

# ● Forschungsdatenmanagement für Daten aus High Performance Measurement and Computing (HPMC)

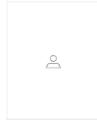
Best-Practice Beispiele und Anwendungen

22.07.2021

## Beteiligte



Prof. Dr.-Ing. Christian Stemmer, Lehrstuhl für Aerodynamik und Strömungsmechanik TUM, Forschungsprojekt NFDI4Ing



Katja Kessler, Universitätsbibliothek TUM, Beratungszentrum Forschungsdatenmanagement (eRIC)



Benjamin Farnbacher, Lehrstuhl für Aerodynamik und Strömungsmechanik TUM, Forschungsprojekt NFDI4Ing



Nils Hoppe, Lehrstuhl für Aerodynamik und Strömungsmechanik TUM, Forschungsprojekt NFDI4Ing



Stephan Peinkofer, Data and Storage Division, LRZ



Dr. Stephan Hachinger, Research Data Management, LRZ

## Agenda

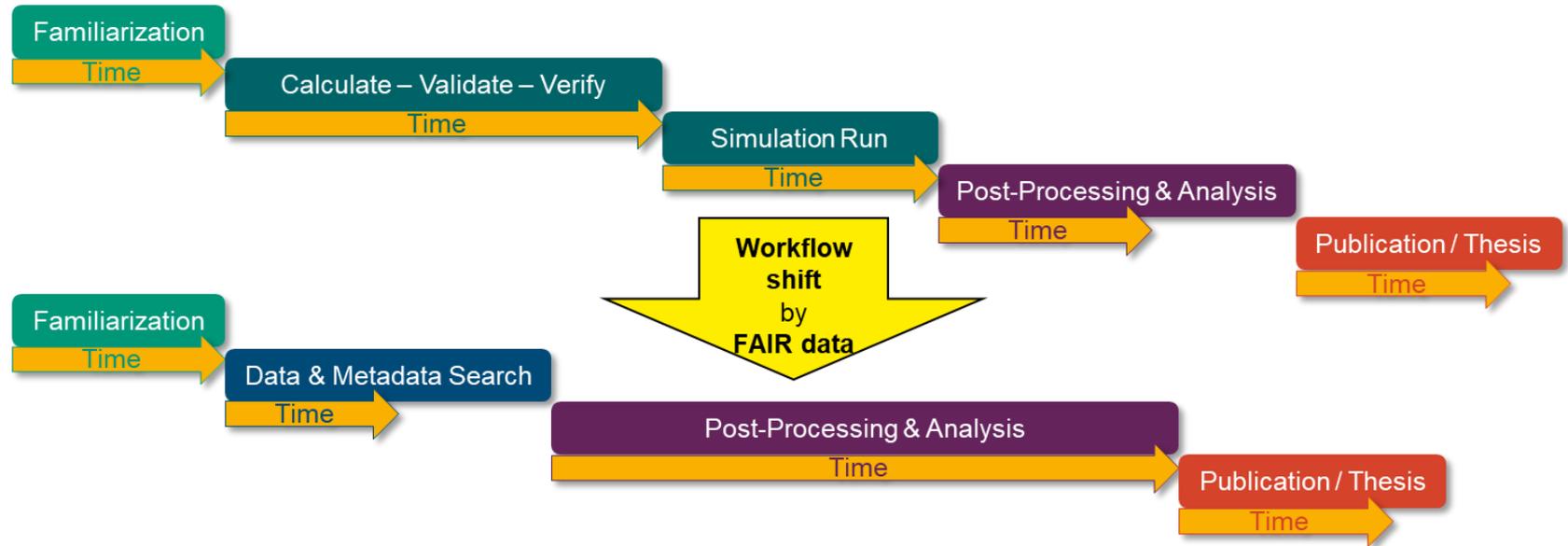
- Forschungsdatenmanagement: Relevanz
- Forschungsdaten
- Was ist Forschungsdatenmanagement?
- Datenmanagementpläne
- FAIR-Prinzipien und deren Umsetzung
- Repositorien und Identifier
  
- mediaTUM – Basics & Features
- Lizenzen
- TUM Workbench
  
- Beispiel für Metadaten und Ontologien
- Archivierung und Veröffentlichung von Daten
- Tools zum Forschungsdatenmanagement
  
- Nationale Forschungsdateninfrastruktur (NFDI)
- NFDI für Ingenieurwissenschaften (NFDI4ing)
- Archetyp DORIS: Forschungsdaten aus HPMC
  
- Fokus: HPMC-Daten
- DSS – Data Science Storage
- globus
- FAIR Supercomputing Data at LRZ

## Warum ist Forschungsdatenmanagement relevant?

### Extrinsische Motivation

- Wissenschaftliche Integrität und Einhaltung der DFG Leitlinien zur Sicherung guter wissenschaftlicher Praxis sowie interner Leitlinien
- Anforderungen der Projektträger, Koppelung der Förderung an nachhaltiges FDM
- Thema von zunehmender politischer Relevanz und Förderung (siehe z.B. Datenstrategie der Bundesregierung, NFDI)
  
- Institutsinterne Datenverfügbarkeit unabhängig von Personalfluktuatation
- Auch: Daten aus Industriekooperationen für nicht öffentliche Daten
  - Daten- und Zugangsmanagement via FDM
  
- Sekundärforschung / Begleitforschung
- Neue Möglichkeiten wie z.B. „Datenpromotion“

## Warum ist Forschungsdatenmanagement relevant?

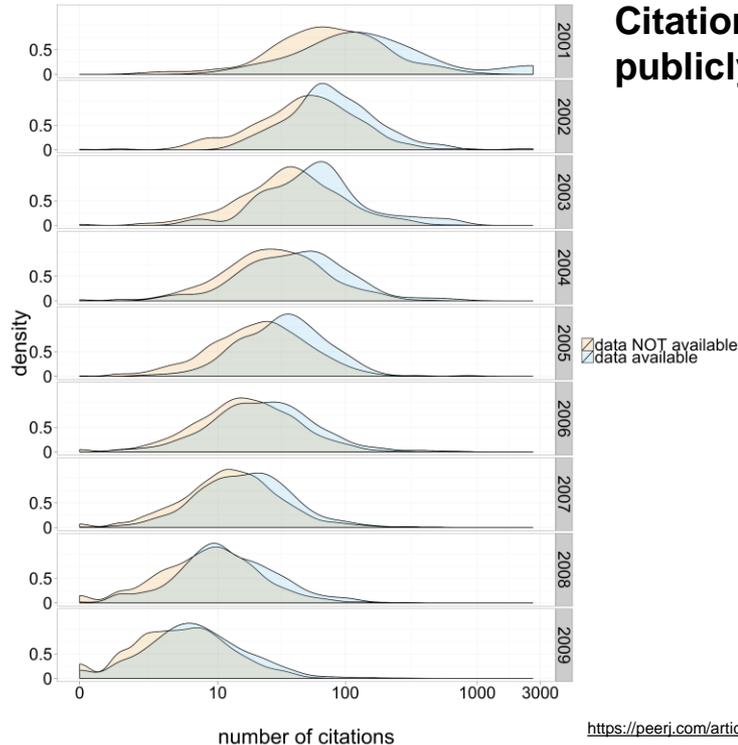


## Warum ist Forschungsdatenmanagement relevant?

### Intrinsische Motivation

- Wiederverwendung oder Weitergabe eigener (alter) Ergebnisse
- Wissenschaftlicher Ruf & Transparenz
- Erhöhte Chancen auf Kollaborationen und Förderungen
- „Standing on the shoulder of giants“
  - Neue Erkenntnisse durch (Meta-)Datenanalyse
  - Verifizierung und Validierung eigener Modelle durch fremde Daten
  - Verkürzung der Publikationszeit durch entfallene Datengenerierung
- Vision / langfristige Perspektive: Datenpublikation (peer reviewed?)
- Erhöhung der Zitationen durch Veröffentlichung von Forschungsdaten (Publikationen 2013, 2016, 2020)

## Warum ist Forschungsdatenmanagement relevant?



**Citation density for papers with and without publicly available data (Piwowar & Vision 2013)**

## Warum ist Forschungsdatenmanagement relevant?

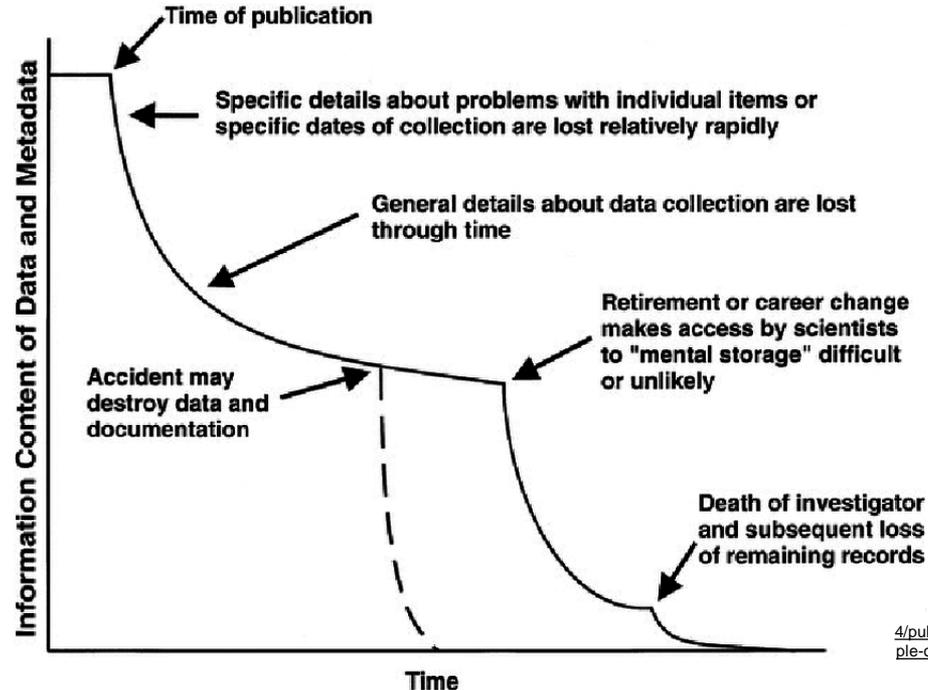
### Sharing data increases citations (Drachen et al. 2016)

Journal papers published in 2010	Astro-physical Journal (ApJ)	Astronomy and Astrophysics (A&A)	Astronomical Journal (AJ)
<b>All</b>			
# of papers	2501	1918	388
# of citations	74,663	40,829	9465
Mean citations/paper	29.9	21.3	24.4
<b>Datalink papers</b>			
# of papers	794	875	174
# of citations	27,936	22,308	4754
Mean citations/paper	35.2	25.5	27.3
<b>No datalink papers</b>			
# of papers	1707	1043	214
# of citations	46,727	18,521	4711
Mean citations/paper	27.4	17.8	22.0

<https://www.liberquarterly.eu/article/10.18352/lq.10149/> (04.05.2021)

## Warum ist Forschungsdatenmanagement relevant?

### The loss of information about data over time (Michener et al. 1997)



<https://www.researchgate.net/profile/Adam-Wilson-4/publication/255571027/figure/fig4/AS:297949465726983@1448048096877/Example-of-the-normal-degradation-in-information-content-associated-with-data-and.png>  
(04.05.2021)

## Was sind Forschungsdaten?

- Daten, die während eines Forschungsprozesses entstehen oder dessen Ergebnis sind
- Alles und nichts, je nach Perspektive
  - Bibliotheken: jeder Notizzettel etc.
  - Projektleitung: nur final publizierte Ergebnisse
  - Erweiterung: Daten, die für die ursächliche Forschungsfrage uninteressant sind, können für Sekundärforschung relevant werden (z.B. Energieverbrauch oder Temperatur im HPC bei Simulationen)

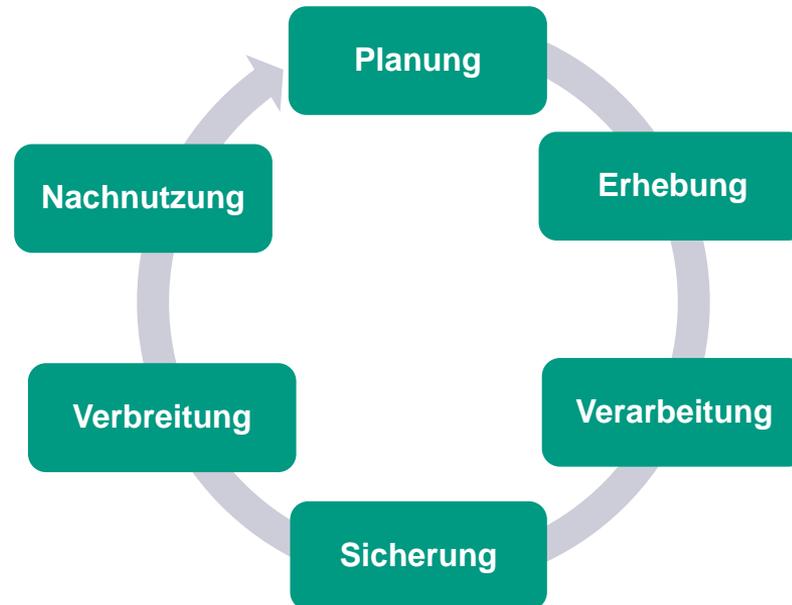
also das

- Organisieren,
- Speichern,
- Wiederverwenden,
- Teilen

von Forschungsdaten

## Was ist Forschungsdatenmanagement (FDM)?

- Maßnahmen zur Sicherung und langfristigen Nutzbarkeit von Forschungsdaten
- Prozess entlang der Stationen des Forschungsdaten-Lebenszyklus



## Wie funktioniert Forschungsdatenmanagement?

### Datenmanagementplan (DMP)

Vorgaben:

- DFG: <https://www.ub.tum.de/dmp-vorgaben-dfg>
- EU: <https://www.ub.tum.de/dmp-vorgaben-eu-kommission>
- TUM: <https://www.it.tum.de/projekte/forschungsdaten-management>

Inhalte:

- Administrative Informationen (Projekt, Förderung, Verantwortliche etc.)
- Beschreibung relevanter Daten (Typ, Format, Größe, Lesbarkeit etc.)
- Datenorganisation (Ordnerstruktur, Benennung, Backup, Zugriff etc.)
- Ethische und rechtliche Belange (Datenschutz, Einwilligung, Anonymisierung etc.)
- Angaben zur Langzeitarchivierung und Veröffentlichung (Repositorium, Nutzungslizenzen etc.)
  
- DMP sind dynamische Dokumente und sollten ca. halbjährlich aktualisiert werden

## Wie funktioniert Forschungsdatenmanagement?

### Datenmanagementplan (DMP): Tools & Templates

- TUM Workbench: <https://workbench.ub.tum.de>
  - TUM generic template for planning your research data management
  - templates for applications to research funding agencies (e.g. Horizon 2020)
- RDMO: <https://rdmorganiser.github.io/en>
  - German tool developed in a DFG project,
  - discipline-specific templates available
- DMPonline: <https://dmponline.dcc.ac.uk>
  - British tool, suitable for applications to British funding agencies and EU
- DMPTool: <https://dmptool.org>
  - US-American tool, suitable for applications to US-American funding agencies
- Short DMP Checklist by the DDC:
  - [www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP\\_Checklist\\_2013.pdf](http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf)

# Wie funktioniert Forschungsdatenmanagement?

## Grundsatz

Einhaltung der FAIR-Prinzipien für ein optimales FDM:

- Findable
- Accessible
- Interoperable
- Reusable

Wilkinson et al.: The FAIR Guiding Principles for scientific data management and stewardship;  
*Scientific Data* volume 3, Article number: 160018 (2016); <https://doi.org/10.1038/sdata.2016.18>

## Wie funktioniert Forschungsdatenmanagement?

### Das F in FAIR (Findable / Auffindbarkeit)

- Daten erhalten persistente Identifikatoren (z.B. DOI)
  
- Daten sollen durch **Metadaten** in lesbarer Form beschrieben sein
  - Beschreibend, strukturell, administrativ etc.
  - Metadaten durch Suchmöglichkeit indizierbar (z.B. via Repositoryum)
  - Entscheidend zur Auffindbarkeit in Repositorien & Datenbanken sowie zur Indizierung durch Suchmaschinen
  - Ontologien kategorisieren Metadaten in einheitlicher Hierarchie und Terminologie

# Wie funktioniert Forschungsdatenmanagement?

## Das A in FAIR (Accessible / Zugänglichkeit)

### Voraussetzung

- Speicherung der Daten durch Forschende
  - Infrastruktur zur Langzeitarchivierung
- Zugriffsmöglichkeiten auf Datensatz
  - Infrastruktur für Rechtemanagement
  - Interner Zugriff vs. Zugriff für Community vs. Open Access
  - Datenschutz & Urheberrecht müssen eingehalten werden
- Barrierefreiheit: Zugriffsmöglichkeit möglichst unabhängig von Hard- und Software (Endgerät, Betriebssystem, Browser etc.)
- Offene Standardprotokolle zum Zugriff für Berechtigte
  
- MediaTUM: Repositorium & Datenbank für TUM-Forschungsdaten
- Vision für HPMC: Speicherung von Metadaten in Repositorium, Speicherung von Forschungsdaten an Rechenzentrum

## Wie funktioniert Forschungsdatenmanagement?

Empfohlene Formate: open source, einfach, häufig genutzt

Inhalt	Empfohlene Formate
Text	PDF/A <i>ohne Formatierung:</i> TXT <i>für Editierbarkeit:</i> ODT, RTF, HTML <i>mit Formeln:</i> LaTeX (TEX)
Tabelle	CSV, TSV <i>numerische Daten:</i> HDF5
Grafik	<i>Rastergrafik:</i> PNG, TIFF <i>Vektorgrafik:</i> SVG, EPS
Multimedia	<i>Container:</i> MKV, WebM, OGG <i>Video-Codec:</i> AV1, VP9 <i>Audio-Codec:</i> FLAC, WAV, Vorbis, Opus
Relationale Datenbank	SIARD, Dump, XML
allg. strukturiere Daten	XML, JSON, YAML

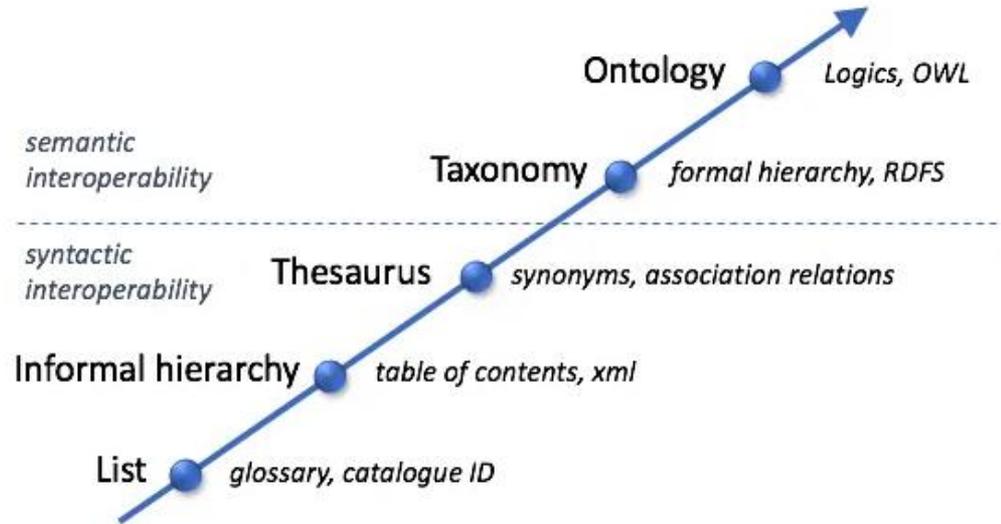
## Wie funktioniert Forschungsdatenmanagement?

### Das I in FAIR (Interoperable / Kompatibilität)

- (Meta-)Daten orientieren sich an verbreiteten Standards
  - Einheitliche Terminologie, z.B.
    - Taxonomie der Lebewesen (...Ordnung > Familie > Gattung > Art)
    - Geoinformation: ISO 19115, Klimaforschung: CERA Data Model,  
Materialwissenschaften: EMMO

# Wie funktioniert Forschungsdatenmanagement?

## Das I in FAIR (Interoperable / Kompatibilität)



[https://i1.wp.com/emmc.info/wp-content/uploads/2019/05/alternatives\\_2.png?w=600&ssl=1](https://i1.wp.com/emmc.info/wp-content/uploads/2019/05/alternatives_2.png?w=600&ssl=1) (04.05.2021)

## Wie funktioniert Forschungsdatenmanagement?

### Das I in FAIR (Interoperable / Kompatibilität)

- Kompatibilitätshindernisse minimieren, z.B.
  - keine Dateinamen mit Sonderzeichen, Leerzeichen, zu vielen Zeichen etc.
  - keine betriebssystemspezifischen Formate/Plug-Ins etc.
  
- Eigene Daten(-formate) sind transparent und kompatibel abzubilden, z.B. durch
  - Bereitstellung einer Dokumentation
  - Bereitstellung von Software zur Interaktion
  - Bereitstellung der ursprünglichen Rechnerumgebung (Virtualization)

## Wie funktioniert Forschungsdatenmanagement?

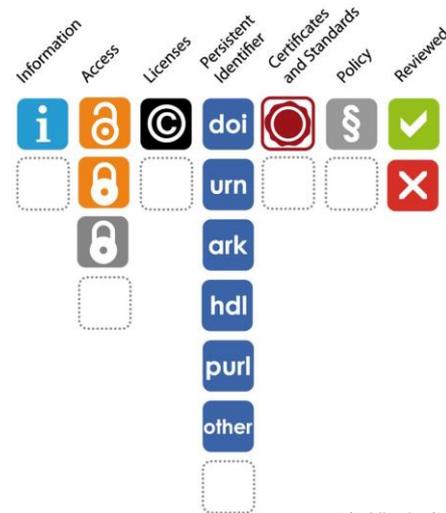
### Das R in FAIR (Reusable / Nachnutzbarkeit)

- Wenn **F**, **A**, **I** eingehalten sind, ist **R** meist erfüllt, dazu
  - Meta(daten) erfüllen die Standards der Community
  - Meta(daten) benötigen eine eindeutige und freie Nutzungslizenz
  - Meta(daten) beschreiben auch die Datengenese/-Provenienz
  
- Achtung: für die Wiederverwertung von HPMC-Daten ist meist **Rechenzeit** an HPC-Zentren notwendig

# Wie publiziere ich meine (TUM-)Forschungsdaten?

## Datenrepositorien

- TUM: <https://mediatum.ub.tum.de/>
- disziplinspezifische Repositorien: <https://www.re3data.org/>
- Filter:



## Wie publiziere ich meine (TUM-)Forschungsdaten?

### Datenrepositorien: DOI (Digital Object Identifier)

- Eindeutiger und dauerhafter Identifikator für Objekte, Events etc.
- Häufigster „persistent identifier“ in Forschungscommunity
- Genormt nach ISO 26324
- Macht veröffentlichte Daten zitierbar
- Bei vergebener DOI werden Forschungsdaten „kalt“
  - >> keine Veränderungen mehr möglich
  - >> Ausnahme: Concept DOI (angeboten von mediaTUM, Zenodo, ...)  
vgl. <http://help.zenodo.org/#versioning>,  
Beispiel: <https://zenodo.org/record/1477956#.W-mC6Gf1OCQ>

## Wie publiziere ich meine (TUM-)Forschungsdaten?

### mediaTUM - Basics

- <https://mediatum.ub.tum.de/>
- Medien- und Publikationsserver der TUM
  - Publizieren & archivieren von wissenschaftlichen Arbeiten und Forschungsdaten
  - Einbindung in Publikationslisten auf Websites
  - Veröffentlichung der Dissertation: <https://www.ub.tum.de/ediss>

Vorgehen zur Veröffentlichung von Forschungsdaten:

1. Ausfüllen des Publikationsantrags mit Informationen zu Daten & Metadaten
2. Bibliothek erstellt Metadatensatz & Verzeichnis für Upload durch Forschende
3. Forschende erteilen Freigabe >> Bibliothek entzieht Schreibrechte  
>> Datensatz wird veröffentlicht

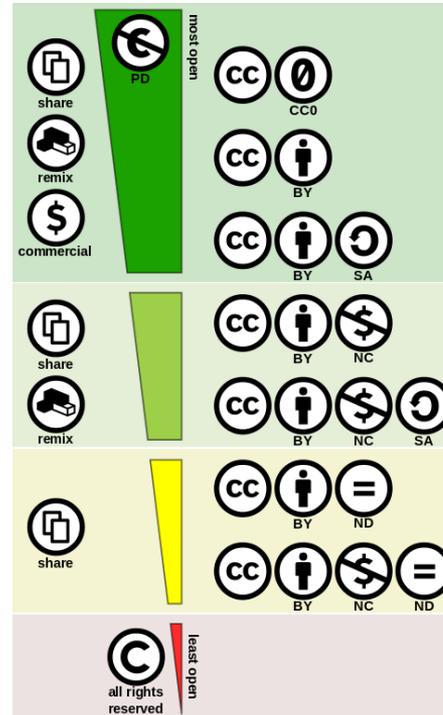
Weitere Informationen: <https://www.ub.tum.de/forschungsdaten-publizieren>

## Wie publiziere ich meine (TUM-)Forschungsdaten?

### mediaTUM - Features

- **Peer-Review:** auf Antrag anonymes Login für Zugriff durch Gutachter (vor Veröffentlichung)
- **DOI und Concept DOI:** Bibliothek kann permanente DOI (Digital Object Identifier) oder Concept DOI (für dynamische Datensätze) vergeben
- Große Datenmengen möglich (einige TB)
- Auswahl von Lizenzen (siehe: [Handreichung für die Veröffentlichung wissenschaftlicher Daten und Software](#))
  
- Beispiele für publizierte Forschungsdaten:
  - [https://mediatum.ub.tum.de/1459250?show\\_id=1578279](https://mediatum.ub.tum.de/1459250?show_id=1578279)
  - <https://mediatum.ub.tum.de/1452571>

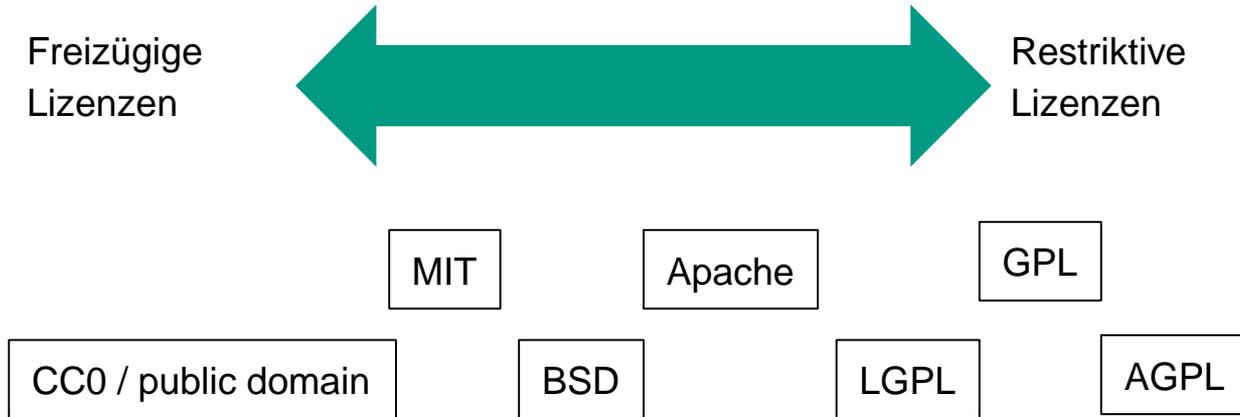
## Creative Commons Lizenzen



CC BY 4.0

[https://creativecommons.org/  
about/downloads/](https://creativecommons.org/about/downloads/)

## Software-Lizenzen



Weitere Informationen:

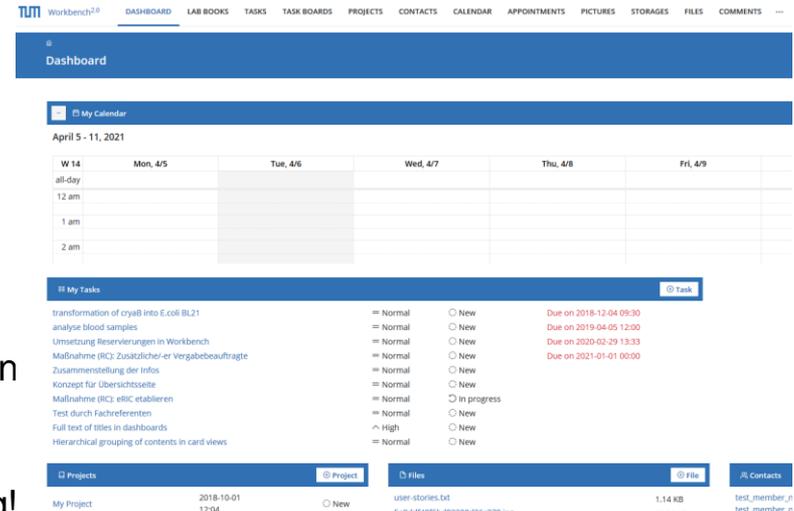
<https://choosealicense.com>

<https://www.ifross.org/lizenz-center>

## TUM Workbench

<https://workbench.ub.tum.de>, siehe auch [www.ub.tum.de/workbench](http://www.ub.tum.de/workbench)

- Virtuelle Forschungsumgebung, entwickelt durch die Universitätsbibliothek
- Modularer Aufbau:
  - Elektronisches Laborbuch
  - Projektmanagement
  - Datenmanagementpläne ...
- Features:
  - Versionskontrolle
  - Metadaten und Suchfunktionen
  - Rechtemanagement ...
- Einloggen mit der TUM-Kennung!



TUM Workbench 2.0

DASHBOARD LAB BOOKS TASKS TASK BOARDS PROJECTS CONTACTS CALENDAR APPOINTMENTS PICTURES STORAGES FILES COMMENTS ...

Dashboard

My Calendar

April 5 - 11, 2021

W 14	Mon, 4/5	Tue, 4/6	Wed, 4/7	Thu, 4/8	Fri, 4/9
all-day					
12 am					
1 am					
2 am					

My Tasks

Task	Status	Due Date
transformation of cryaB into E.coli BL21	Normal	Due on 2018-12-04 09:30
analyse blood samples	Normal	Due on 2019-04-05 12:00
Umsetzung Reservierungen in Workbench	Normal	Due on 2020-02-29 13:33
Maßnahme (RC): Zusätzlicher Vergabebeauftragte	Normal	Due on 2021-01-01 00:00
Zusammenstellung der Infos	Normal	
Konzepte für Übersichtsseite	Normal	
Maßnahme (RC): sllnc etablieren	Normal	In progress
Test durch Fachreferenten	Normal	New
Full text of titles in dashboards	High	New
Hierarchical grouping of contents in card views	Normal	New

Projects

Project	Created	Status	File	Size	Contacts
My Project	2018-10-01 17:04	New	user-stories.txt	1.14 KB	test_member_# test_member_#

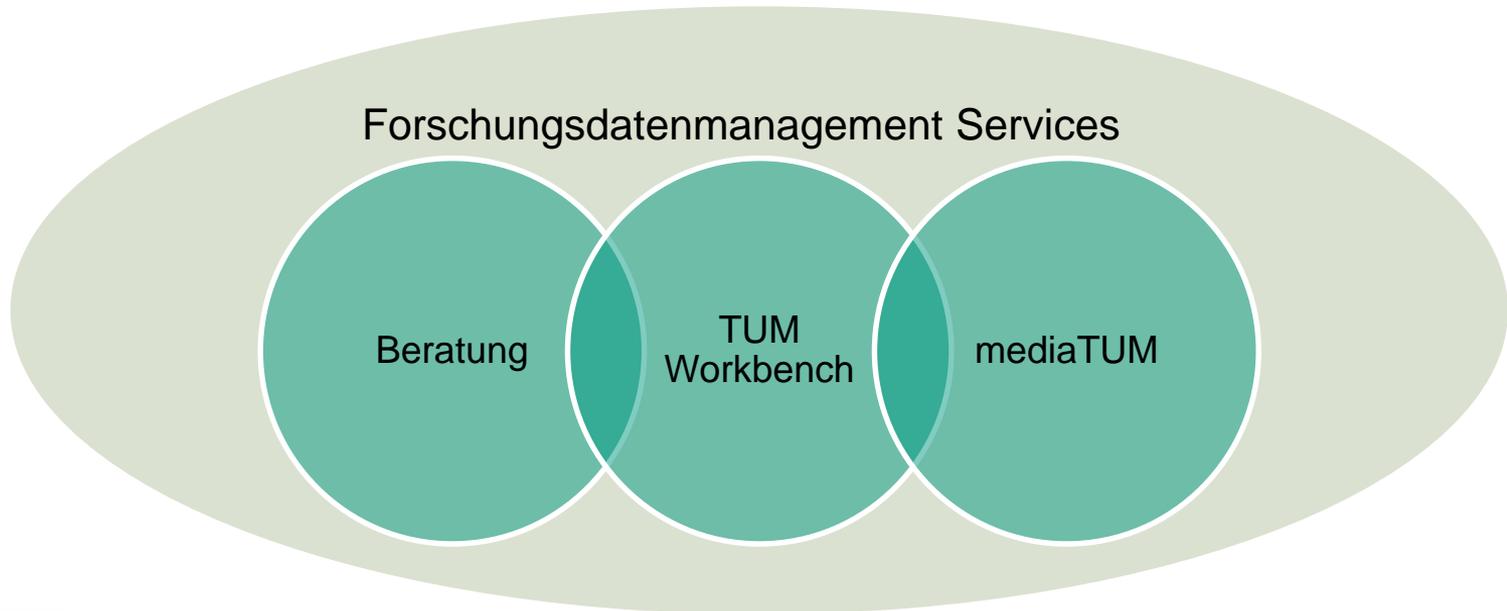
## Unterstützung der TUM Universitätsbibliothek

[eric@ub.tum.de](mailto:eric@ub.tum.de)

[www.ub.tum.de/forschungsdaten](http://www.ub.tum.de/forschungsdaten)

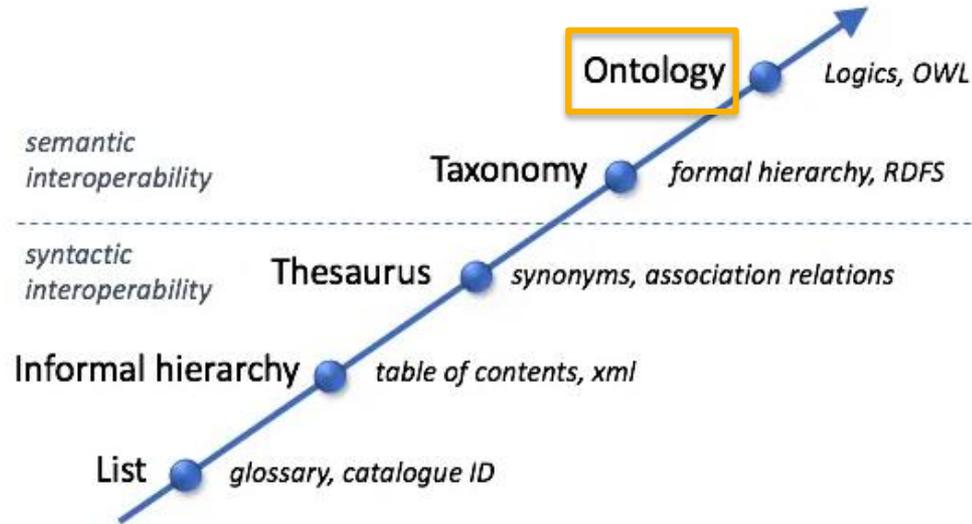


eResearch - Infrastructure and Communication



# Wie funktioniert Forschungsdatenmanagement?

## Das I in FAIR (Interoperable / Kompatibilität)



[https://i1.wp.com/emmc.info/wp-content/uploads/2019/05/alternatives\\_2.png?w=600&ssl=1](https://i1.wp.com/emmc.info/wp-content/uploads/2019/05/alternatives_2.png?w=600&ssl=1) (04.05.2021)

## Beispiel: Eintrag von Metadaten in eine Ontologie

### Roboterarm öffnet Kronkorken

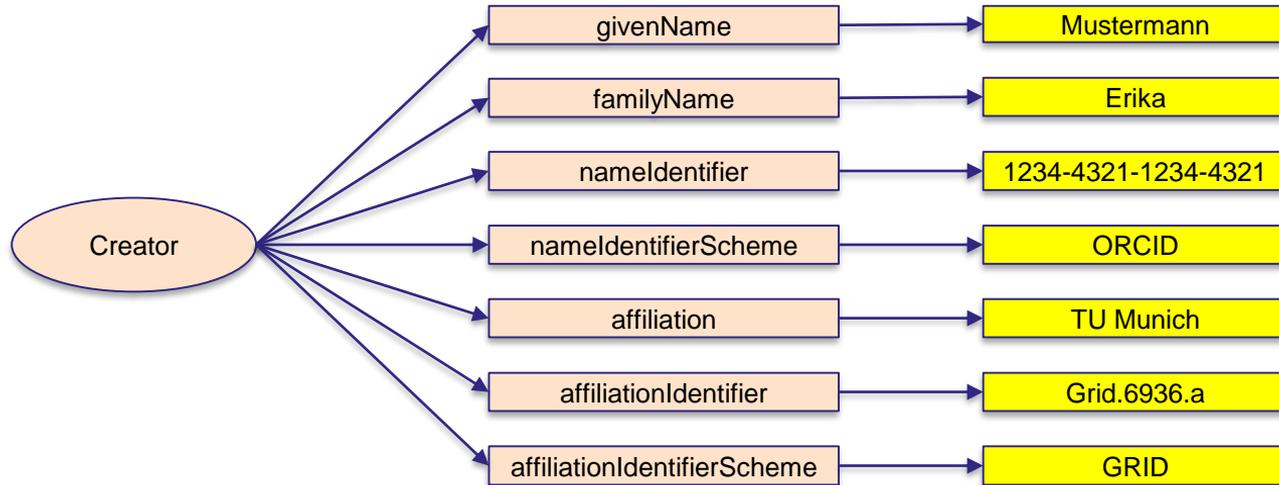
- Experiment: Roboterarm (Modell CXY-3250 Firma X. AG) wurde eingekauft. Mittels Software Y (Version 12.4.4) wurde eine Routine aufgespielt. Diese erlaubt mit dem angebrachten Werkzeug (Flaschenöffner Model Open von Firma Trink AG) Flaschen zu öffnen. Erstellt wurden Videoaufnahmen mit Kamera (Click von Firma S AG)
- Die Forschungsdaten (hier: Videoaufnahmen) werden zusammen mit der Ontoinstanz (Metadaten) an ein Repository übergeben (z.B. Universitätsbibliothek). Dort werden DOIs vergeben und (Meta)daten auffindbar und zugänglich gemacht (siehe FAIR)

## Beispiel: Eintrag von Metadaten in eine Ontologie

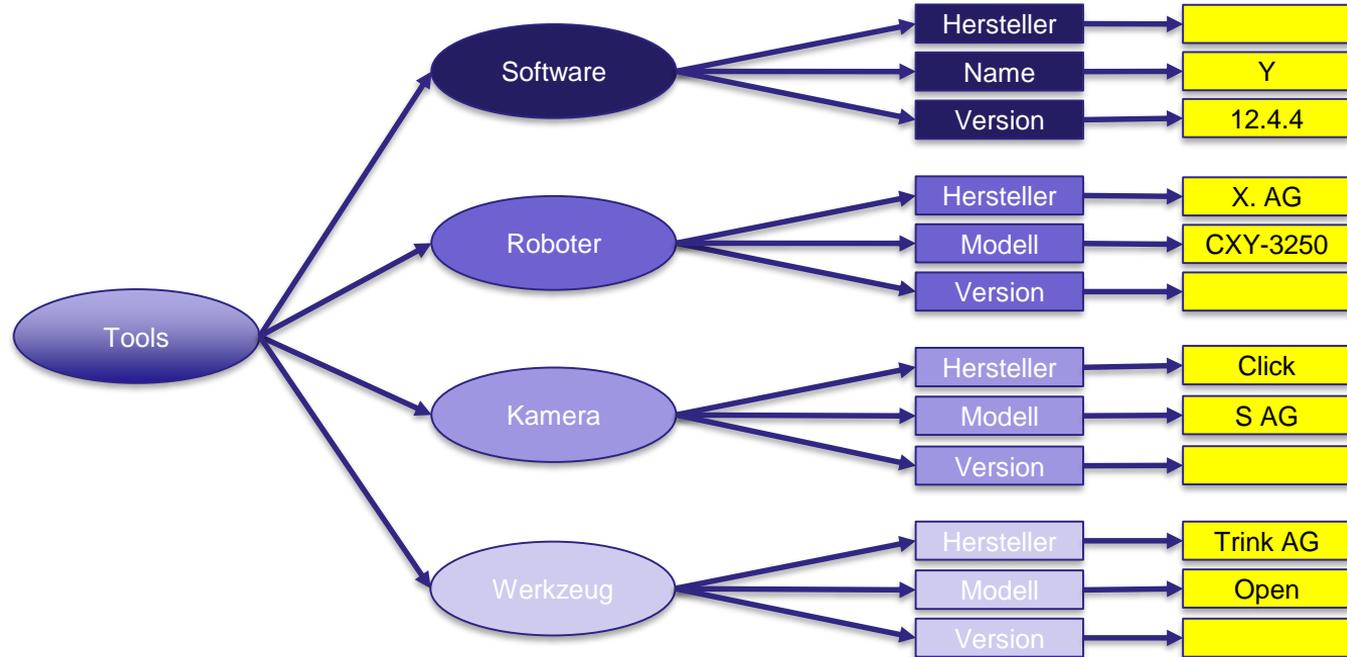
### Roboterarm öffnet Kronkorken

- Verwendet wird die SimpleRobot-Ontologie für die Forschungs(meta-)Daten und das DataCite-Schema für den Forschenden
- **Aufgabe 1:** Metadaten als Instanz in die Ontologien eintragen.
- ➡ Automatisierung ist Trumpf
- **Aufgabe 2:** Welche Metadatenfelder fehlen? Wie kann die Ontologie erweitert werden?

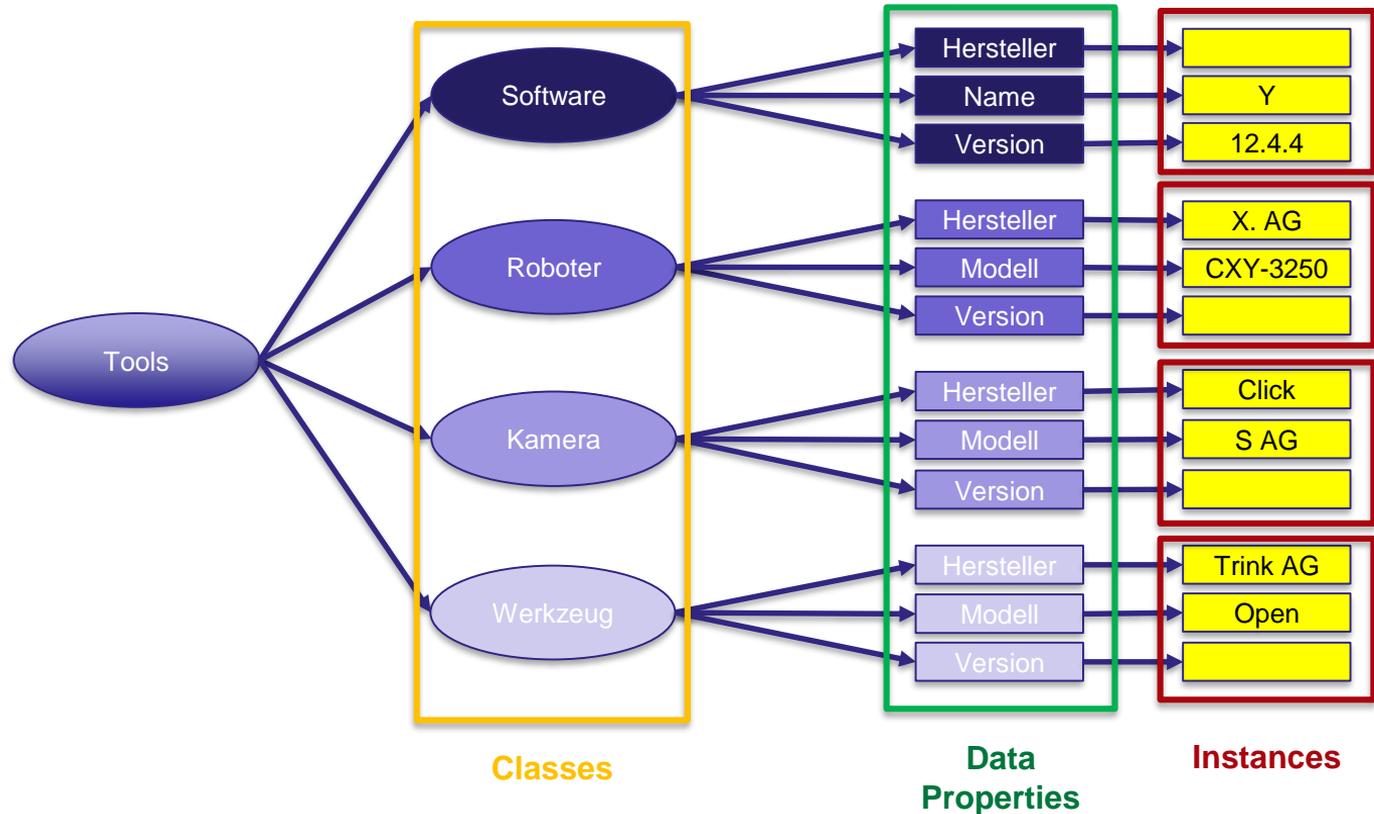
## Beispiel: DataCite Metadata Schema for Research Output



## Beispiel: SimpleRobot-Fantasie-Ontologie

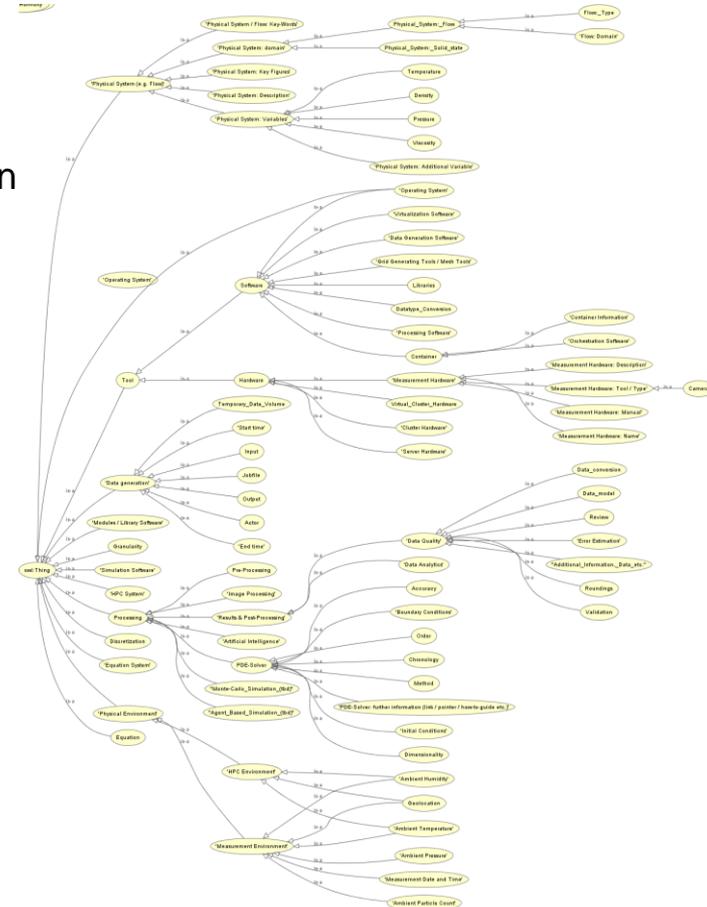


## Beispiel: SimpleRobot-Fantasie-Ontologie



## In Entwicklung (NFDI4Ing)

- Ontologie für alle Ingenieurwissenschaften
- Einbindung einer HPMC-Ontologie



## Archivierung & Veröffentlichung: Wann? Wie? Wo?

### Was ist zu beachten?

- Veröffentlichte Daten sind nicht mehr veränderbar
  - Kleine Dateien müssen in größere Archiven gepackt werden (tar, zip etc.)
  - Daten sollen nach DFG-Leitlinien 10 Jahre archiviert werden (*mittelfristig Änderungen für HPC-Daten geplant*)
  - Metadaten sollen dauerhaft gespeichert werden und ausreichende Informationen zur Reproduktion erhalten.
  - Beim Rechtemanagement können öffentlich zugängliche und nicht zugängliche Daten unterschieden werden.
  - Archivierungsoptionen sind abhängig von lokalen / institutionellen Gegebenheiten (Infrastruktur, Policy etc. an Lehrstühlen, Bibliotheken, Rechenzentren)
- ➡ Automatisierte Archivierung und Veröffentlichung als Vision

## Tools

### Metadaten

- Diverse Ontologien verfügbar / in diversen Communities herrscht Konsens (z.B. EMMO für Materialwissenschaften)
- Sammlungen existierender Ontologien, z.B. DAML, NFDI4Ing, Stanford
- NFDI4Ing erarbeitet Ontologie für Ingenieurwissenschaften mit HPMC-Subontologie
- Tools zur automatisierten Extraktion von Metadaten gemäß Ontologie (z.B. ExtractING der Uni Stuttgart in Java; Python-basierte Weiterentwicklung ist in Arbeit)

## Tools (am Standort München)

### Speicher & Austausch

- DSS (LRZ)
- TUM-NAS
- Sync+Share (LRZ)
- ISP (LRZ, ehem. TSM)
- DFN Gigamove



# Nationale Forschungsdateninfrastruktur (NFDI)

## Entstehung

2018: Bund-Länder-Vereinbarung zu Aufbau und Förderung einer Nationalen Forschungsdateninfrastruktur.

*„Die NFDI soll Standards im Datenmanagement setzen und als digitaler, regional verteilter und vernetzter Wissensspeicher Forschungsdaten nachhaltig sichern und nutzbar machen. Auf diese Weise wird eine unverzichtbare Voraussetzung dafür geschaffen, neue wissenschaftliche Erkenntnisse zu gewinnen und Innovationen in Forschung und Gesellschaft zu ermöglichen.“*

2020: Gründung des Vereins NFDI.

*„Mit der NFDI soll ein nachhaltiger digitaler Wissensspeicher geschaffen werden als unverzichtbare Voraussetzung für neue Forschungsfragen, Erkenntnisse und Innovationen.“*

# NFDI4Ing

## Entstehung

2020: Aufnahme von 9 aus 22 Konsortien in die Bund-Länder-Förderung:

- NFDI4Culture - Forschungsdaten zu materiellen und immateriellen Kulturgütern
- KonsortSWD - Sozial-, Bildungs-, Verhaltens- und Wirtschaftswissenschaften
- GHGA - Deutsches Humangenom-Phenomarchiv (**TUM**)
- NFDI4Health - personenbezogene Gesundheitsdaten
- DataPLANT - Daten in Pflanzen-Grundlagenforschung
- NFDI4BioDiversität - Biodiversität, Ökologie und Umweltdaten
- NFDI4Cat - Wissenschaften mit Bezug zur Katalyse (**TUM**)
- NFDI4Chem - Fachkonsortium Chemie
- NFDI4Ing – **Ingenieurwissenschaften** (**TUM**)

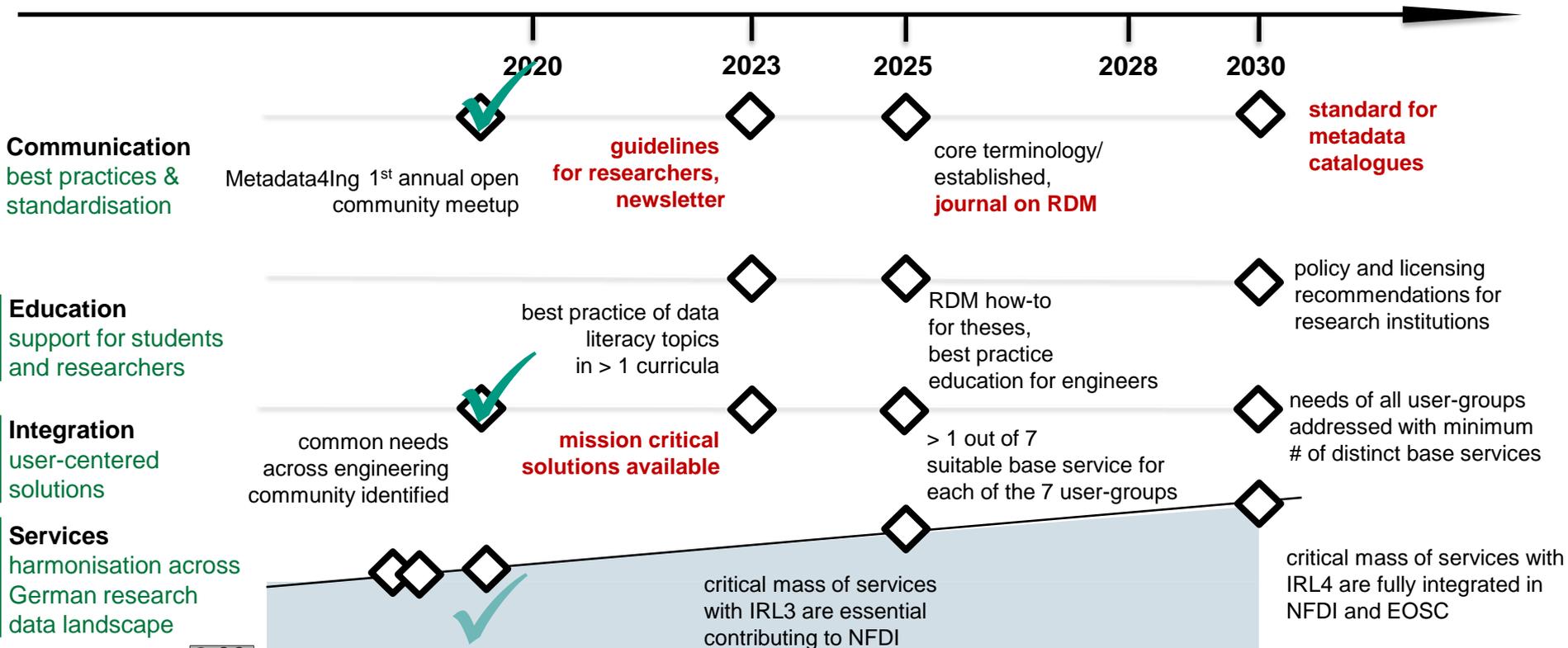
Aufnahme weiterer Konsortien geplant!

## NFDI4ing: Struktur

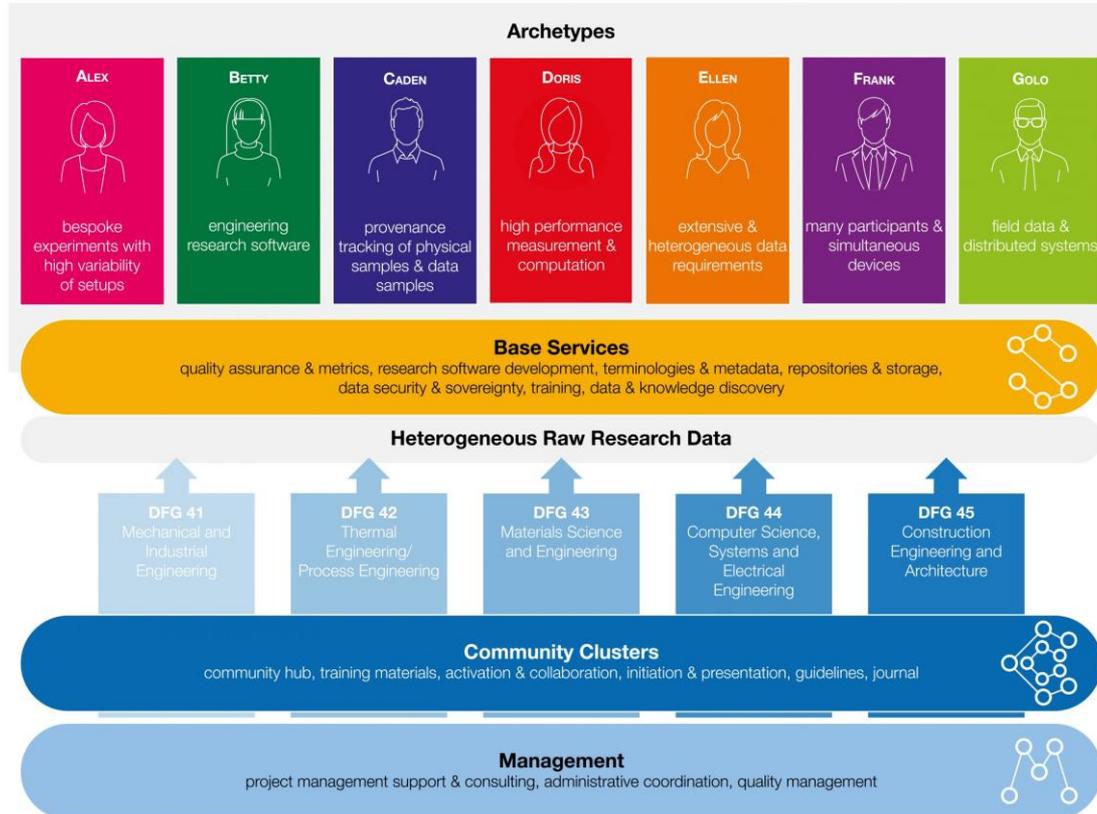
23 steering committee members @ 14 Institutions



## NFDI4ing: Deliverables



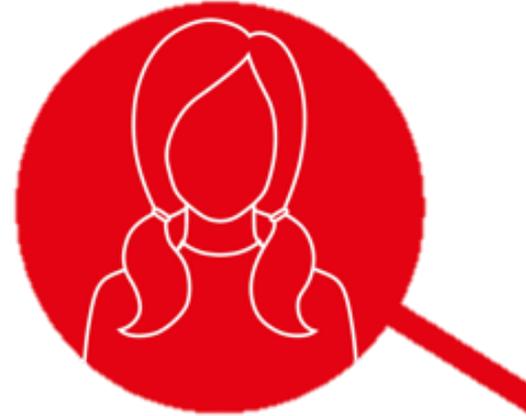
## NFDI4ing: Archetypen



## Archetyp DORIS

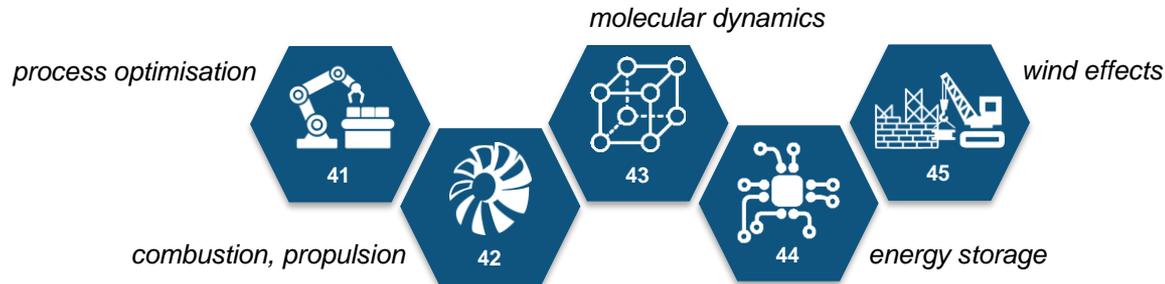
... I'm an engineer conducting and post-processing high-resolution and **high-performance measurements and computation** (simulation) **with very large data** on HPC systems.

The data sets I work with are extremely large and as such are largely immobile. This mandates tailored, hand-made software."



### My needs are

- Enable **exchange** of **huge** high-quality **datasets**.
- Provision of HPC-data to foster **wide-spread usage**.
- Drive NFDI-wide **new methodological** sharing approaches.



DORIS's patron is  
Christian Stemmer

## DORIS: Team

### TUM: LS für Aerodynamik und Strömungsmechanik (Koordination)

Prof. Dr.-Ing. Christian Stemmer | [christian.stemmer@tum.de](mailto:christian.stemmer@tum.de) (Sprecher)  
Benjamin Farnbacher | [benjamin.farnbacher@tum.de](mailto:benjamin.farnbacher@tum.de)  
Nils Hoppe | [nils.hoppe@tum.de](mailto:nils.hoppe@tum.de)

### High-Performance Computing Center Stuttgart (HLRS)

Dr.-Ing. habil. Martin Thomas Horsch | [martin.horsch@hlrs.de](mailto:martin.horsch@hlrs.de)  
Dr.-Ing. Björn Schembera | [bjorn.schembera@hlrs.de](mailto:bjorn.schembera@hlrs.de)

### Leibniz-Rechenzentrum (LRZ)

Dr. Stephan Hachinger | [stephan.hachinger@lrz.de](mailto:stephan.hachinger@lrz.de)  
Johannes Munke | [johannes.munke@lrz.de](mailto:johannes.munke@lrz.de)

### RWTH Aachen: Lehrstuhl für Strömungslehre und Aerodynamisches Institut (AIA)

Dr.-Ing. Marco Braun | [m.braun@aia.rwth-aachen.de](mailto:m.braun@aia.rwth-aachen.de)

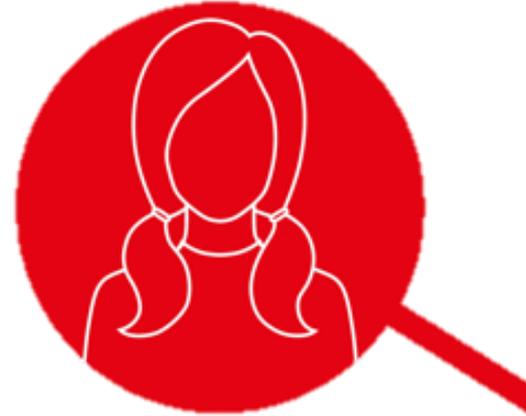
### TUM: Munich Data Science Institute (MDSI)

tbd

### TUM: Universitätsbibliothek (UB)

Team eRIC | [eric@ub.tum.de](mailto:eric@ub.tum.de)  
Dr. Manuel Hora | [manuel.hora@ub.tum.de](mailto:manuel.hora@ub.tum.de)  
Katja Kessler | [katja.kessler@ub.tum.de](mailto:katja.kessler@ub.tum.de)

... dazu diverse Pilotnutzer aus der Anwendung



## DORIS: Aufgaben und Maßnahmen

- Accessibility and access rights, data security and sovereignty
- Metadata definitions & terminologies, support to data-generating groups
- Storage & archive for very large data
- Reproducibility on large-scale high-performance systems
- **Support for third-party users & community-based training**, provision of post-processing algorithms and modules
- **Anmeldung zum DORIS-Newsletter:** Newsletter: [https://lists.tu-darmstadt.de/mailman/listinfo/nfdi4ing\\_taskarea\\_doris](https://lists.tu-darmstadt.de/mailman/listinfo/nfdi4ing_taskarea_doris)

## DORIS: aktuelles

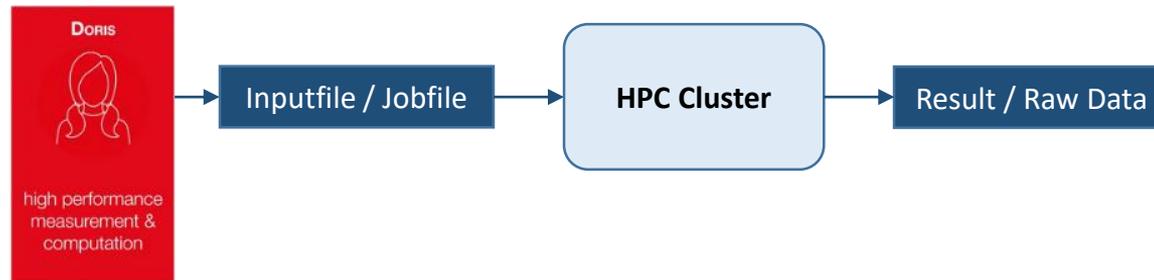
### Daran wird aktuell gearbeitet

- Mitarbeit an einem einheitlichen NFDI4Ing Metadatenchema (Ontologie)
- Entwurf eines best-practice Metadatenchemas für HPMC-Daten (Sub-Ontologie)
- Entwicklung eines Metadata-Crawlers zur automatisierten Auslese von Metadaten und Eingabe in Ontologie

## DORIS: aktuelles

### Daran wird aktuell gearbeitet

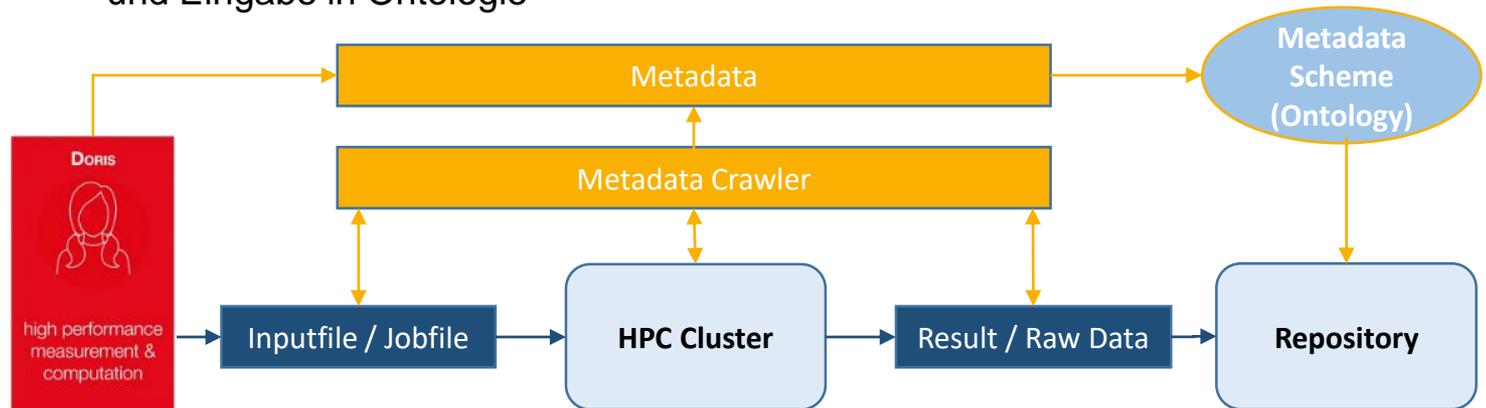
- Mitarbeit an einem einheitlichen NFDI4Ing Metadatenchema (Ontologie)
- Entwurf eines best-practice Metadatenchemas für HPMC-Daten (Sub-Ontologie)
- Entwicklung eines Metadata-Crawlers zur automatisierten Auslese von Metadaten und Eingabe in Ontologie



## DORIS: aktuelles

### Daran wird aktuell gearbeitet

- Mitarbeit an einem einheitlichen NFDI4Ing Metadatenchema (Ontologie)
- Entwurf eines best-practice Metadatenchemas für HPMC-Daten (Sub-Ontologie)
- Entwicklung eines Metadata-Crawlers zur automatisierten Auslese von Metadaten und Eingabe in Ontologie



## DORIS: aktuelles

### Daran wird aktuell gearbeitet

- Mitarbeit an einem einheitlichen NFDI4Ing Metadatenchema (Ontologie)
- Entwurf eines best-practice Metadatenchemas für HPMC-Daten (Sub-Ontologie)
- Entwicklung eines Metadata-Crawlers zur automatisierten Auslese von Metadaten und Eingabe in Ontologie
- Test von Container- und Postprocessing-Tools für HPMC-Daten
- Testprojekte in den großen Rechenzentren (LRZ, HLRS, JSC)
- Test von Transfertools und Transferprotokollen

## Fokus: HPMC-Daten

### Besonderheiten

- Daten werden in HPMC-Centern in personalisierten Accounts erzeugt und gespeichert  
→ keine Indizierung durch Repositorien oder Suchmaschinen
- Spezielle Hard- & Software zum Erzeugen, Lesen und Bearbeiten notwendig
- Terabyte bis Petabyte → Daten nicht mobil
- „Daten“ bestehen aus diversen Komponenten (Code, Inputfile, Rohdaten, postprozessierte Daten, Metadaten etc.)
- Keine Best-Practice-Beispiele für Datenmanagementpläne, Terminologien etc. vorhanden
- DSS bietet attraktive Möglichkeiten

### Umsetzung der FAIR-Prinzipien

**Findable:** durch Speicherung in persönlichen Accounts nicht auffindbar



**Accessible:** s.o., dazu kein Zugang zu HPMC-Centern für Externe



**Interoperable:** abhängig von Speicherung und Struktur durch Erzeugende



**Reusable:** abhängig von F, A & I



## DSS – Data Science Storage (LRZ)

- \* Doku: <https://doku.lrz.de/display/PUBLIC/Data+Science+Storage>
- \* Hochperformanter Speicher direkt „im“ Rechenzentrum (LRZ)
- \* Für große Datensätze ausgelegt.
- \* Teilen der Daten im LRZ-Ökosystem und mit Externen Partnern möglich
- \* Zugriff von außerhalb via Globus Online
- Backup und Archivierung möglich
- \* Servicebezug auf unterschiedlichen Wegen
- **Kostenlos aka „Jemand anderes Zahlt“:**
  - TUM-DSS: „Joint-Venture“ verschiedener TUM Fakultäten
  - Jedes LRZ LinuxCluster Projekt bekommt 10TB DSS als Grundversorgung
  - SuperMUC-Projekte können DSS-Speicher mitbeantragen
- **Bezahlt:**
  - DSS on Demand: Pay per Use
  - DSS Joint Project Offer: Eigener DSS Buildingblock
- \* *"Zugriffsmanagement und Rechtevergabe für TUM-DSS komplex*
- \* *Was muss der TUM-User konkret machen, wenn er DSS-Speicher braucht? Wo findet man den Data Curator etc.? Wie läuft der Antrag über das LRZ?"*



## DSS: Introduction

### Aufbau

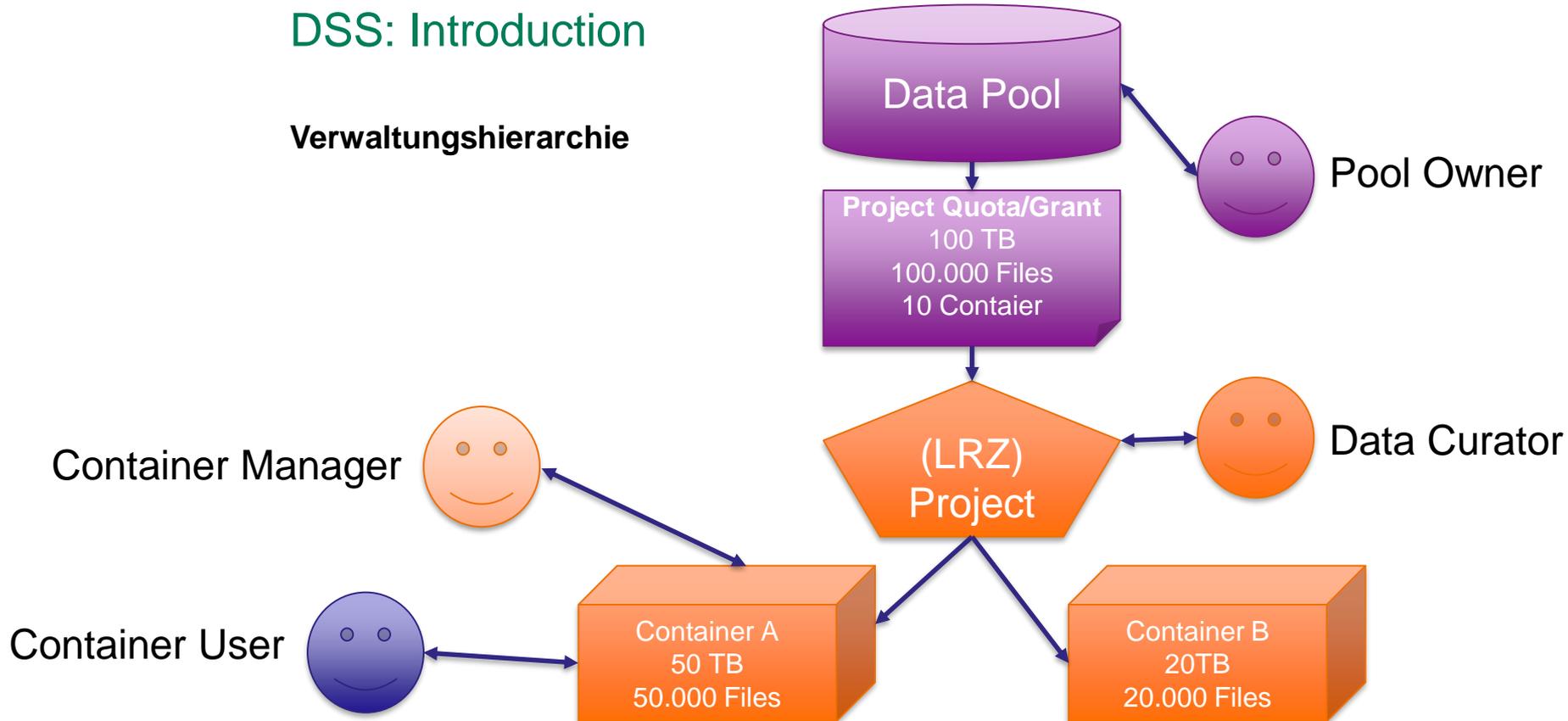
- \* Auf Linux Cluster (und teilweise SuperMUC-NG) automatisch eingehängt wie normales (Festplatten-)Laufwerk

```
nils@DESKTOP-R88DNLK: ~  
ga56tik3@cm2login4:~> cd /dss/dsstumfs01/pn34ka/pn34ka-dss-0000/  
ga56tik3@cm2login4:/dss/dsstumfs01/pn34ka/pn34ka-dss-0000> ls  
GlobusTest  RiemannCubes  
ga56tik3@cm2login4:/dss/dsstumfs01/pn34ka/pn34ka-dss-0000>
```

- \* Für Zugriff von Compute Cloud aus ist NFS Export des Containers möglich
- \* Ermöglicht kollaboratives Arbeiten von unterschiedlichen LRZ Maschinen aus
- \* Austausch mit Externen via Globus Sharing
- \* Integration in TUM-Workbench möglich => UB Kontaktieren
- \* DSS Speicherbereiche/Container werden vom sog. Data Curator verwaltet (z.B. Institutsleiter, Fakultätsbeauftragter, Master User, ...)

## DSS: Introduction

### Verwaltungshierarchie



## DSS: Introduction

### Konfiguration: [dssweb.dss.lrz.de](https://dssweb.dss.lrz.de)

- \* LRZ bietet ein Command-Line Tool sowie eine Webinterface für die DSS-Konfigurationen an
- \* Data Curator untergliedert Speicher (z.B. 10TB) in sogenannte Container. Container können unterschiedlich groß sein und unterschiedliche konfiguriert werden (z.B. Freigaben)
- \* Curator kann Konfiguration auf Container-Ebene auf einen Manager übertragen
- \* Nutzer werden auf Container-Ebene eingeladen.
- \* Alle weiteren Einstellungen erfolgen auf Container-Ebene

## DSS: Introduction

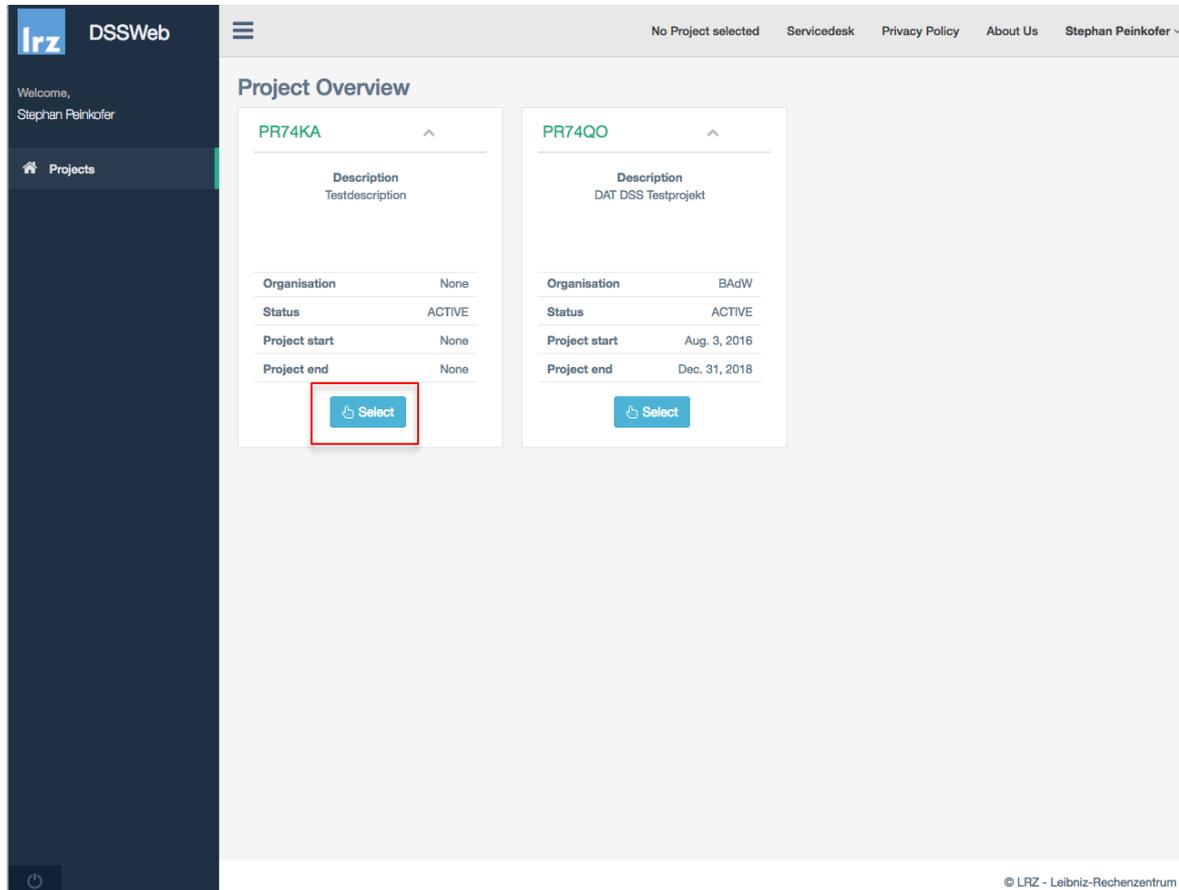
### Konfiguration: [dssweb.dss.lrz.de](https://dssweb.dss.lrz.de)

- \* Curator/Manager können einstellen:
  - Container Modus (Jeder hat Zugriff auf alles/Jeder hat ein privates Verzeichnis/usw)
  - \* Zugriffsrechte für MWN-Benutzer (LRZ Ökosystem)
  - \* Freigaben als NFS-Laufwerk (nur innerhalb des LRZ Rechnerwürfels)
  - \* Globus-Freigaben für externe Benutzer
  - \* Backup-Regelungen
  
- \* Standardeinstellung sind generell ein guter Start
- \* **ACHTUNG:** Einige Einstellungen sind nur einmalig beim Erstellen der Container möglich

## TUM-DSS: Data Curator

### DSS Ansprechpersonen / Data Curators

Fakultät	Person	E-Mail
Bau Geo Umwelt	Dr. Alexander Braun	<a href="mailto:alex.braun@tum.de">alex.braun@tum.de</a>
Chemie	Dr. Josef Homolka	<a href="mailto:josef.homolka@ph.tum.de">josef.homolka@ph.tum.de</a>
Elektrotechnik und Informationstechnik	Dr. Claus von Rücker	<a href="mailto:ruecker@tum.de">ruecker@tum.de</a>
Informatik	Dr. Benjamin Hof	<a href="mailto:rbg-system@in.tum.de">rbg-system@in.tum.de</a>
Maschinenwesen	Prof. Christian Stemmer	<a href="mailto:christian.stemmer@tum.de">christian.stemmer@tum.de</a>
Physik	Dr. Josef Homolka	<a href="mailto:josef.homolka@ph.tum.de">josef.homolka@ph.tum.de</a>
Mathematik	Dr. Benjamin Hof	<a href="mailto:rbg-system@in.tum.de">rbg-system@in.tum.de</a>
Luftfahrt, Raumfahrt, Geodäsie	Mario Brunner	<a href="mailto:mario.brunner@tum.de">mario.brunner@tum.de</a>
Nicht auf der Liste? Fragen Sie Ihren Information Officer: <a href="https://www.it.tum.de/cioio-gremium/mitglieder/">https://www.it.tum.de/cioio-gremium/mitglieder/</a>		



**lrz DSSWeb**

Welcome,  
Stephan Peinkofer

Projects

No Project selected | Servicedesk | Privacy Policy | About Us | Stephan Peinkofer

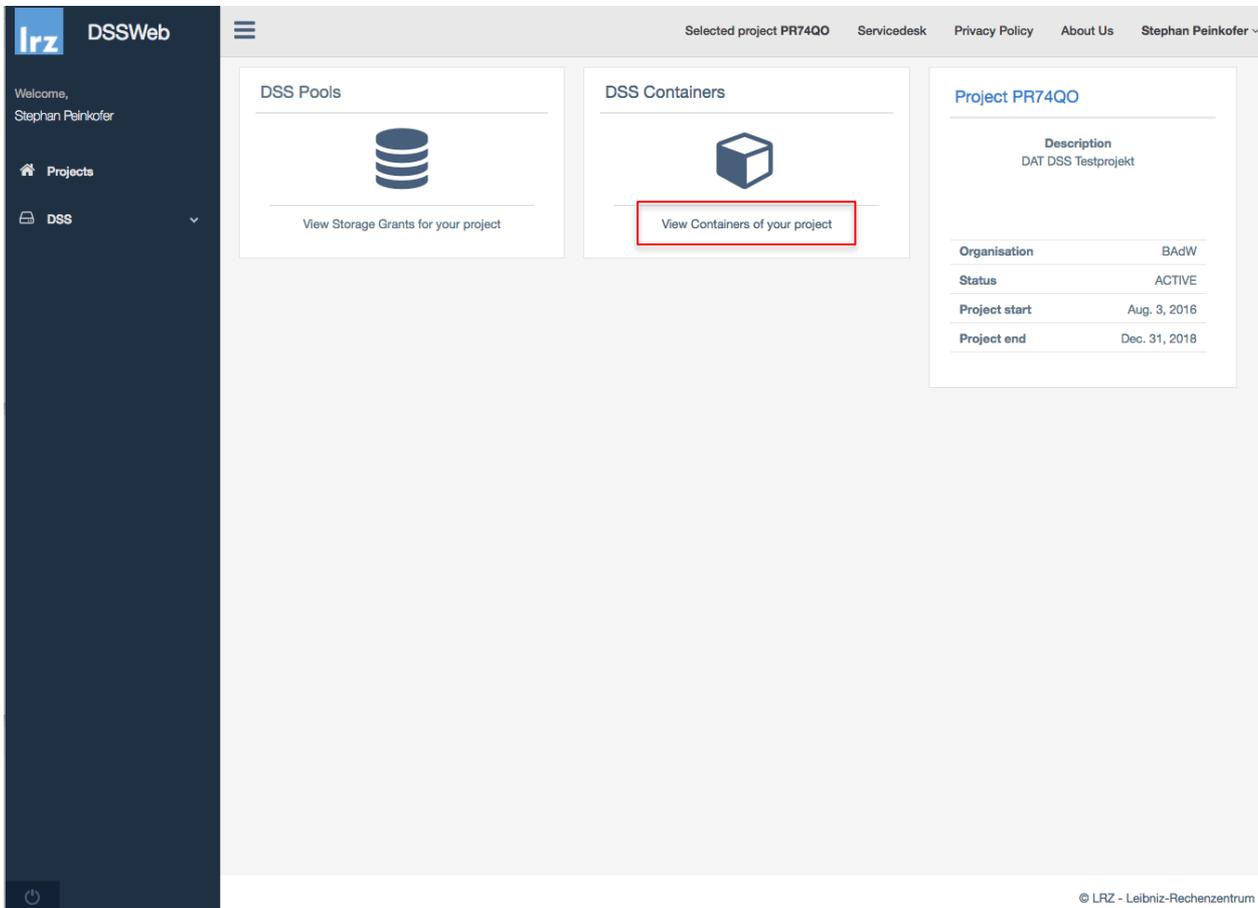
### Project Overview

Project ID	Description	Organisation	Status	Project start	Project end
PR74KA	Testdescription	None	ACTIVE	None	None
PR74QO	DAT DSS Testprojekt	BAdW	ACTIVE	Aug. 3, 2016	Dec. 31, 2018

**Select** (highlighted in red box)

**Select**

© LRZ - Leibniz-Rechenzentrum



**lrz** DSSWeb

Welcome,  
Stephan Peinkofer

Projects

DSS

Selected project **PR74QO** Servicedesk Privacy Policy About Us Stephan Peinkofer

### DSS Pools



View Storage Grants for your project

### DSS Containers



View Containers of your project

### Project PR74QO

Description  
DAT DSS Testprojekt

Organisation	BAdW
Status	ACTIVE
Project start	Aug. 3, 2016
Project end	Dec. 31, 2018

© LRZ - Leibniz-Rechenzentrum

**Irz** DSSWeb

Welcome,  
Stephan Peinkofer

- Projects
- Storage Pool Quotas
- DSS

Selected project PR74QO    Servicedesk    Privacy Policy    About Us    Stephan Peinkofer ▾

### DSS Container Details

**PR74QO-DSS-0000**

Description  
Testcontainer

Filesystem Path  
/dsa/dstests01/pr74qo/pr74qo-dss-0000

Pool: DSS Testpool

Quota GB: 1,024

Used GB: 0

Quota Files: 100,000

Used Files: 1

Status: ●

Backup Mode: NONE

ID Mode: NORMAL

Globus Sharing: ●

[Edit](#)

Used Space



**Project PR74QO**

Description  
DAT DSS Testprojekt

Organisation: BAuW

Status: ACTIVE

Project start: Aug. 3, 2018

Project end: Dec. 31, 2018

Container User

SIM User    Globus User    Auto Group Links

Username	Firstname	Lastname	Status	Access Type	Quota GB	Quota Files	Inviter	Expires at	Actions
...	...	...	<span style="color: green;">●</span>	RW	None	None	AutoCuratorInviter	None	<span style="color: blue;">ⓘ</span> <span style="color: green;">✉</span> <span style="color: orange;">✎</span> <span style="color: red;">✖</span>

[+ Add new user](#)

NFS Exports

IP	Access Type	Status	Mount Path	Expires
10.156.8.53	RW	<span style="color: green;">●</span>	datdscst102.dss.irz.de/dsa/dstests01/pr74qo/pr74qo-dss-0001	None
10.156.8.54	RW	<span style="color: green;">●</span>	datdscst102.dss.irz.de/dsa/dstests01/pr74qo/pr74qo-dss-0001	None

## DSS: Introduction

### Temperatur

- \* DSS basiert auf online Speichersystemen und ermöglicht Kooperation in der performanten Genese der Daten
- \* Was wenn die Genese abgeschlossen ist, die Daten aber noch aufbewahrt werden müssen?
  - Übergabe an die UB und Publikationsmöglichkeit (Geplant)
    - Verlinkung im mediaTUM (mit DOI) und die Daten werden ordentlich archiviert, (Lese-)Zugriff bleibt erhalten.
  - LRZ Data Science Archive
    - Momentan in Friendly User Phase für SuperMUC-NG
    - Öffnung für TUM/LMU in Planung
    - Quasi DSS auf Tape (AWS S3 vs. AWS Glacier)

## DSS: Introduction

### Pitfalls

- \* DSS ist in das MWN-Benutzerverzeichnis integriert
  - NFS Exports an Compute Cloud VMs erwarten, dass auf der VM die gleichen UIDs wie im MWN-Benutzerverzeichnis verwendet werden
    - <https://doku.lrz.de/display/PUBLIC/DSS+documentation+for+data+curators#DSSdocumentationfordatacurators-NFSHelperfunctions>
  - DSS ist nicht (optimal) geeignet für:
    - Betrieb von Datenbanken
      - Auch wenn manche Kunden vielleicht was anderes behaupten
    - Viele Millionen kleiner Dateien
    - Besondere Kategorien personenbezogener Daten gemäß Art. 9 DSGVO
  - DSS Einladungen werden an die im IDM hinterlegten Mailadressen gesendet
    - Kontaktdaten aktuell halten

# Globus Online: Introduction

## Globus Architektur

- \* Hybrides System
  - Globus Endpoints: Datenquelle/-senke
  - Globus Service: Orchestrator/Kümmerer
  
- Fire and Forget Transfers
  - Globus kümmert sich darum dass die Daten korrekt ankommen
  
- Premium Addon: Globus Sharing
  - Für einen Teilbereich wird das ACL Management an den Globus Service outsourced

# Globus Online: Introduction

## Globus Endpoints

- \* Basiert auf GridFTP Protokoll
  - \* Verknüpfungen von „Endpoints“
  - \* Viele Rechenzentren stellen öffentliche Globus Endpoints zur Verfügung
    - Globus Connect Server
    - Optimiert für schnellen Datentransfer
    - Multi-User
  - \* Durch Installation von (freier) Software kann beliebiger Rechner als „persönlicher“ Endpunkt fungieren
    - Globus Connect Personal
    - Limitiert durch Netzanbindung des Rechners
    - Single-User

## Globus Online: Introduction

**<https://app.globus.org/>**

- \* US-hosted!
  - Globus Online ist eine Einrichtung der University of Chicago
  - Daten fließen immer direkt zwischen den Transferendpunkten
  - Metadaten (z.B. Pfadnamen, Benutzername, Vorname, Nachname, Email, Short lived access tokens) gehen über Globus
  - Rechtsgrundlage: Standardvertragsklauseln -> Freiwillige Einwilligung
- \* Login via Institution (Shibboleth)
  - In der Regel sollte hier das LRZ als IdP und nicht die TUM verwendet werden

## Globus Online: Introduction

### Globus Online vs. Globus Sharing

- Globus Online/Mapped Endpoints
  - DSS Endpunkt: Leibniz Supercomputing Centre's DSS – CILogon
  - [https://app.globus.org/file-manager?origin\\_id=a9ab7d72-f77b-11e6-ba7f-22000b9a448b](https://app.globus.org/file-manager?origin_id=a9ab7d72-f77b-11e6-ba7f-22000b9a448b)
  - Mapped in Globus eingeloggten Benutzer auf LRZ/TUM/LMU Account
  - Ist der Endpunkt den alle "normalen" DSS Benutzer verwenden sollten
  - Ist "immer Aktiviert"
  - Besonderheit beim Austausch zwischen zwei LRZ-Accounts:  
<https://doku.lrz.de/display/PUBLIC/DSS+How+to+transfer+data+with+Globus+Online+between+two+LRZ+user+accounts>
  
- Globus Sharing/Shared Endpoints
  - DSS Endpunkte: LRZ DSS Container XXXXXX-DSS-YYYY
  - Ist nur für Data Sharing mit Externen ohne LRZ/TUM/LMU Account gedacht
  - Ist nur Verfügbar wenn Globus Sharing für den Container enabled wurde

## Globus Online: Introduction

### Globus Online vs. “klassisches” GridFTP vs. kommerzielle Tools

#### \* Globus Online/Sharing

- Von manchen Rechenzentren nicht unterstützt
- Relativ Komfortabel/Einfach zu Nutzen
- Einfache Installation auf eigenen Rechnern möglich (aber ggf. vom Datenkurator nicht gerne gesehen?)
- Auf für PBs geeignet - passender Endpunkt vorausgesetzt :)

#### \* Klassisches GridFTP

- Kompliziertes Setup
- Komplizierte Nutzung
- Kein Sharing

#### \* One Drive, Dropbox etc.

- Datenschutz und Datensicherheit nicht gewährleistet
- Ggfls. konträr zu Vorgaben von Projektträger, Arbeitgeber etc.
- Nur für „kleinere“ Datenmengen geeignet

# FAIR Supercomputing Data at LRZ

Towards a new Research Data Management (RDM)  
Service

# HPC data vs. FAIR principles

Wilkinson et al. (2016) – Scientific Data, 3, 160018



Findable



Accessible



Interoperable

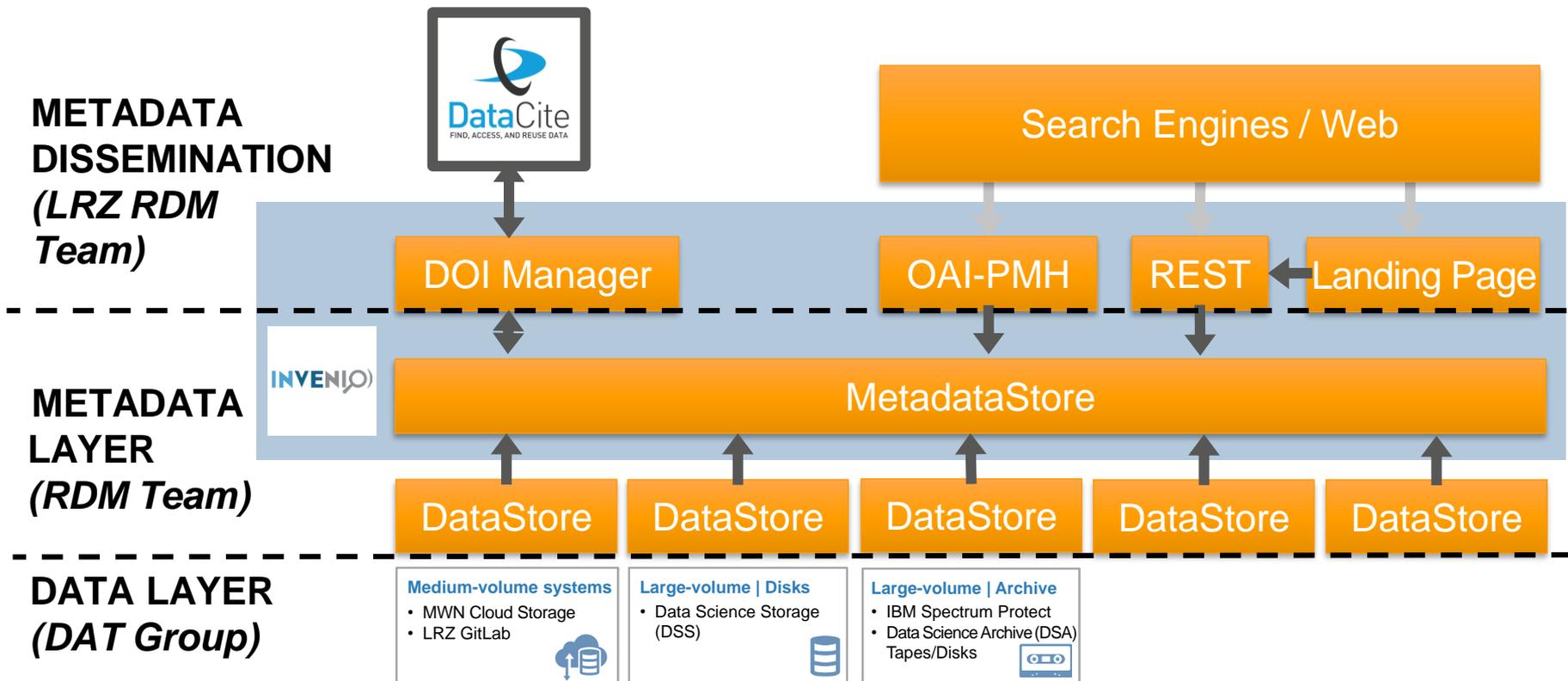


Reusable



- You'd like to make your data „citable“, getting a DOI for it?
- Your funding agency forces you to „publish“ your data?
- You want others to find your data, via data or web search engines?
- Your boss told you to deposit a description („metadata“) with your data?

Typical data from LRZ HPC machines is too large for repositories!



# How will this work?

## Idea:

- You will mark folders which you want to be „published“
- Folders will contain YAML (or similar files...
  - ...with metadata or
  - ...with instruction of how to get metadata (e.g. from your folder names or file metadata)

## LRZ core service:

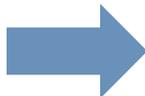
- DataCite metadata (domain-agnostic, necessary for DOI)

Domain-specific metadata: To be coordinated with NFDI consortia.

```
/dss/dssfr~fig.yaml [-M--] C: 1 L:[ 49/327 49+ 0] *(1531/11681b)= 32 0x20
- creatorName:
  default: "Stephan Hachinger"
  nameType:
    default: "Personal"
  givenName:
    default: "Stephan"
  familyName:
    default: "Hachinger"
  nameIdentifiers:
    - nameIdentifier:
        default: "0000-0001-8341-1478"
        nameIdentifierScheme:
          default: "ORCID"
        schemeURI:
          default: "http://orcid.org/"
  affiliations:
    - affiliation:
        default: "Bavarian Academy of Sciences and Humanities"
        affiliationIdentifier:
          default: "https://ror.org/001rdaz60"
        affiliationIdentifierScheme:
          default: "ROR"
        schemeURI:
          default: "https://ror.org"
  titles:
    title:
      default: "ViNA Global Simulation Data"
  publisher:
    default: "Leibniz Supercomputing Centre"
  PublicationYear:
    default: "2020"
  subjects:
    - subject:
        default: "Hydrology"
        subjectScheme:
          default: "Library of Congress"
        schemeURI:
          default: "https://id.loc.gov/authorities/subjects.html"
        valueURI:
          default: "https://id.loc.gov/authorities/subjects/sh85063458.html"
    - subject:
        default: "Water efficiency"
        subjectScheme:
          default: "Library of Congress"
```

# User-friendly metadata workflow envisaged

web tool to  
edit metadata  
specification



```
/dss/dssf-fig.yaml [-M--] C: 1 L:[ 49/327 49+ 0] *(1531/11681b)- 32 0x20
- creatorName:
  default: "Stephan Hachinger"
  nameType:
    default: "Personal"
  givenName:
    default: "Stephan"
  familyName:
    default: "Hachinger"
  nameIdentifiers:
  - nameIdentifier:
    default: "0000-0001-8341-1478"
    nameIdentifierScheme:
      default: "ORCID"
    schemeURI:
      default: "http://orcid.org/"
  affiliations:
  - affiliation:
    default: "Bavarian Academy of Sciences and Humanities"
    affiliationIdentifier:
      default: "https://ror.org/001rdaz60"
    affiliationIdentifierScheme:
      default: "ROR"
    schemeURI:
      default: "https://ror.org"
  titles:
  title:
    default: "ViWA Global Simulation Data"
  publisher:
    default: "Leibniz Supercomputing Centre"
  PublicationYear:
    default: "2020"
  subjects:
  - subject:
    default: "Hydrology"
    subjectScheme:
      default: "Library of Congress"
    schemeURI:
      default: "https://id.loc.gov/authorities/subjects.html"
    valueURI:
      default: "https://id.loc.gov/authorities/subjects/sh85063458.html"
  - subject:
    default: "Water efficiency"
    subjectScheme:
      default: "Library of Congress"
```



# InvenioRDM-based Web Interface for LRZ RDM service



The screenshot shows a web browser window with the following elements:

- Browser Tab:** LRZ LTDS
- Address Bar:** <https://rdm.lab.lrz.de>
- Page Header:** lrz Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities
- Hero Section:** A blue-tinted background image of server racks with the text "LRZ LTDS InvenioRDM" and a search bar containing the text "Wilson Inc's".
- Main Content:** A large white area with the text "Welcome to the LRZ LTDS InvenioRDM instance!".
- Footer:** A circular diagram with four segments: "CREATE" (grey), "STORE" (blue), "RE-USE" (red), and "USE" (orange).

# InvenioRDM-based Web Interface



File Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

Search results

https://rdm.lab.lrz.de/search?q="Wilson Inc's"&ip=1&cs=10&sort=newest

Leibniz Supercomputing Centre  
of the Bavarian Academy of Sciences and Humanities

"Wilson Inc's"

2 result(s) found

Newest

July 1994 (v0.0.1) Book section Metadata-only View

**Wilson Inc's gallery**

Mitchell, John, Blair, Paul, Cline, Paul

Worry dinner bank experience wonder best newspaper. Behind almost prevent. Myself adult of. Travel part card prepare where. Fish federal win morning indicate lead task. Available number stage indicate number. Debate line wish interest physical. Happen tend development something painting prevent realize. Instead part along be about. Husband langu...

smile remember

Uploaded on May 31, 2021

Sep 5, 1970 -- Jul 11, 2010 (v0.0.1) Technical note Metadata-only View

**Wilson Inc's gallery**

Gutierrez, Joseph, Aguiar, Timothy, Dillon, Paul

Process couple between ahead time. Land arm crime event fine artist material. Road parent south pull still. Place upon operation fund base room. Simply radio what war. Anyone people year behavior voice. Brother rule structure gun. Note artist student west share second. Sea increase election leader none. Class cold home during black good nature. ...

market deep

Uploaded on May 31, 2021

< 1 >

10 results per page

**Contact**

Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities  
Boltzmannstrasse 1  
85748 Garching near Munich  
RDM Team E-Mail: rdm@lists.lrz.de  
Website: www.lrz.de

**Social Media**

LRZ on Facebook  
LRZ on Twitter  
LRZ on YouTube

# InvenioRDM-based Web Interface



The screenshot shows a web browser window displaying the record for 'Wilson Inc's gallery'. The browser address bar shows the URL <https://rdm.lab.lrz.de/records/dy92n-2zv81>. The page header includes the lrz logo and 'Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities'. The main content area features the title 'Wilson Inc's gallery' with a date range 'Sep 5, 1970 – Jul 11, 2018' and version 'v0.0.1'. There are buttons for 'Technical note' and 'Metadata-only'. A 'Versions' sidebar shows the current version details. Below the title, there are sections for 'Rights Holder(s)', 'DOI' (10.21979/datacite.dy92n-2zv81), and 'Subjects' (market, deep). A descriptive paragraph follows. At the bottom, a 'Details' table provides metadata.

Details	
Resource type	Technical note
Publisher	InvenioRDM
Additional titles	subtitle a research data management platform
Additional descriptions	methods Responsibility them garden practice ground listen difficult. Economy onto responsibility. Last ready season friend great record hit thank. Police toward great listen.
Formats	application/pdf
Funding	Award OpenAIRE ( 246686 )
	Funder European Commission

+ instructions on how to obtain data  
+ later: link to GLOBUS, etc.

TUM Library is providing consulting / courses on Research Data Management via <https://www.ub.tum.de/research-data> / [eric@ub.tum.de](mailto:eric@ub.tum.de).

For non-TUM users:

LRZ RDM consulting service

via <https://servicedesk.lrz.de> (under Beratung/Consulting)

“Pure” storage questions please directly to DAT experts (select appropriate service).

# Thanks for your attention!

Stephan Hachinger – LRZ RDM Team Lead – [hachinger@lrz.de](mailto:hachinger@lrz.de)

LRZ RDM Team – [rdm@lists.lrz.de](mailto:rdm@lists.lrz.de)

## Weitere Informationen

### Downloads

- Folien, Aufzeichnung: <https://zenodo.org/communities/nfdi4ing?page=1&size=20>
- Software: vsl. GitLAB
- Publikationen: MediaTUM, Zenodo

### Events

- NFDI4Ing Konferenz am 27. & 28. September (online)

### *Stand und Zukunft des Forschungsdatenmanagements in den Ingenieurwissenschaften*

- <https://nfdi4ing.de/konferenz/>

## Weiterführende Links

### EngMeta

- <https://doi.org/10.18419/darus-500>
- [doi.org/10.1007/978-3-030-14401-2\\_12](https://doi.org/10.1007/978-3-030-14401-2_12)
- <https://dx.doi.org/10.1504/JMSO.2020.107792>

### DIPL-ING

- <https://doi.org/10.2314/KXP:1693393980>
- <https://books.ub.uni-heidelberg.de/heibooks/catalog/book/598/c8422>
- <https://darus.uni-stuttgart.de/>
- <https://dataverse.org/>
- <https://github.com/bjschembera/ExtractIng>
- <https://github.tik.uni-stuttgart.de/darus/DaRUSApp> (Passwort)

### Weitere

- Projekt AIMS:  
<https://www.wzl.rwth-aachen.de/go/id/ivong>
- Meneveau et al. (2017), Charles and Marusic, Ivan “Turbulence in the Era of Big Data: Recent Experiences with Sharing Large Datasets.” In A. Pollard, L. Castillo, L. Danaila and M. Glauser (eds.), Whither Turbulence and Big Data in the 21st Century?. Springer, (2017): 497-507