


RDM: A New Focus of the Leibniz Supercomputing Centre (LRZ)

S. Hachinger, A. Götz, T. Weber, J. Munke, M. Hayek, H. Nguyen, W. Kurtz, R. García-Hernández, A. Batsaikhan, A. Frank, M. Sharikadze, J. Weismüller and D. Kranzlmüller¹

all: Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften, Garching b. München;
¹Institut für Informatik, Ludwig-Maximilians-Universität München

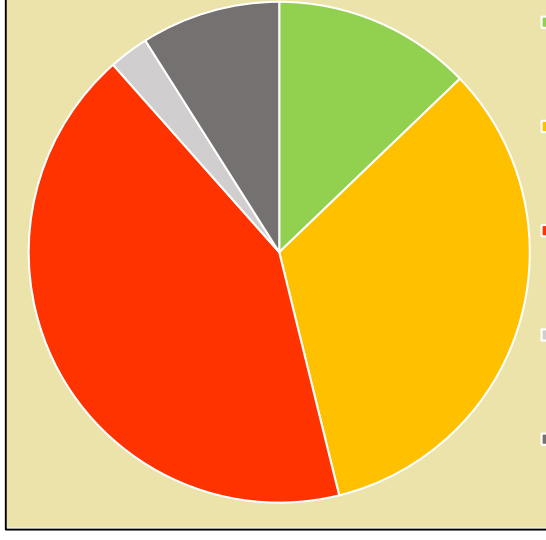
COLLABORATION & POLICY MAKING

Activities in RDA



RDARI IG (Research Data Architectures in Research Institutions)

- IG for getting an overview over architectures & harmonizing them.
- First output: Governance (and infrastructure) survey (by J.A.J. Wilson, UCL, et al.); Results / Data: <https://doi.org/10.5522/04/10283540>




RIDA (Repository Interfaces for Data Analytics) BoF

- Launched by G. Greene (NIST), T. Weber / A. Götz (LRZ) and colleagues.
- BoF in P14 – continuation planned.

rdmuc

Meetings to harmonize RDM Activities in the Munich Area

- Includes University Libraries, LMU ITG (Digital Humanities), Bavarian State Library, LRZ and MPCDF.
- First output: recommendation on exact use of DataCite metadata scheme (“how to fill it in”) by S. Kümmer, S. Lücke, J. Schulz, M. Spenger, T. Weber: <https://zenodo.org/record/3559800>; <http://www.kit.gwi.uni-muenchen.de/?p=51272>
- Further Topics include:
 - Usage of PIDs,
 - Data Management Plans,
 - Making data findable (OAI-PMH, alternatives).



Support for NFDI Consortia

With our team at LRZ, we are participant in several initiatives and offer, e.g.:

- Consulting on RDM and respective services,
- Support with making data findable via search portals for research data (GeRDI, B2FIND, BASE, Google Dataset Search etc.),
- Unified RDM approaches (see below) to integrate domain-specific repositories and legacy “data silos”.

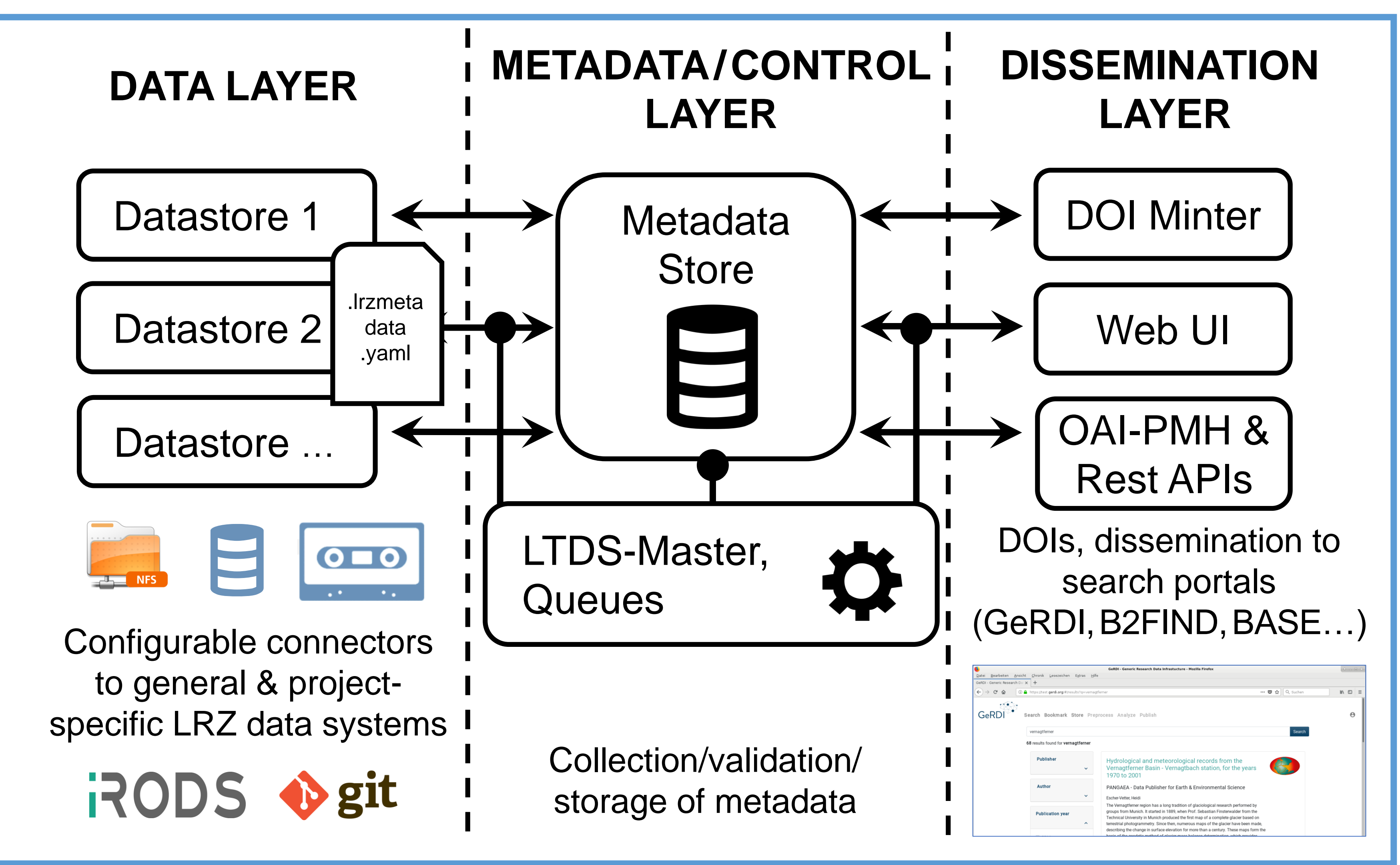
SYSTEMS FOR FAIR DATA

“Let The Data Sing” (LTDS): Lightweight RDM for Big Data – under construction –

- Microservice-based, open RDM framework to make “data silos” FAIR
- Intended for very large datasets at LRZ
- General purpose (largely domain agnostic)
- Low-threshold: DataCite metadata as a minimum requirement

OBJECTIVES – MAKING DATA FAIR:


- Add metadata to data at LRZ
- Get DOIs for data at LRZ
- Make data findable via research data search portals



SUPPORT FOR PROJECTS & RESEARCHERS

RDM in Research Projects with LRZ Participation


LRZ participates in various projects which require research data management (focused on domain-science or on data/compute aspects). Where possible, we aim to co-develop LTDS with these projects. The projects can thus make use of LTDS, and the framework can be tested and grown. Also, in domain-driven projects, novel services such as data-analytics portals or computing-on-demand facilities are developed.



- PI: Prof. R. Ludwig, LMU
- Web: www.climex-project.org
- Funding: Bavarian Ministry of Environment & Consumer Protection

PROJECT & RDM:


- One of the largest regional climate-simulation datasets (Bavaria/Quebec)
- Indexing, annotation, publication of several 100TB.



- PI: Prof. W. Mauser, LMU
- Web: viwa.geographie-muenchen.de
- Funding: German Federal Ministry of Education & Research

PROJECT & RDM:


- >300 TB of hydrology simulation data
- Metadata & portal for downloading geo-selected data



- PI: Prof. M. Bittner, DLR/UAU
- Web: www.alpendac.eu
- Funding: Bavarian Ministry of Environment & Consumer Protection

PROJECT & RDM:

- Building Platform with On-Demand Computing & Data Analytics facilities
- Use LTDS e.g. for DOI minting



- PI: Dr. J. Martinovic, IT4I (CZ)
- Web: lexis-project.eu
- Funding: European Commission

PROJECT & RDM:

- Optimized execution/control of Cloud-/HPC-Workflows (HPC-Cloud-Convergence)
- LRZ task:iRODS/EUDAT-based Distributed Data Infrastructure
- LTDS useful for data dissem.