

Data Management: The Data Life Cycle



Pina Springer, Dirk Fleischer, Hela Mehrstens, Carsten Schirnick, Andreas Czerniak
datenmanagement@geomar.de

An Integrated Data Management Approach

The scientific site of Kiel provides support for projects with data management requirements due to project size or interdisciplinarity. This infrastructure is the Kiel Data Management Infrastructure (KDMI) and was initially created by SFB574, SFB754, Excellence Cluster 'The Future Ocean' and the GEOMAR | Helmholtz Centre for Ocean Research Kiel.

To achieve public data availability from publicly funded projects by the end of the funding period it is necessary to initiate the data acquisition during the data creation process. Accordingly the KDMI uses a three level approach to achieve this goal in SOPRAN III. Data management is already involved in the planning of expeditions or experiments. The resulting schedule for data files can be used by the project coordination to increase the efficiency of data sharing within SOPRAN

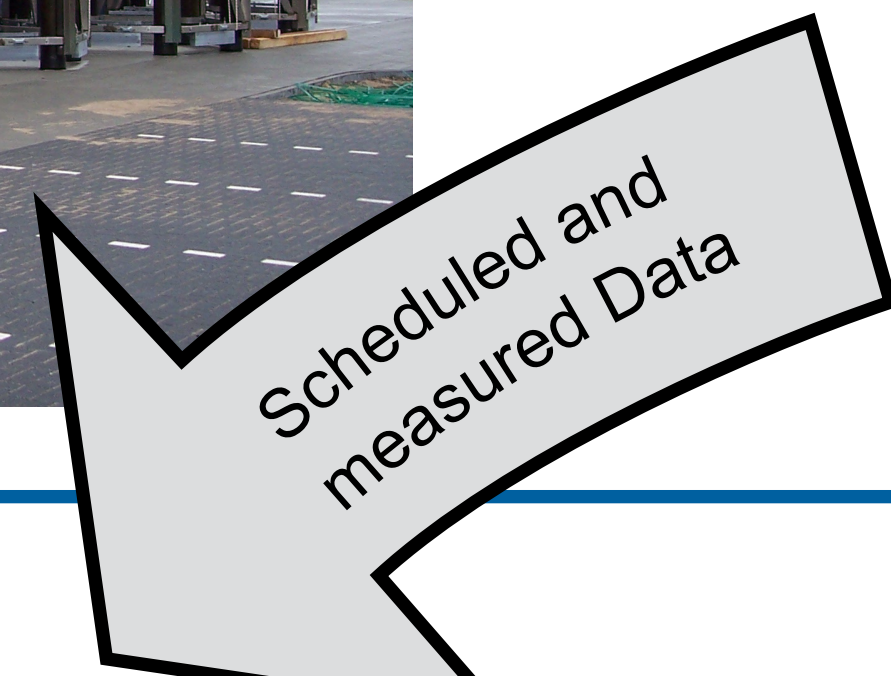
III. The scientists provide files with basic metainformation, which are available within the virtual research environment as soon as possible to all project members. Final data will be transferred to PANGAEA for long term availability when the data are analysed and interpreted in a scientific publication or by the end of SOPRAN III.

The Kiel Data Management Team offers a portal for all GEOMAR and University Kiel marine projects. This portal will be used in SOPRAN III in combination with PANGAEA to fulfill the project's data management requirements and to enhance the data sharing within SOPRAN III by a file sharing environment for preliminary data not yet suitable for PANGAEA.



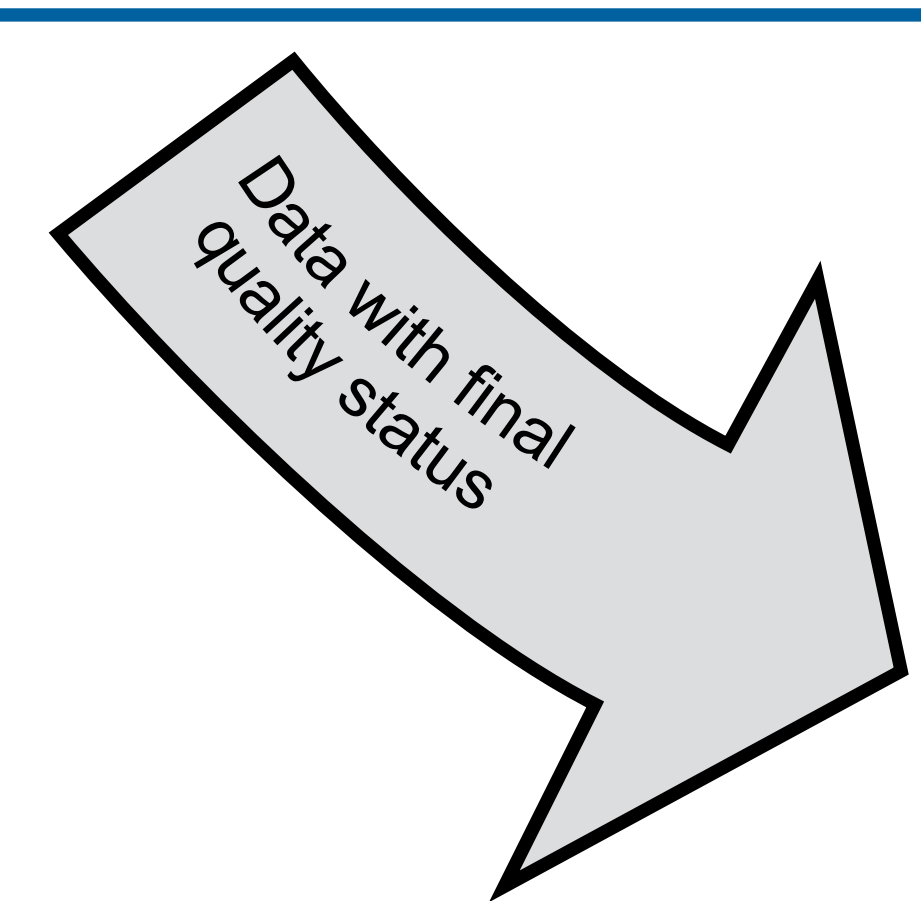
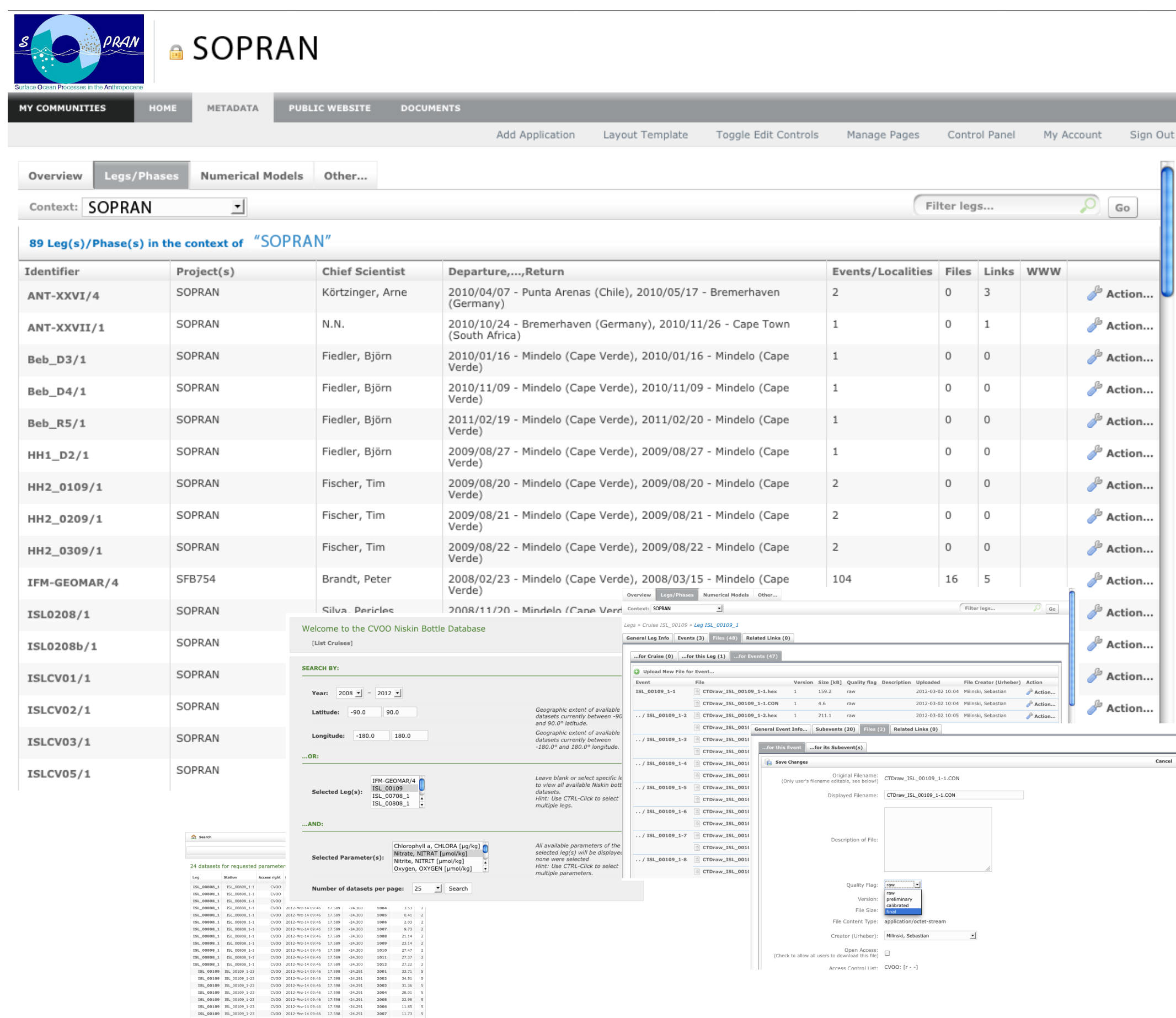
Planning of Data Acquisition Cruises, Field Work, Experiments & Modelling

- Data checklists and PIs
- Schedule of cruises
 - Preliminary Data Availability
 - Finalization of Data Sets
- Experimental Setup
 - Preliminary Data Availability
 - Finalization of Data Sets
- Models
 - Data Availability

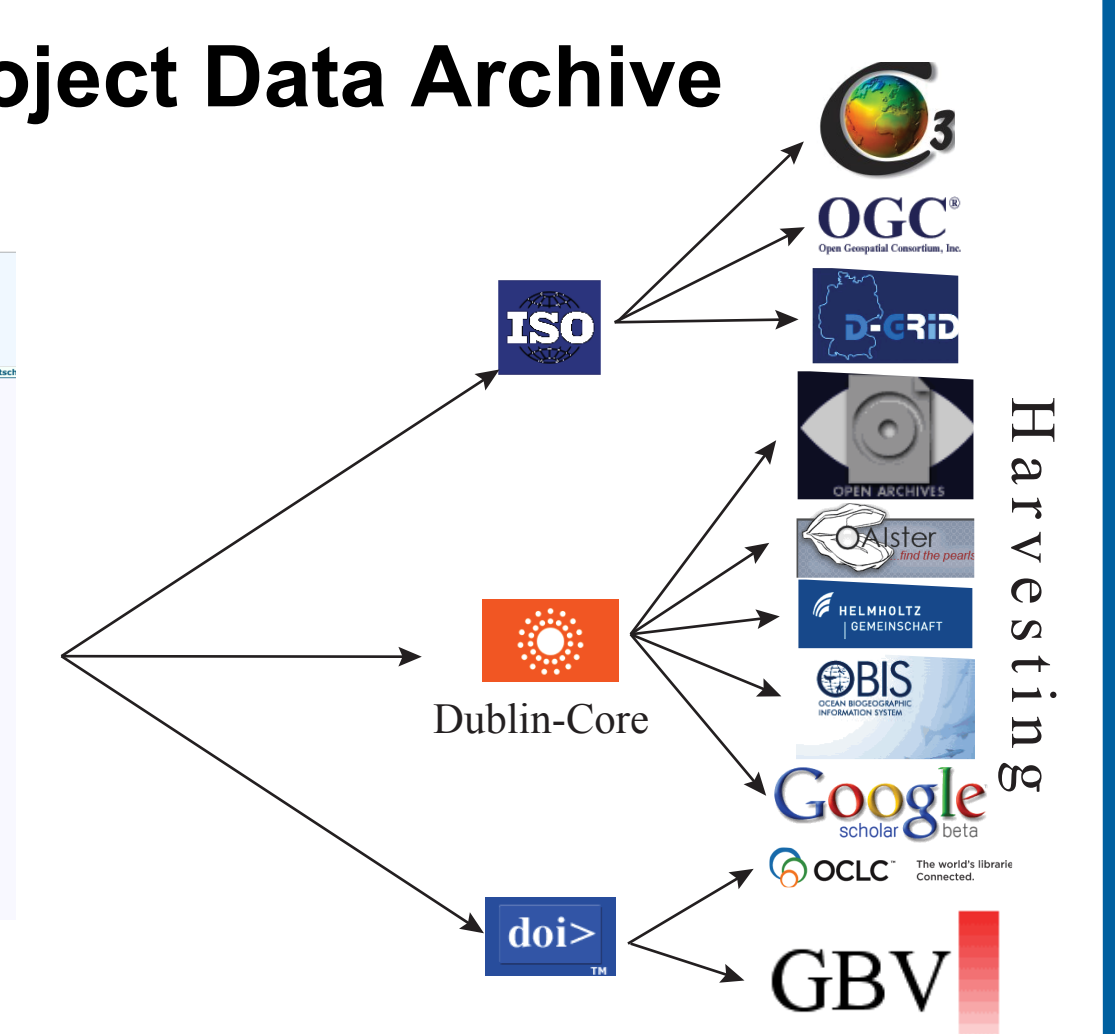


Virtual Research Environment

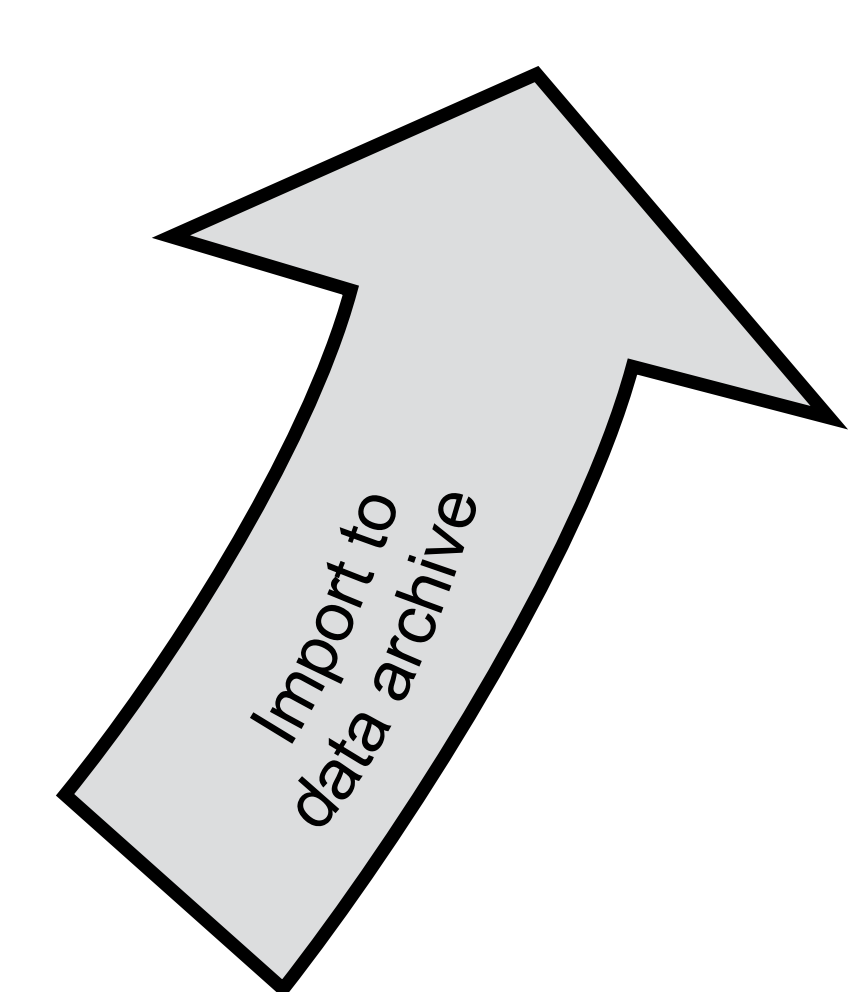
The collaboration portal (<https://portal.geomar.de>) is based on a community concept. The registered users can build groups (communities) of any scale and are provided with a lot of standard features such as a Wiki for collaborative knowledge acquisition and writing but also a document library with version control for file sharing. On request additional features can be made available to the community for more specific tasks. Such features can be a version control for model source code, the publication repository (<https://portal.geomar.de/oceanrep>), the metadata acquisition and file sharing tool for field expeditions and model output and/or the database for Niskin Bottle Data.



PANGAEA - Project Data Archive



- long-term preservation, publication, open access of metadata and data
- each dataset is citeable by a unique and persistent Digital Object Identifier (DOI)
- availability of SOPRAN data to the global scientific community beyond the lifetime of the project



Online data submission and editorial system

- Data Publication Tracking
- Author Approval
- Quality control
- Completeness of Meta-data

