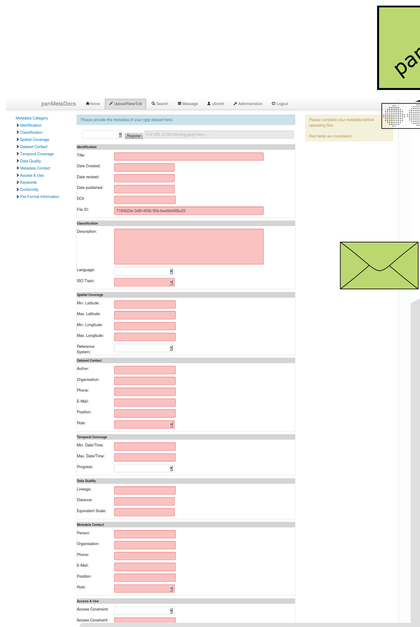
GFZ
Helmholtz Centre
POTSDAM

panMetaDocs and Datasync
providing a convenient way to share and publish research data

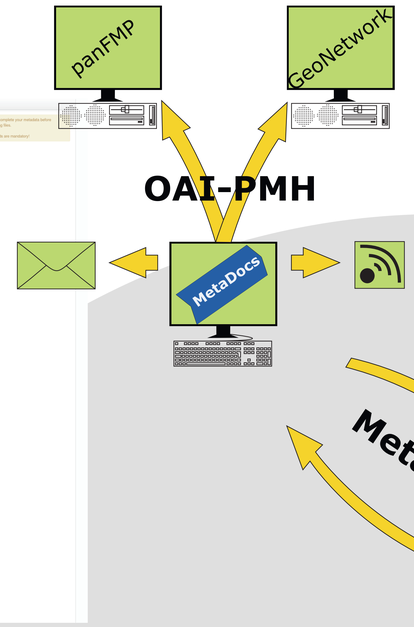
Damian Ulbricht and Jens Klump, GFZ German Research Centre for Geoscience, Potsdam, Germany
E-Mail: {ulbricht, klump}@gfz-potsdam.de

IN41B-1605



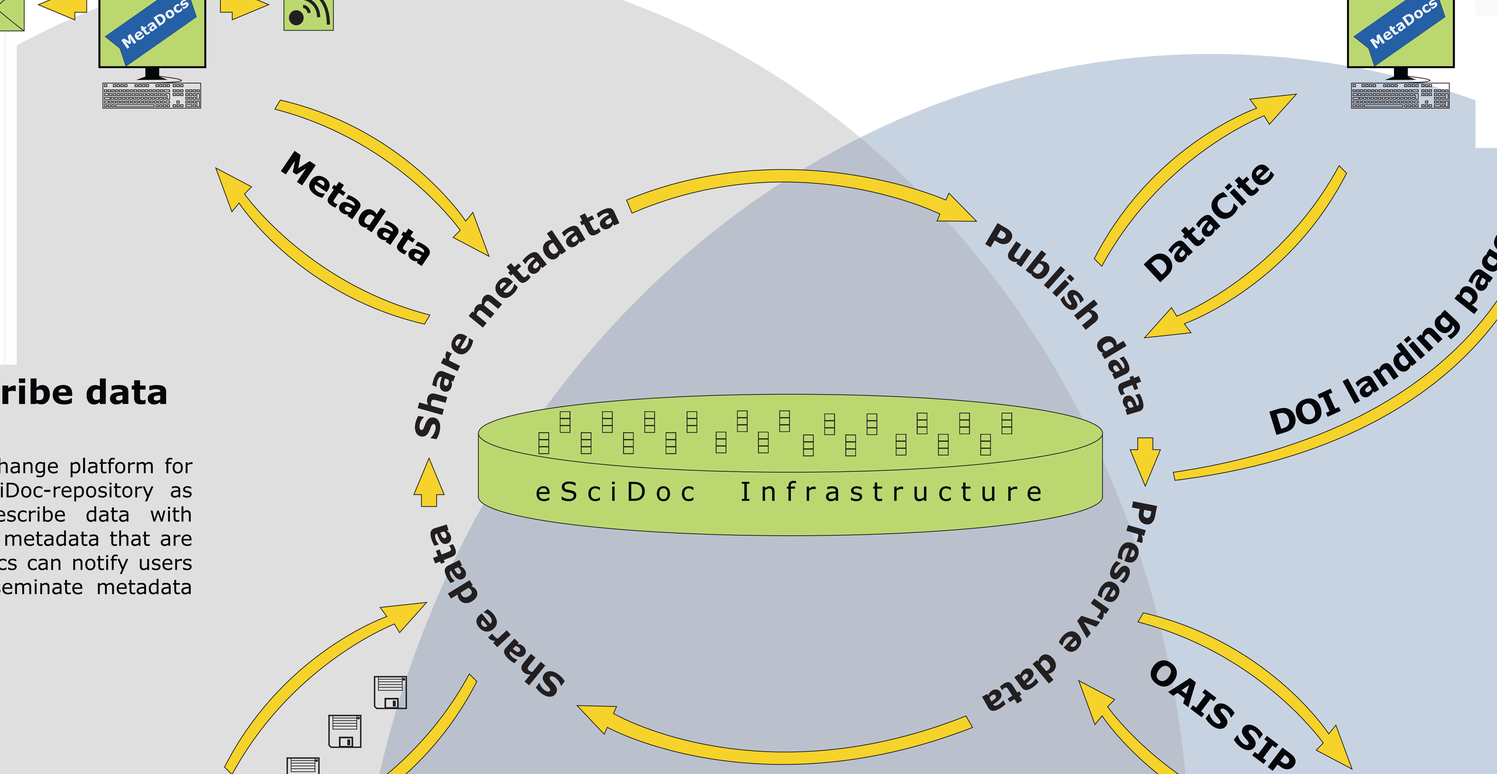
panMetaDocs - describe data

panMetaDocs is a web-based data exchange platform for federated projects that uses an eSciDoc-repository as storage. The software allows to describe data with metadata and make use of supporting metadata that are stored in a project setting. panMetaDocs can notify users about recent changes and it can disseminate metadata through various channels.



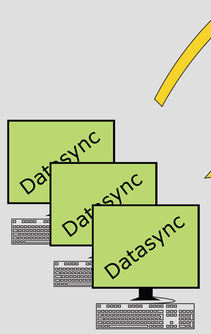
Publishing Data

When it comes to publication panMetaDocs is used to complement the data with DataCite metadata to receive a dataset DOI and finally publish the data. Then, the eSciDoc-item that holds data and metadata changes only its public status while all other information we stored before is left untouched.
The publicly visible download page is generated from the metadata in the eSciDoc-item by XSL stylesheet transformation.



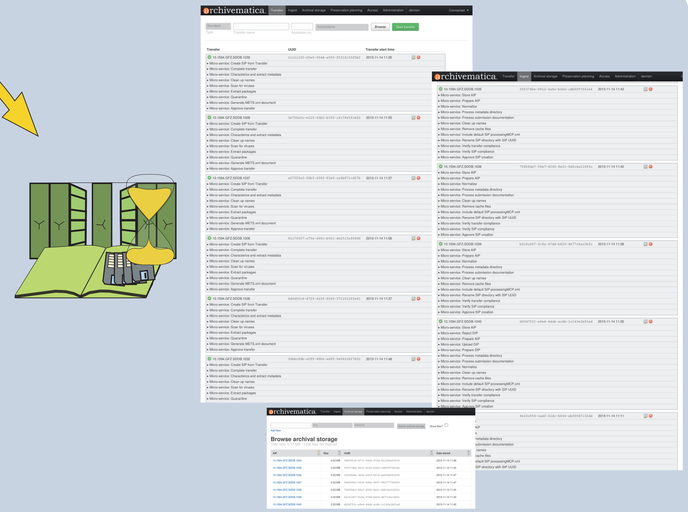
Starting point for preservation

Registration of dataset DOIs comes with the obligation to keep the data accessible in the future. To achieve this we tested the ingest (import) of research data into OAIS compliant archives.



Datasync - share data

Datasync provides a way to synchronize folders with an eSciDoc-repository in the backend. The software is written in JAVA and can be controlled via the command line. Currently, we are looking for partners to help us develope a more advanced graphical user interface.



eSciDoc

GFZ uses eSciDoc as common institutional repository framework. The software acts as middleware (enterprise service bus) for various software solutions and thus allows us to decouple software components from storage infrastructure and allow simultaneously access via the eSciDocs REST API. Datasync acts as deamon to synchronise and distribute files while panMetaDocs is used to upload, describe with metadata, and finally publish data. One use case for eSciDoc is the interface of the GFZ Virtual Mass Spectrometer (Virtual SIMS) to the GFZ data infrastructure. Please visit poster IN53A-1556 (Friday PM session).

Resources:
http://github.com/ulbricht/datasync : DataSync software
http://panmetadocs.sf.net : panMetaDocs project
http://github.com/escidoc : eSciDoc project

Acknowledgements

* This work is funded by DFG through the projects KOMFOR and EWIG.

HELMHOLTZ ASSOCIATION