

Interoperability of institutional data management systems

Hela Mehrrens, Pina Springer, Carsten Schirnack, Claas Faber, Doris Maicher, Lisa Paglialonga, Sören Lorenz

GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany

Motivation

- Awareness of FAIR data requirements is rising
- We have come FAR but:
- Interoperability is the remaining challenge for productive systems

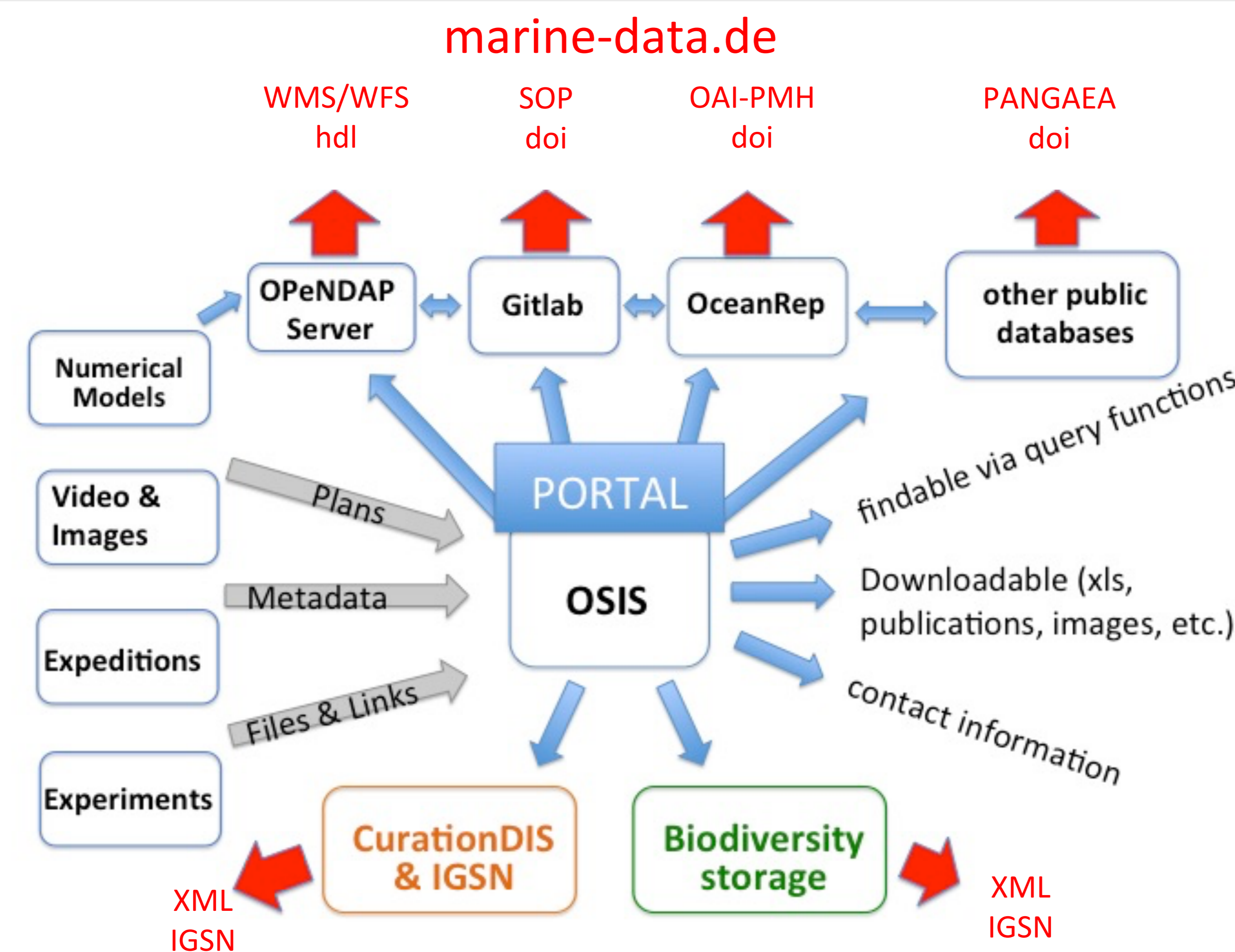
Here we present

- Description of the existing infrastructures and interfaces at GEOMAR
- Challenges to reach interoperability
- A roadmap with planned steps to go

Systems

1. the Ocean Science Information System OSIS is the central data information hub with open access metadata and internal data exchange
2. OceanRep is the institutional open access repository
3. the Liferay portal integrates services in the context of individual communities
4. the THREDDS/OPeNDAP server provides access to gridded data
5. the Gitlab service enables software development and continuous integration as well as workflow descriptions

We are also working on establishing or updating data systems for biological and geological sample management, and for data from large scale devices such as robotic underwater devices. All these systems can provide contributions to the community by a common portal (marine-data.de) via standard interfaces and protocols (marked in red).



GEOMAR digital research services

Challenges

- reach interoperability for the marine community
 - locally established workflows adapt to external requirements
 - local operational constraints need a technical redesign
1. Machine readable provision of metadata has to be implemented at an early stage of data acquisition and tools for researchers to allow them to provide accurate and relevant attributes for their data from the beginning
 2. Enhancing the use and machine actionable functionality of persistent identifiers along the data and sample collection and managing cycle
 3. Describing all existing workflows in a standardized way (SOP = standard operation procedure)
 4. Offering infrastructures for the marine community to visualize data, analyse data by providing tools to use directly on the original storage and make data usage visible via common data metrics

Roadmap

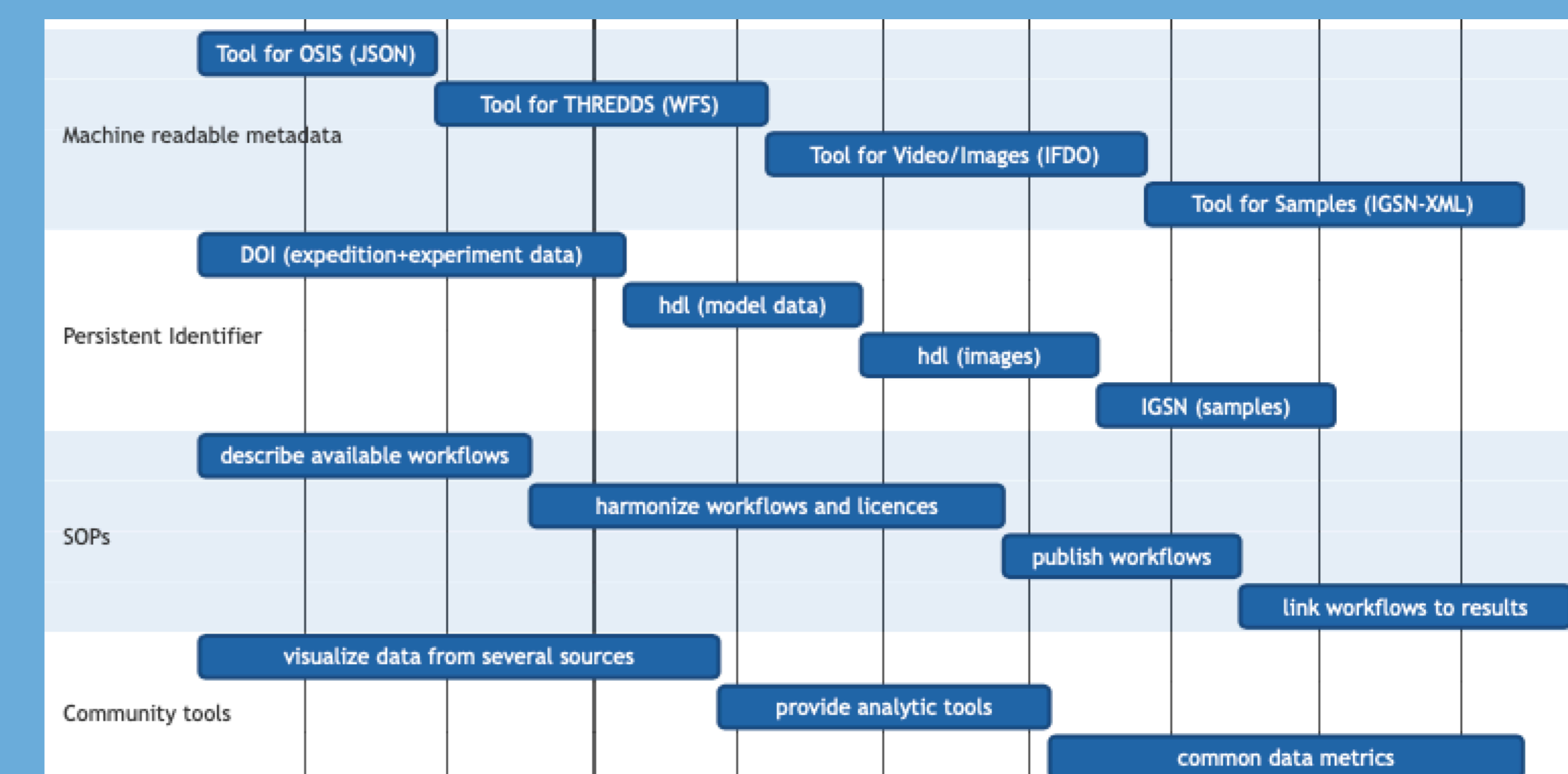
GEOMAR aims at providing a collaborative RDM platform for marine projects as a community service.

How to achieve this on the operational level is currently developed jointly with other marine research institutions in community projects, e.g. within the **DAM** (German Alliance of Marine Research) and **MareHUB**, the marine research focused centres within the Helmholtz research area Earth and Environment which is reaching out to other disciplines, for example atmosphere and solid earth and even beyond. In the next 2-5 years we are contributing with content and interfaces from our institutional systems.

We aim to provide **machine-readable metadata** for expedition data from OSIS, model data from thredds, fair digital objects for images (iFDO) and sample metadata in standard IGSN-XML. Different **persistent identifier** will be used for these cases, in line with the metadata development.

To establish documented, standardised and mostly automated workflows for research data we plan to work on standard operation procedures (SOPs), which is also a longer process, where data management, curators and research divisions at GEOMAR are cooperating.

To provide the institutional infrastructures as **community tools**, to visualize, analyze and show the impact of research data, will not only connect the marine Helmholtz centers but will make a contribution to national and international efforts making data FAIR.



Roadmap of Interoperability