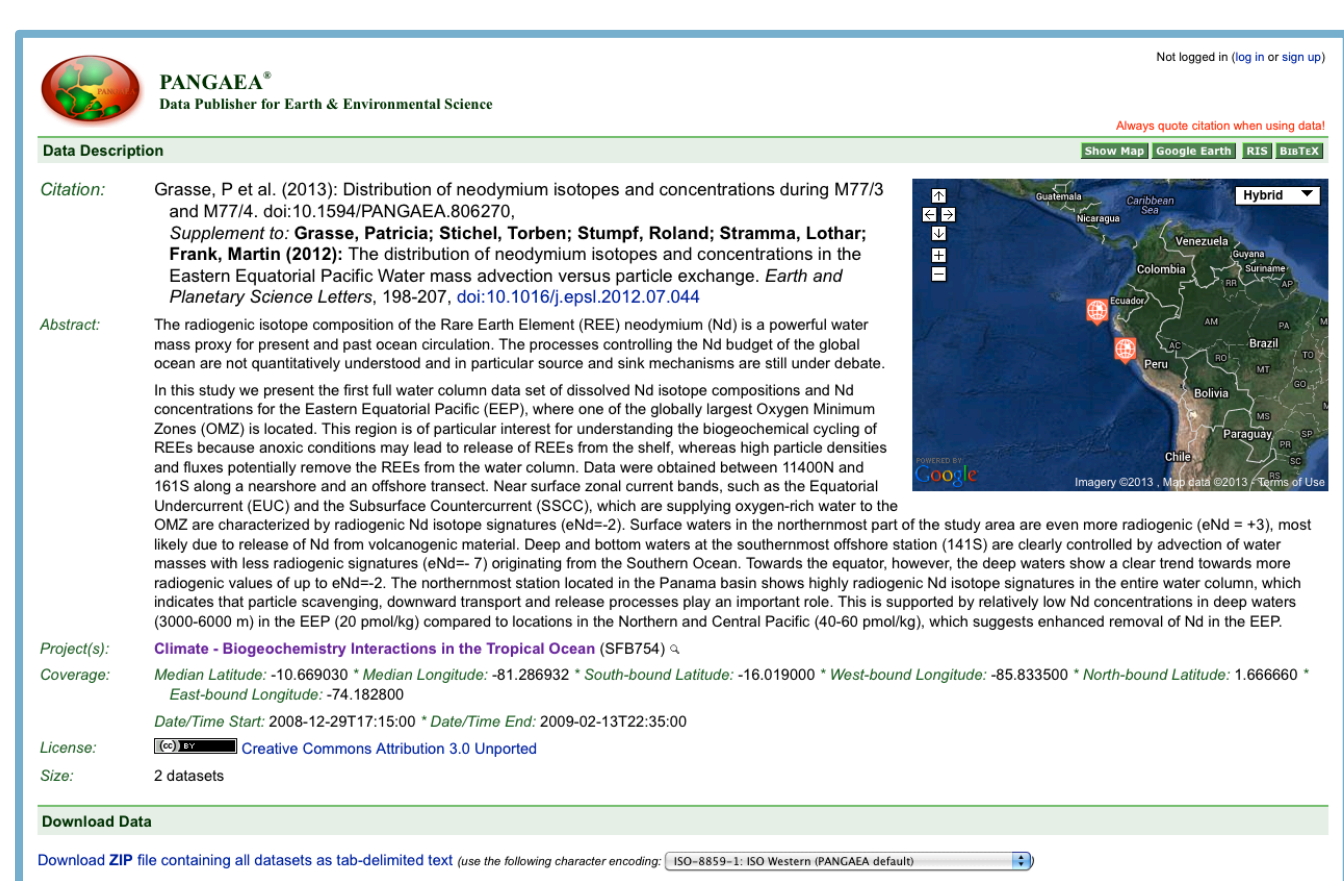

Fig. 3: Project publication list with additional icons for connected data. A web service provided by the data library PANGAEA.

Fig. 4: Search cruise tracks, print publications, sampling locations, glider tracks, sample descriptions and detailed cruise information as georeferenced information in GoogleEarth®.

Publish data and literature linkage

After a moratorium data should be publicly accessible (Fig. 5). A workflow for publishing data in a World Data Centre is developed in a joint collaboration with the DFG project *PubFlow* at CAU. The EPrints based repository OceanRep lists paper publications of GEOMAR and related projects. Customized project publication lists are displayed on their public websites and include links to corresponding expeditions or models in OSIS. Related supplementary data available at WDC PANGAEA are visualized and linked by an icon (Fig. 3).

Fig. 4/a: Detailed cruise information with literature and links to data sets.

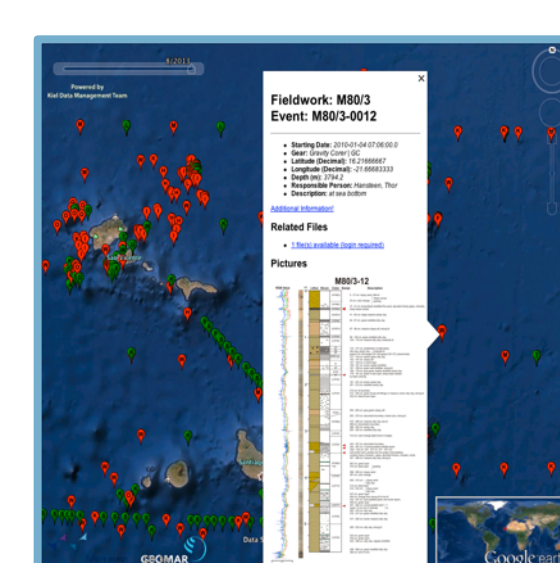*Fig. 4/b: Underwater tracks of autonomously measuring gliders.*

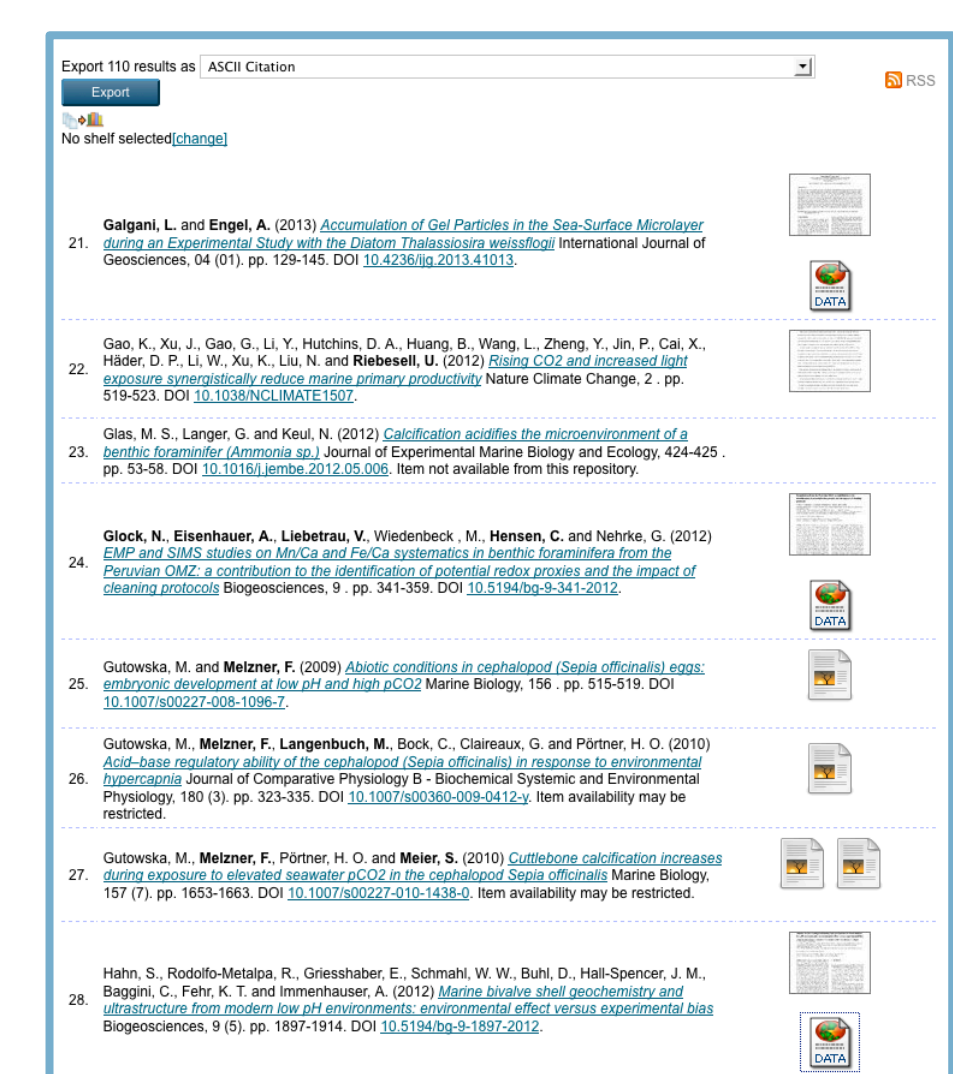*Fig. 4/c: Detailed sampling description and storage location.*

Search data with KDMI GeoSearch

Cruises and expeditions are visualized with the 3D virtual globe GoogleEarth® (Fig. 4). Each sampling event is geographically referenced and reveals sampling technique, responsible researcher, images and further description. Information about available data files, publications and other related links is displayed and linked back to the Ocean Science Information System (OSIS) and the institutional repository OceanRep.