

Search and Dissemination of Metadata

The modular design of the DataCite software and our non-invasive changes allows us to re-use the search component of the original software. Currently search in the metadata is limited to the "Dublin Core"-style fields of the DataCite schema. Modifications to the OAI-PMH provider make it possible to disseminate metadata in an additional DIF XML-schema to data portals.

What is the IGSN?

The International Geo Sample Number (IGSN) was developed by the System for Earth Sample Registration (SESA) with funding of the US National Science Foundation. By implementing a registry of samples, the IGSN aims to provide unique identifiers, that allow unambiguous identification of probed objects in written publications. This facilitates synthesis studies, re-use of data and enables to build catalogues of registered objects. Catalogues, that would assist in finding the current location or help determining the current state of a sample.

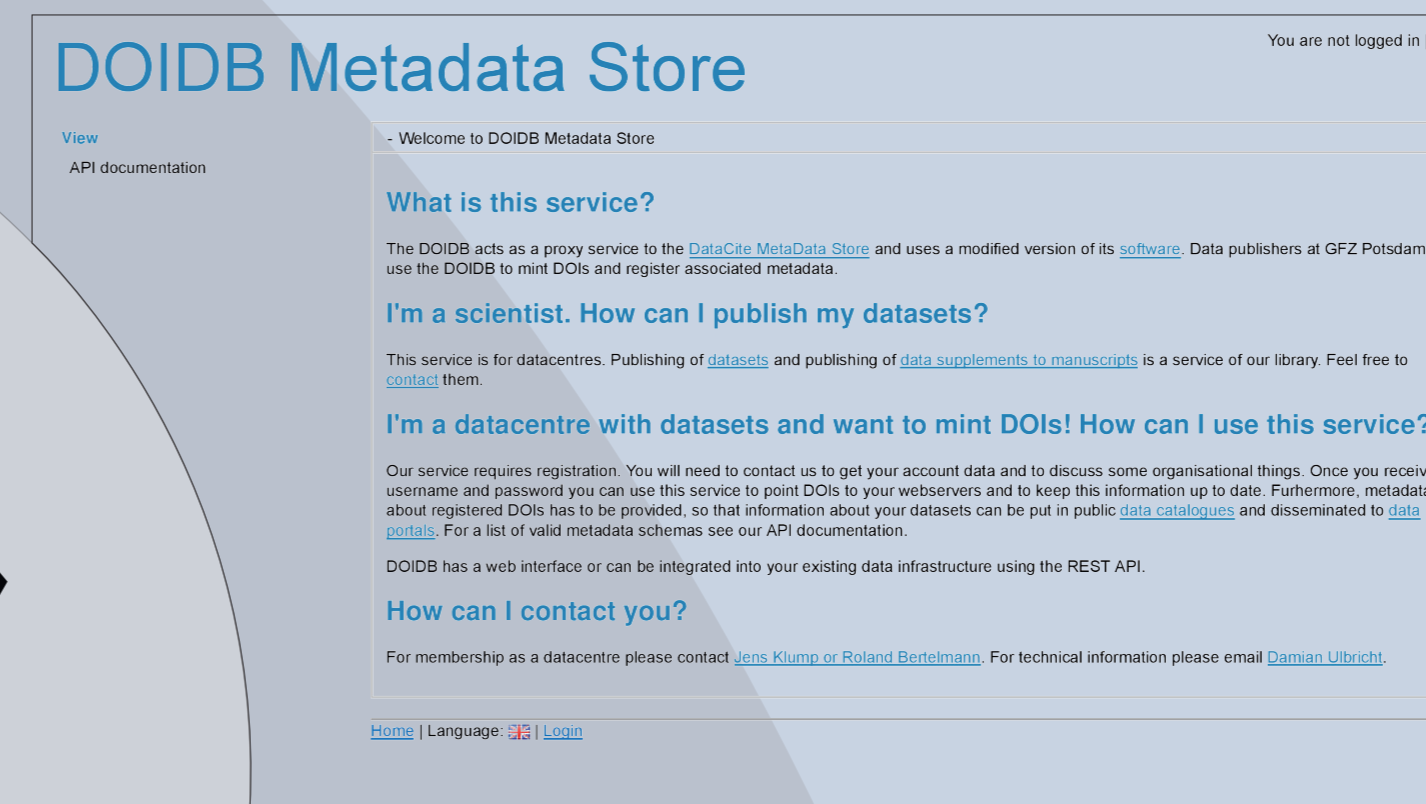
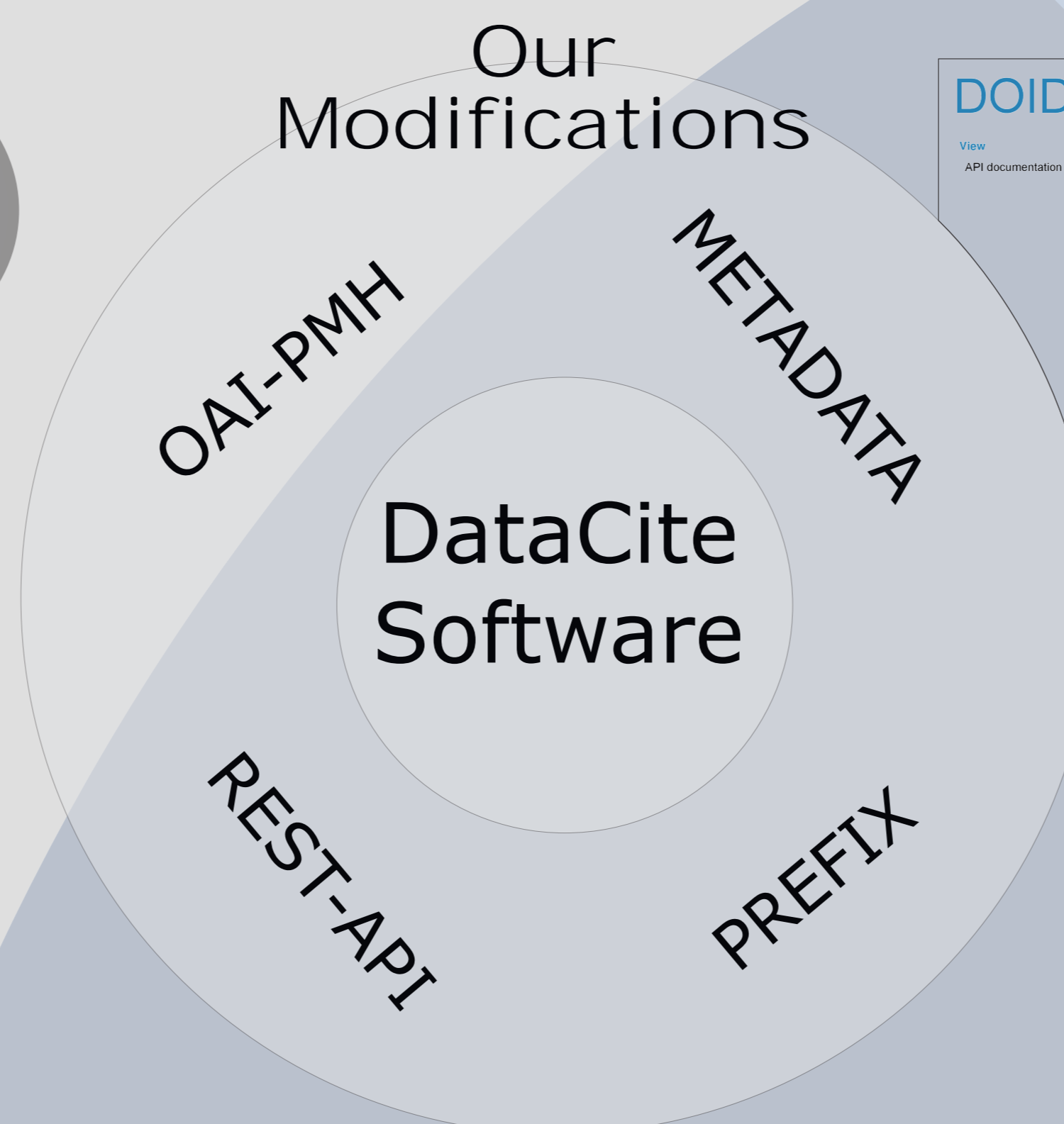
The GFZ DOI DB

The DOI DB is a key component in publication of datasets at GFZ. In combination with panMetaDocs and existing (semi-) automated systems it is used to register dataset DOIs through DataCite. Since the DataCite Metadata Store communicates directly with the Handle System and stores only metadata in a DataCite XML-format, we modified the software to allow both forwarding of REST-requests from GFZ systems to the DataCite Metadata Store and deposition of metadata in a DIF XML-format. Finally we eased restrictions in the DataCite MDS handle prefix structure to allow us to follow GFZ handle policies. This was done also in preparation of the IGSN Registry.

Albert Search Engine

GFZ Data Portal

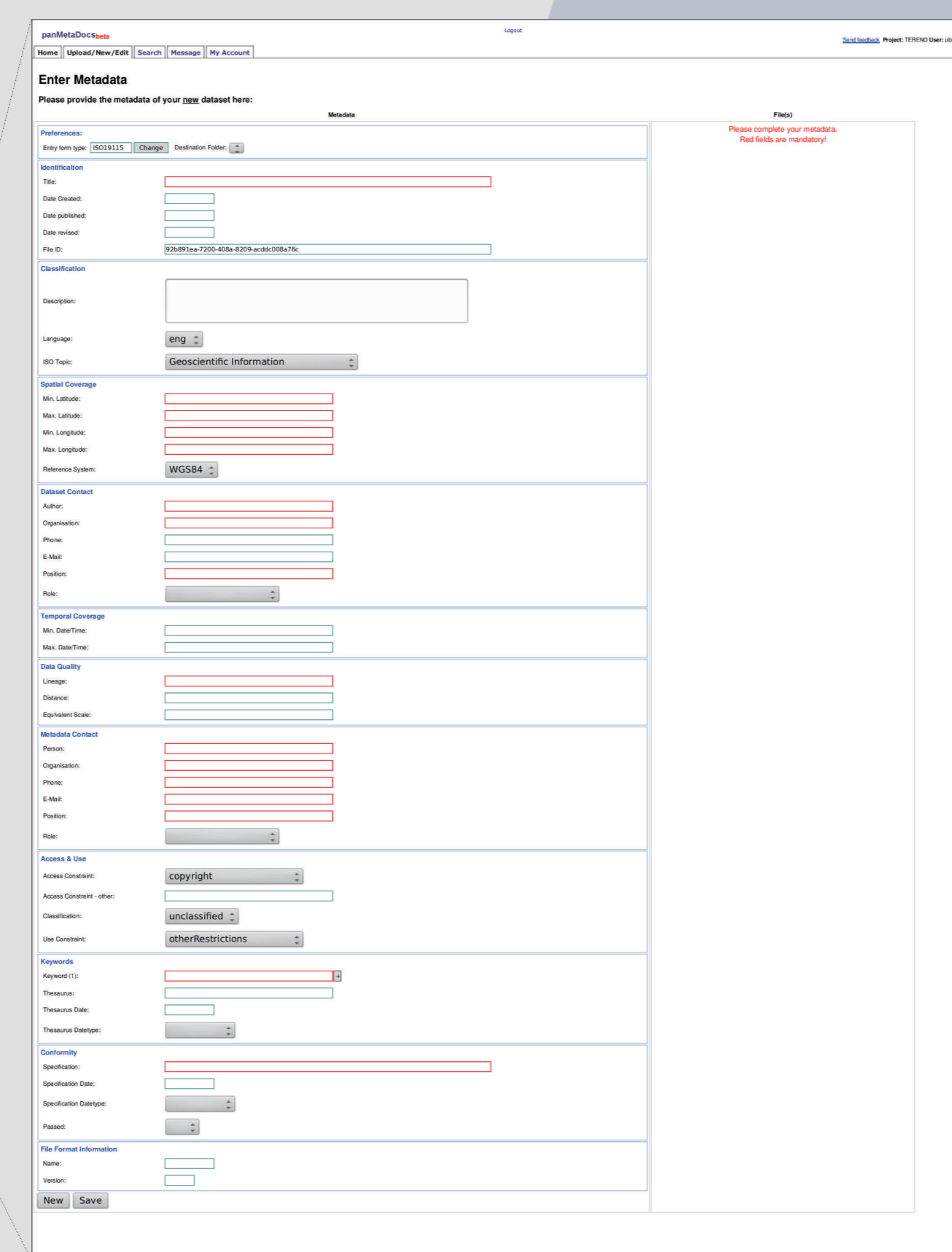
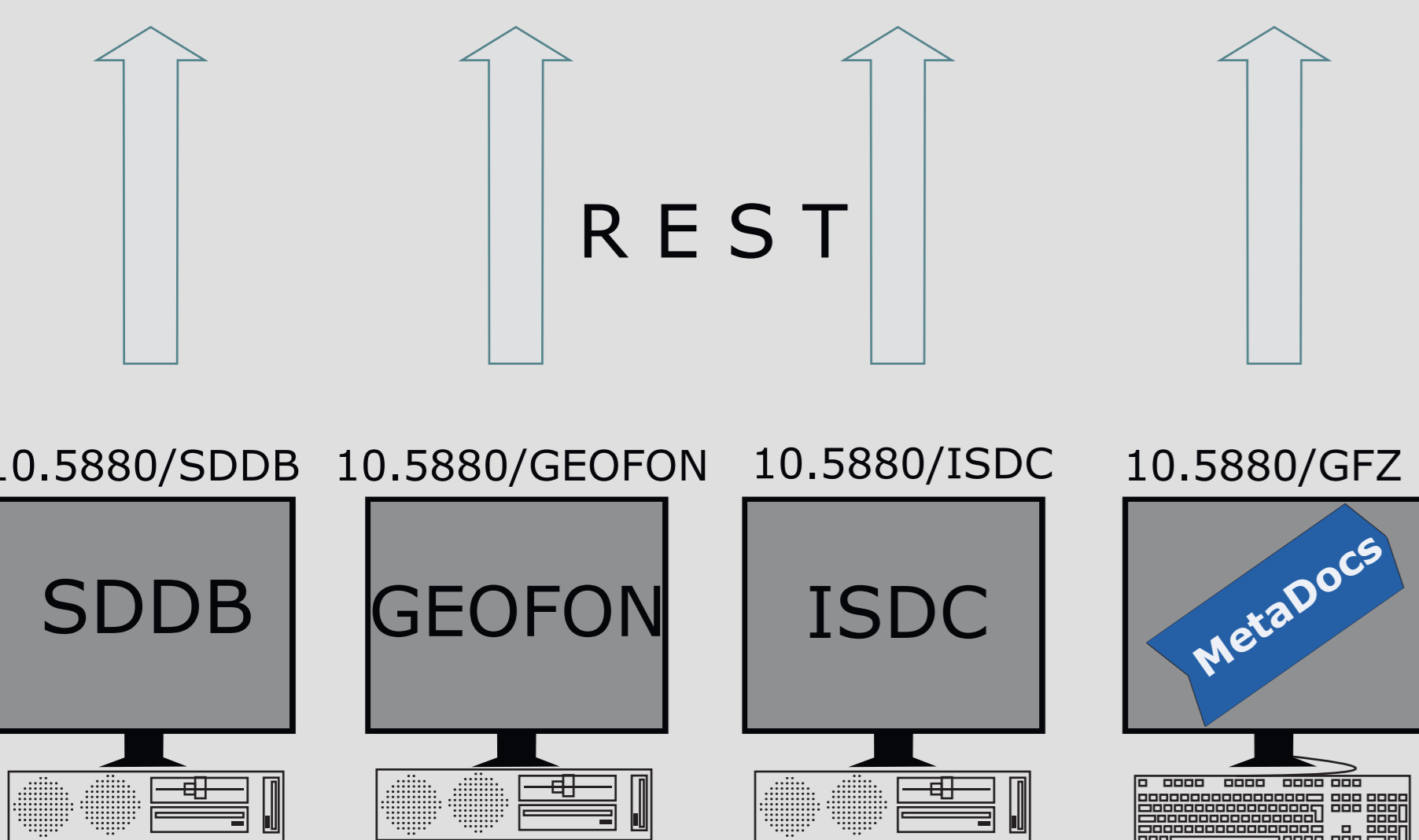
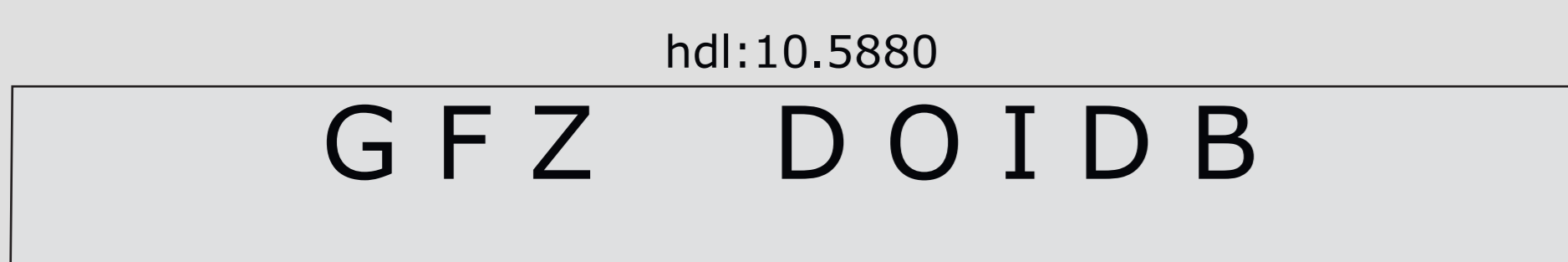
OAI-PMH



The IGSN Registry

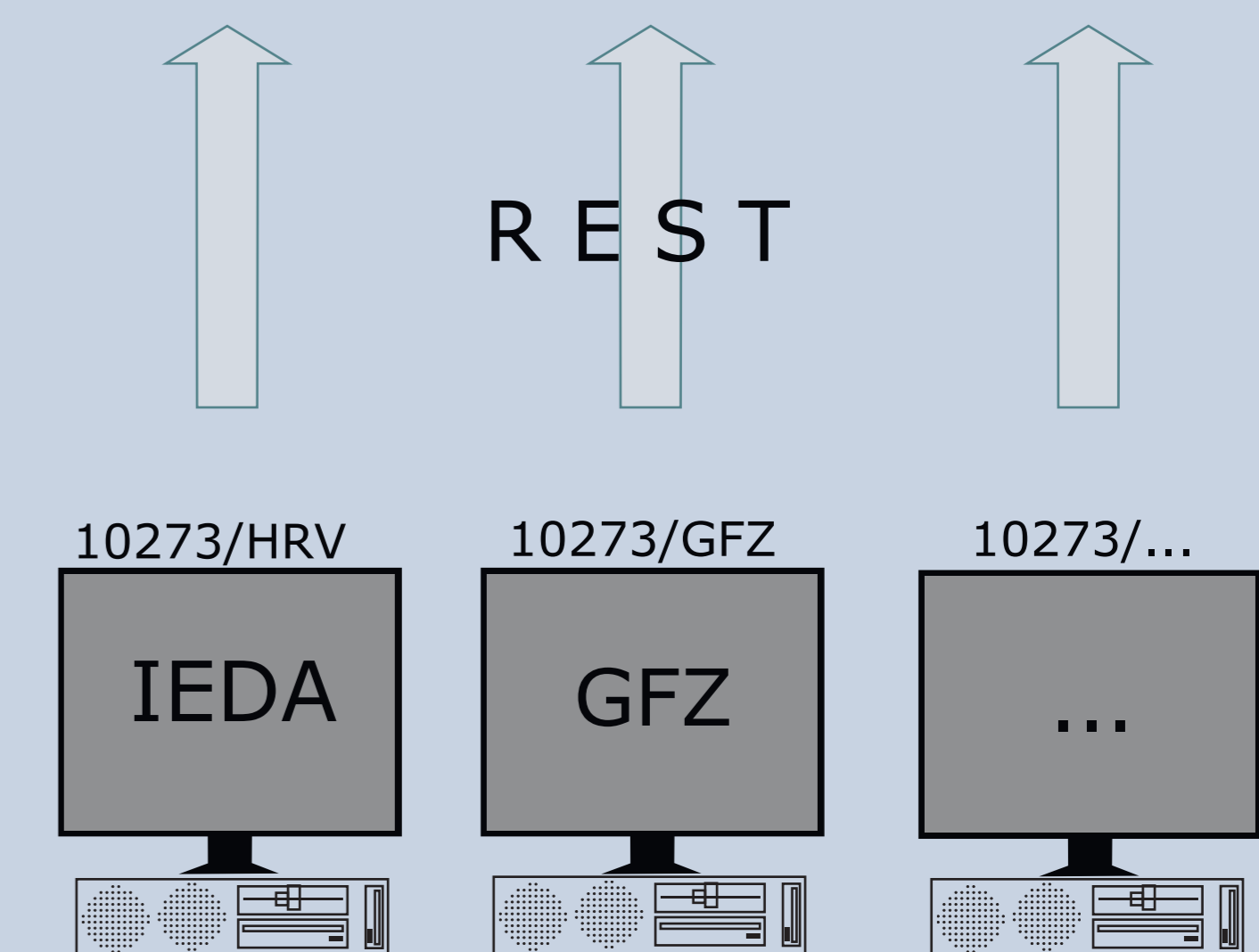
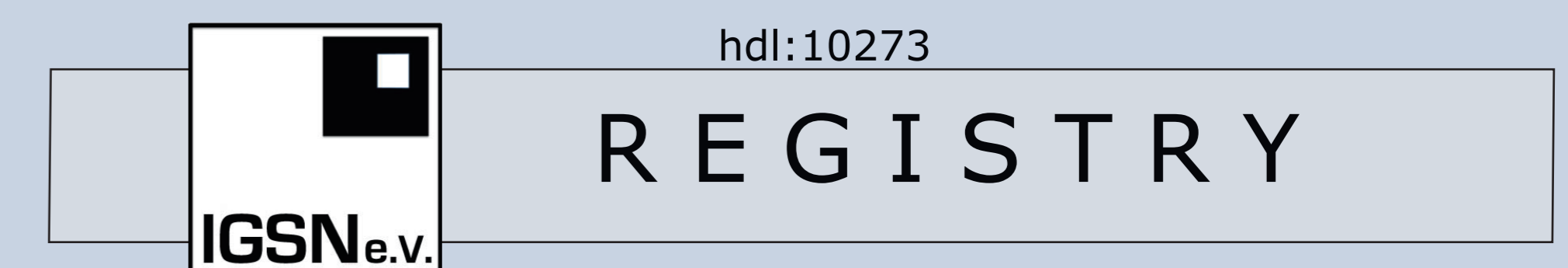
Technical the IGSN is closely related to the DOI system of persistent identifiers. Both use the Handle System to map persistent identifiers to a changing web-address and both create a policy to decide which objects are valid for registration.

The IGSN Registry is a correspondent to the DataCite Metadata Store with a different Handle Prefix namespace and with different metadata. We modified the DataCite Software to accept the type of handles and the type of metadata used for IGSN registration.



The REST-API and integration of automated systems

The REST-API of the DataCite Metadata Store is crucial, when automated systems are to be connected. For re-use as IGSN Registry we modified parts of the DataCite REST API to allow specification of IGSN during handle minting. Furthermore adjustments in the process of metadata registration were necessary to enable indication of DOI or IGSN of metadata schemas, that do not contain this information and allow correct mapping into the schema database.



Resources:

<http://github.com/datacite> : DataCite software
<http://panmetadocs.sf.net> : panMetaDocs project

Acknowledgements

* panMetaDocs is based on the source code of panMetaWorks. The source code is provided by R. Huber of **marum**
* Funding for the development of these adjustments was provided by **DFG** through the projects EWIG and KOMFOR.
* DOI DB and IGSN-Registry are based on the source code of DataCite Metadata Store