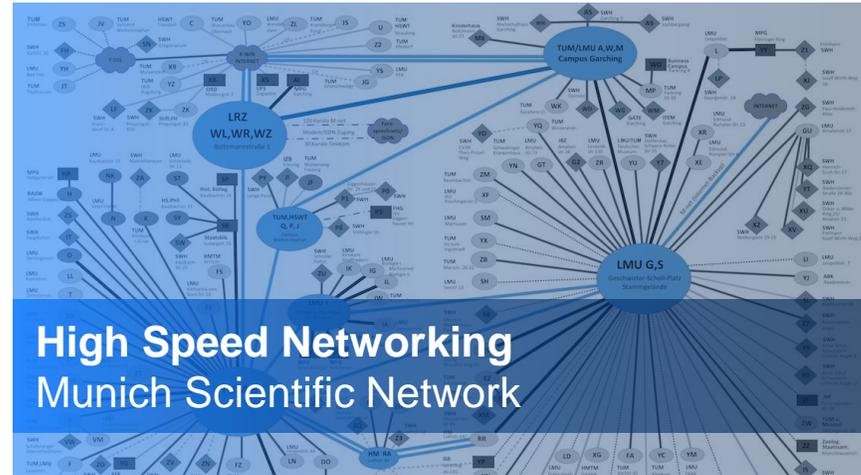


Der Weg zu einem
**Serviceangebot „Forschungsdatenmanagement“
des Leibniz-Rechenzentrums (LRZ)**

LRZ – Rechenzentrum für die Münchner Universitäten, Bayern, GCS (D) und PRACE (EU) “Cutting-Edge” IT-Infrastruktur



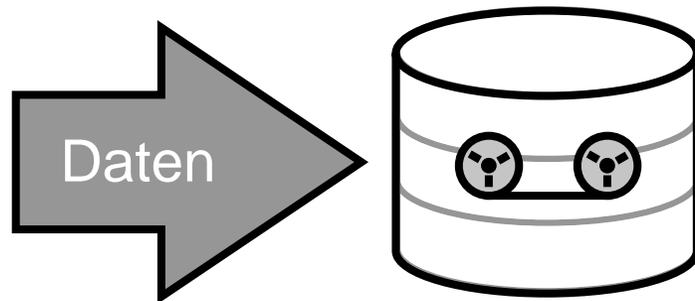
- Email
- Network
- Storage
- Cloud Computing
- Cluster
- HPC
- Training
- Consultancy



Warum Forschungsdatenmanagement am Rechenzentrum?

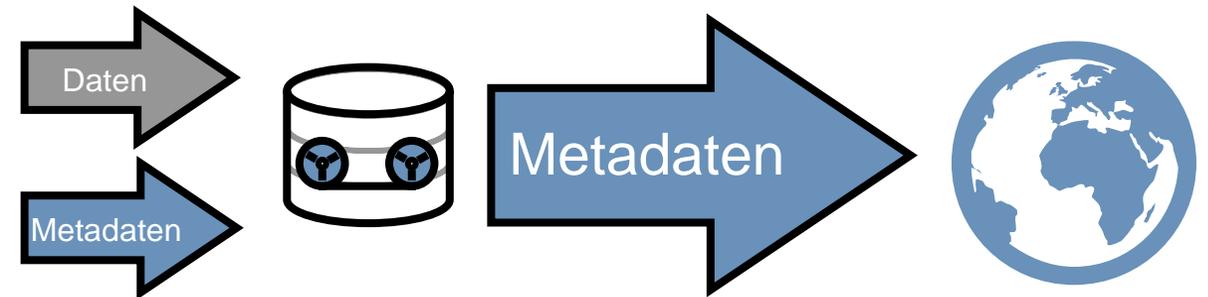
LRZ-Speichersysteme

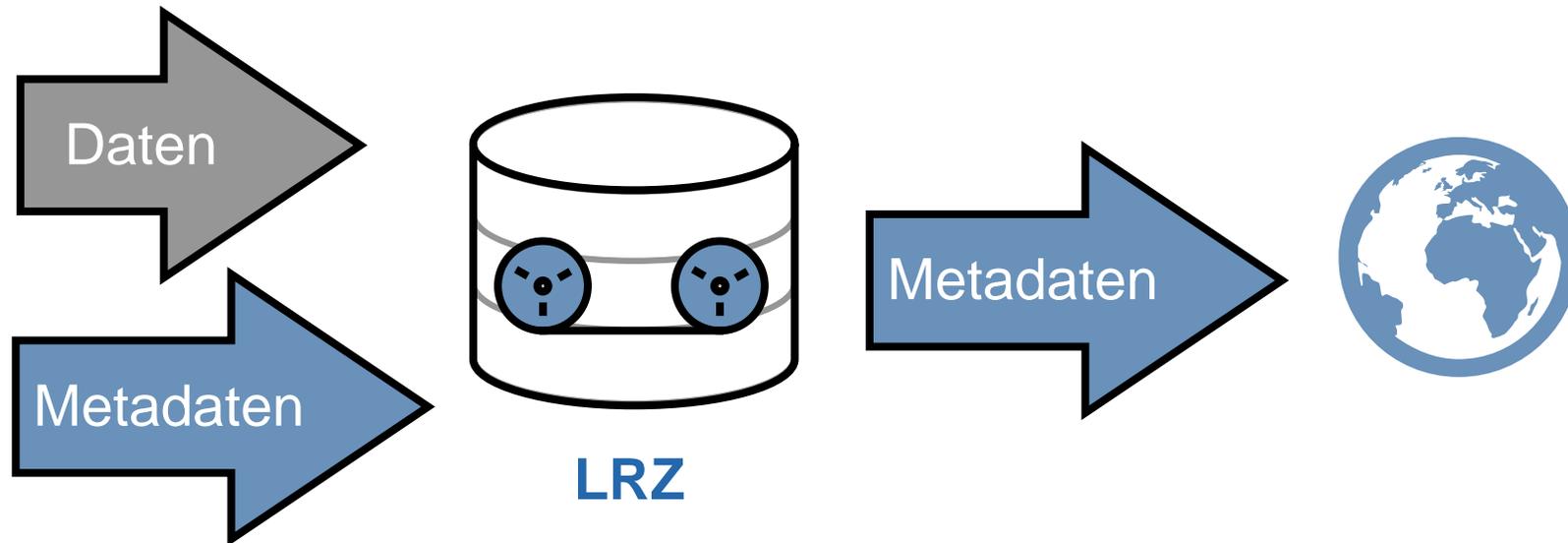
- Hochleistungsfähig
- Enthalten teils riesige, nicht weg-bewegbare Datensätze



Bibliotheken / Repositorien / Forschungsdatenmanagement

- „Verwaltbare“ Datenmengen
- Up- und Download-Portale
- Suche möglich / Auffindbarkeit gegeben





Schwerpunkte

- Domänen-agnostisch, etwas Fokus auf Naturwissenschaften/Ingenieure
- Basis-Metadaten („bibliographische Daten“ - DataCite) und DOIs
- Automatisiertes Metadaten-Sammeln und Extraktion von verschiedenen Systemen her

Serviceangebot „Forschungsdatenmanagement“ des LRZ: Was gibt es schon?

Kleine und mittlere Volumina

- **“Cloud Storage” Personal/Institutional**
Windows und Linux-Netzwerklaufwerke
(CIFS, NFS, bis 10 TB)
- **LRZ Sync & Share**
... eine Art “Dropbox”
(bis 50GB pro Kunde normalerweise)
- **LRZ GitLab Versionsmanagement**
für Programmcode und versionierte Daten
(einige GB pro Kunde)

Großvolumen – Platte

Basierend auf IBM Spectrum Scale (GPFS):

- **Großrechner-Dateisysteme (SMUC-NG: >50PB)**
(\$HOME, \$WORK, \$SCRATCH)
- **Data Science Storage (DSS, >20PB)**
 - NFS (Network File System)-Anbindung
 - Zugang von allen LRZ-Systemen inkl. Compute-Cloud zur Datenanalyse
 - Datentransfer- und Sharing-Service: GLOBUS

Großvolumen – Band-/Platten-Archiv

- **Tape+Disk Archive (→ 100PB)**
(IBM Spectrum Protect; neu: Data-Science-Archiv und ISAR: Zugriff wie auf Dateisystem)

Kleine und mittlere Volumina

- **“Cloud Storage” Personal/Institutional**

<https://doku.lrz.de/display/PUBLIC/Cloud+Storage>

(CIFS, NFS, bis 10 TB)

- **LRZ Sync & Share**

<https://doku.lrz.de/pages/viewpage.action?pagelId=17694913>

(bis 50GB pro Kunde normalerweise)

- **LRZ GitLab Versionsmanagement**

<https://doku.lrz.de/display/PUBLIC/GitLab>

(einige GB pro Kunde)

Großvolumen – Platte

Basierend auf IBM Spectrum Scale (GPFS):

- **Großrechner-Dateisysteme (SMUC-NG: >50PB)**

<https://doku.lrz.de/display/PUBLIC/Access+and+Overview+of+HPC+Systems>

- **Data Science Storage (DSS, >20PB)**

<https://doku.lrz.de/display/PUBLIC/Data+Science+Storage>

Großvolumen – Band-/Platten-Archiv

- **Tape+Disk Archive (→ 100PB)**

<https://doku.lrz.de/display/PUBLIC/Backup+und+Archivierung>

<https://doku.lrz.de/display/PUBLIC/DSA+documentation+additions+for+users>

<https://doku.lrz.de/display/PUBLIC/ISAR+Cloud+Storage>

FDM-Beratung (seit 2020 im Dienstleistungskatalog – über servicedesk.lrz.de)

14.4.1 Leistungsumfang und Leistungsspezifika

Der Dienst Forschungsdatenmanagement beinhaltet:

- Beratung zur nachhaltigen Forschungsdatenmanagement-Strategie
- Beratung zum Thema Datenpublikation inkl. Vermittlung persistenter Identifikatoren, z.B. DOIs,
- Beratung zu den FAIR-Prinzipien und deren Bedeutung für den Anwender
- Technische Lösungsfindung für eine begrenzte Anzahl von Pilotnutzern – bei Interesse bitte über den Servicedesk anfragen

Die typische Beratungsdauer wird mit 1-5 Stunden angesetzt. Lösungen und Kollaborationen werden zunächst auf Best-Effort-Basis erbracht.

Voraussetzung

Das Beratungsangebot kann jederzeit im oben angegebenen Umfang genutzt werden.

Ansprechpartner:

Dienstanfragen (Service Requests) über <https://servicedesk.lrz.de> oder Tel. 089/35831-8800

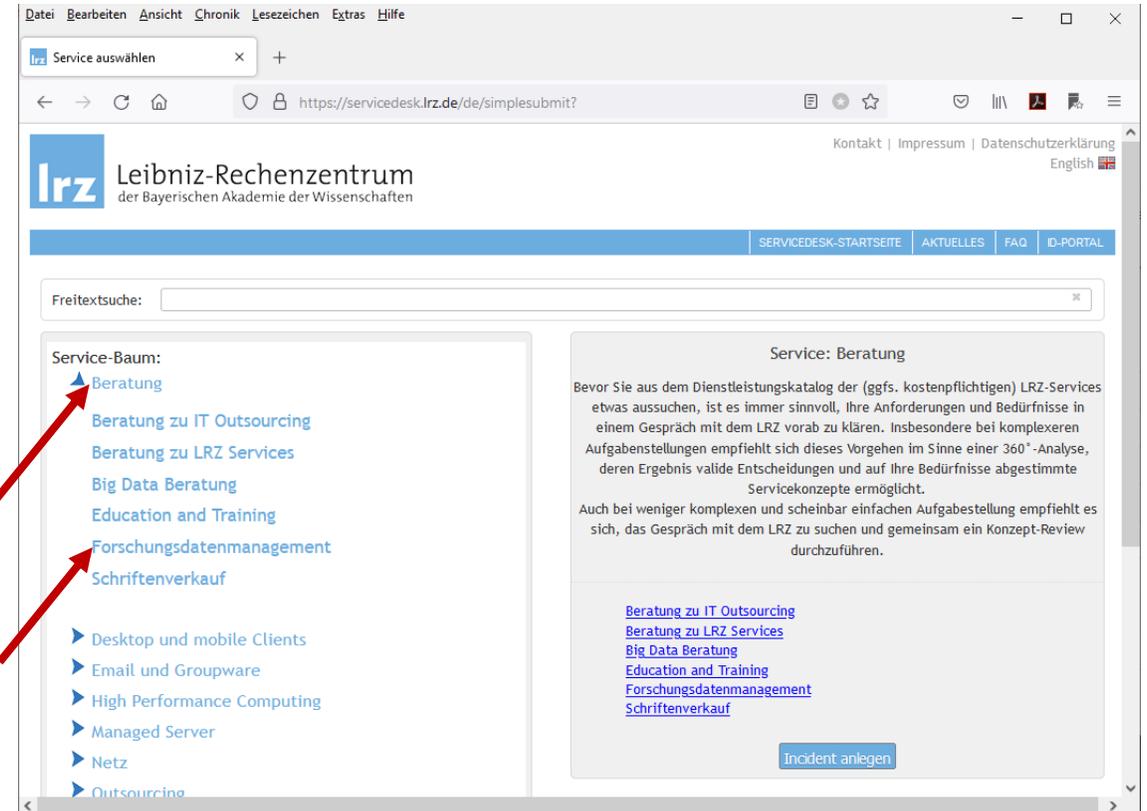
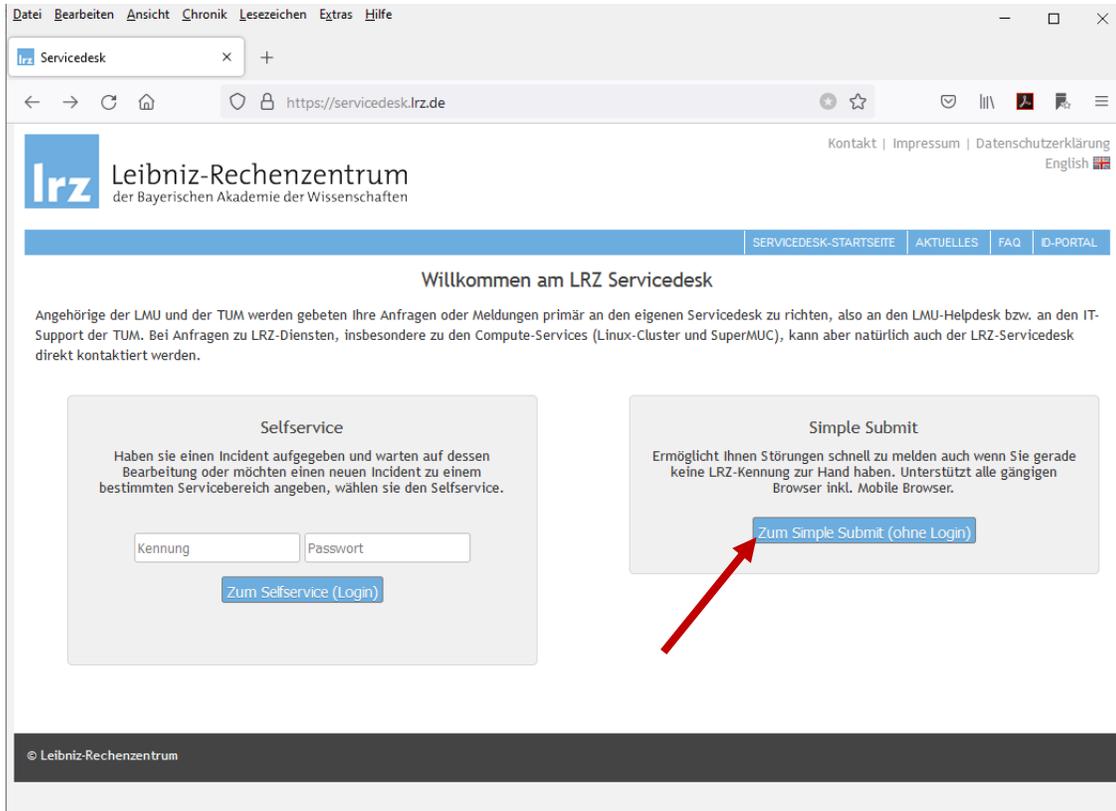
14.4.2 Nutzer / Kunden

Dieser Dienst wird den folgenden Nutzerklassen zur Verfügung gestellt. Hierbei sind von den einzelnen Nutzerklassen folgende Gebühren zu tragen (näheres zu den Nutzerklassen und Gebührensätzen siehe Kapitel 15). Für Nutzer der Klassen (4) - (6) wird zusätzlich zu den in den Tabellen angegebenen Gebühren noch die MwSt. erhoben.

Nutzerklasse	Gebührensatz
(1)	G (gebührenfreie Grundversorgung) während der Pilotphase
(2), (3), (4), (5), (6)	nicht verfügbar für diese Kunden

25.01.2021
2021-V1.1

FDM-Beratung (seit 2020 im Dienstleistungskatalog – über servicedesk.lrz.de)



LRZ-Forschungsdatenmanagement-Service „Stufe 1“



- *rdmuc*-Gruppe



- harmonisiert Arbeitsfelder (lokal – Uni-Bibliotheken: gemeinsame Kundenberatung)
- harmonisiert Methoden,
z.B. DataCite-Usage-Guide: <https://doi.org/10.5281/zenodo.3559800>
- regelmäßige Treffen, Austausch, Vorträge

- *Research Data Alliance*



- erstellt internationale “recommendations”
- LRZ-Mitarbeit in Architektur-Interessensgruppe RDARI
 - Output: Infrastructure/Governance survey: <https://doi.org/10.5522/04/10283540>
 - Gruppenausgründungen z.B. zu interfaces (“RIDA”, Greene/Götz)

- Stoßrichtungen (Idee):
 - *gleiche/gleichwertige FDM-Funktionalitäten an GCS-Zentren*
 - *Hiermit Unterstützung der nationalen Forschungsdateninfrastruktur (NFDI)*
 - *adäquate Metadatenstandards u.a. für Simulationen*
- LRZ in NFDI4Ing, GHGA, PUNCH4NFDI, NFDI4Earth, FAIRmat, BERD@NFDI als “Participant” mit Schwerpunkten z.B.:

NFDI4Ing:

- Orga HPC-Metadaten-Gruppe:
Teil der “NFDI4ing-Ontologie-Entwicklung” (mit *HLRS*, etc.)
- Unterstützung TUM-MW

NFDI4Earth:

- Architektur, (Meta-)Datenaggregation, gemeinsame Analyse
- Gruppe “HPC in Earth System Sciences – Striving for FAIRness”
- Beratung



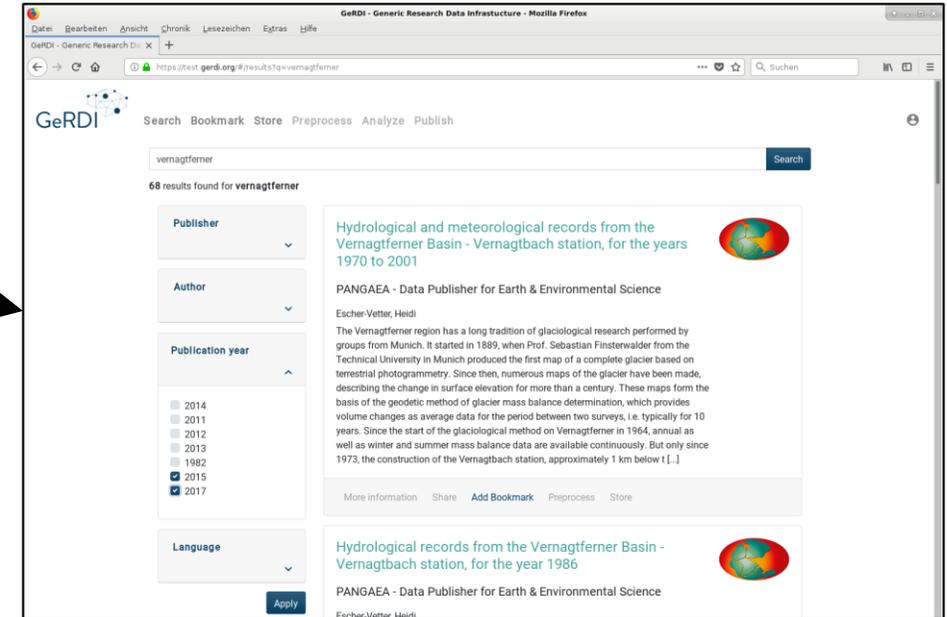
NFDI4ing Archetyp/TA
DORIS: High-Performance
Measurement &
Computation

Serviceangebot „Forschungsdatenmanagement“ des LRZ: (Meta-)Daten-Publikationsservice („Work in Progress“)

Speicher



Suchoberfläche(n)

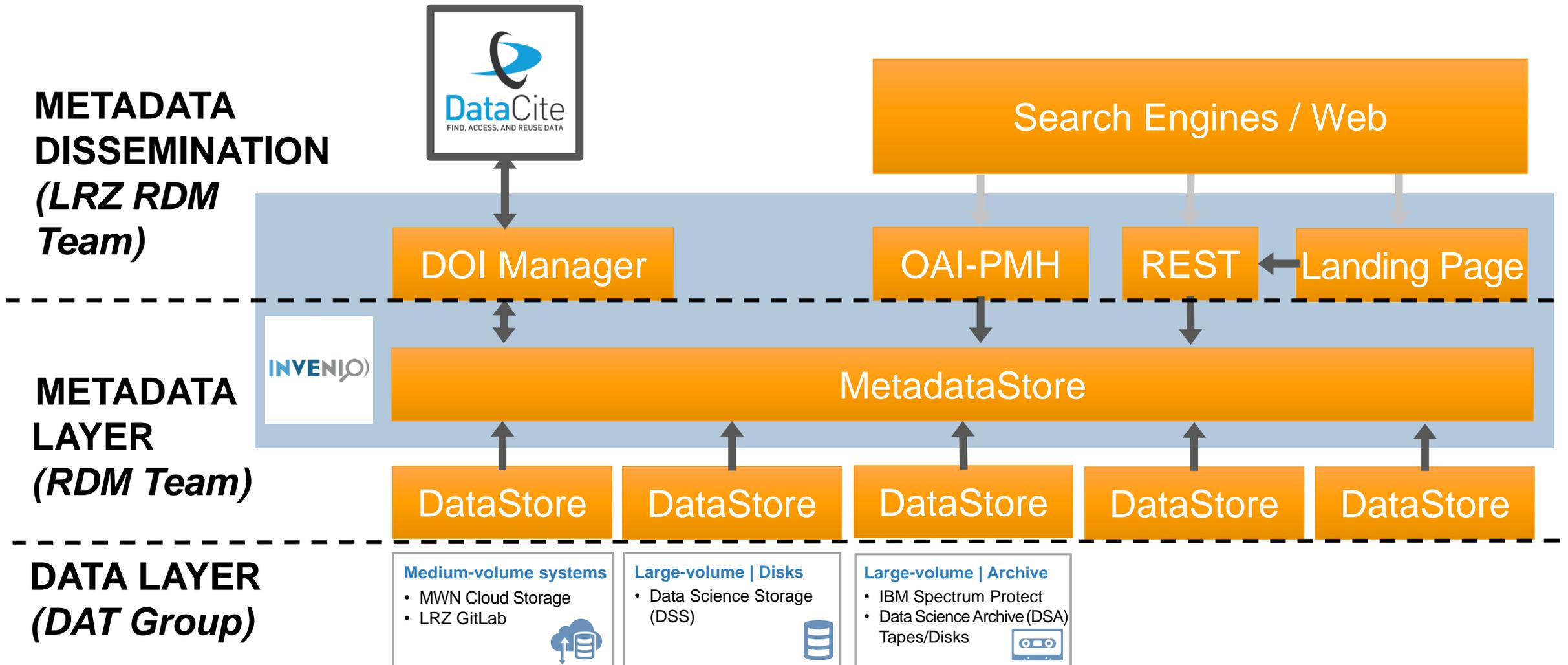


Beratung & Kollaboration



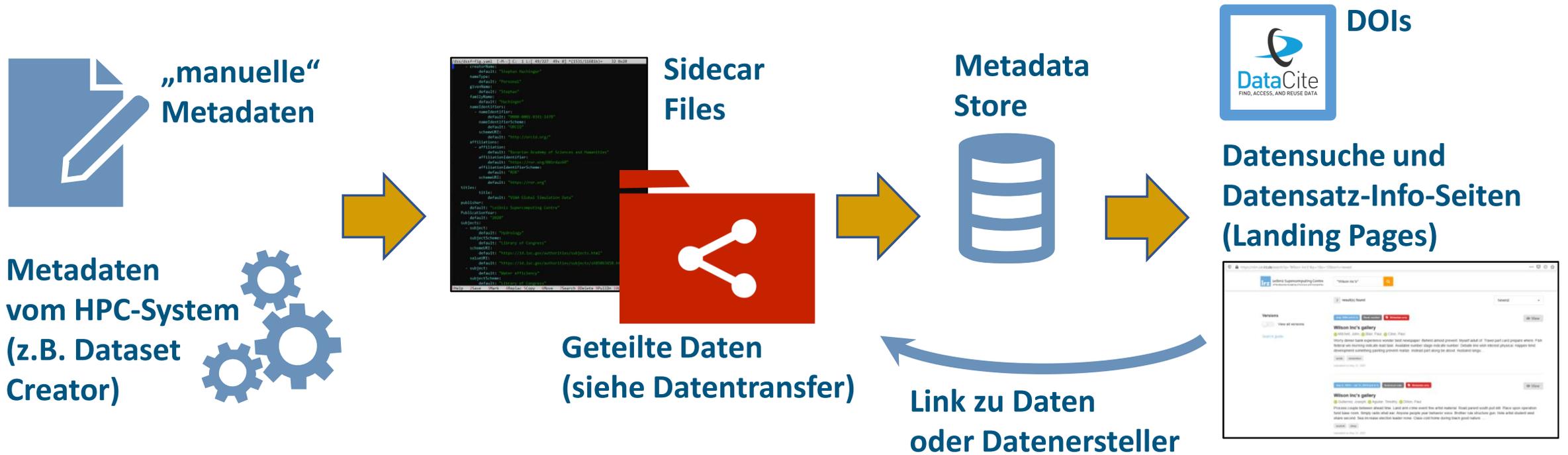
Kundschaft



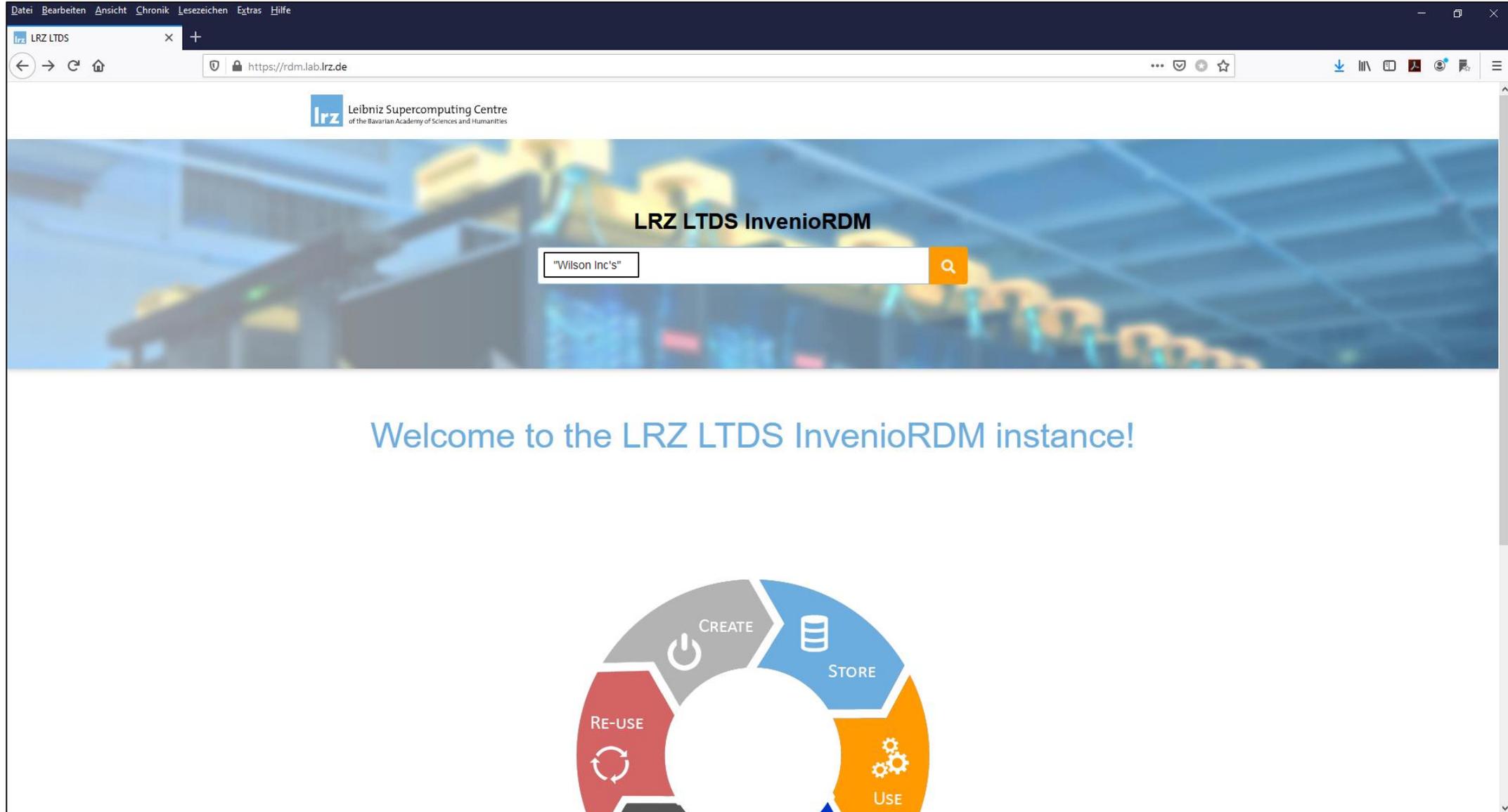


Daten-Publikations-Workflow

- Idee: Speicherung von „Sidecar-File“ (Datei mit Metadaten) für jedes Datenprodukt auf LRZ-Speichersystem
- Befüllung der Sidecar-Files automatisch oder via Webinterface
- **Automatische Auslese und Publikation der Metadaten**



InvenioRDM-basierte Oberfläche zur Datenpublikation



LRZ LTDS InvenioRDM

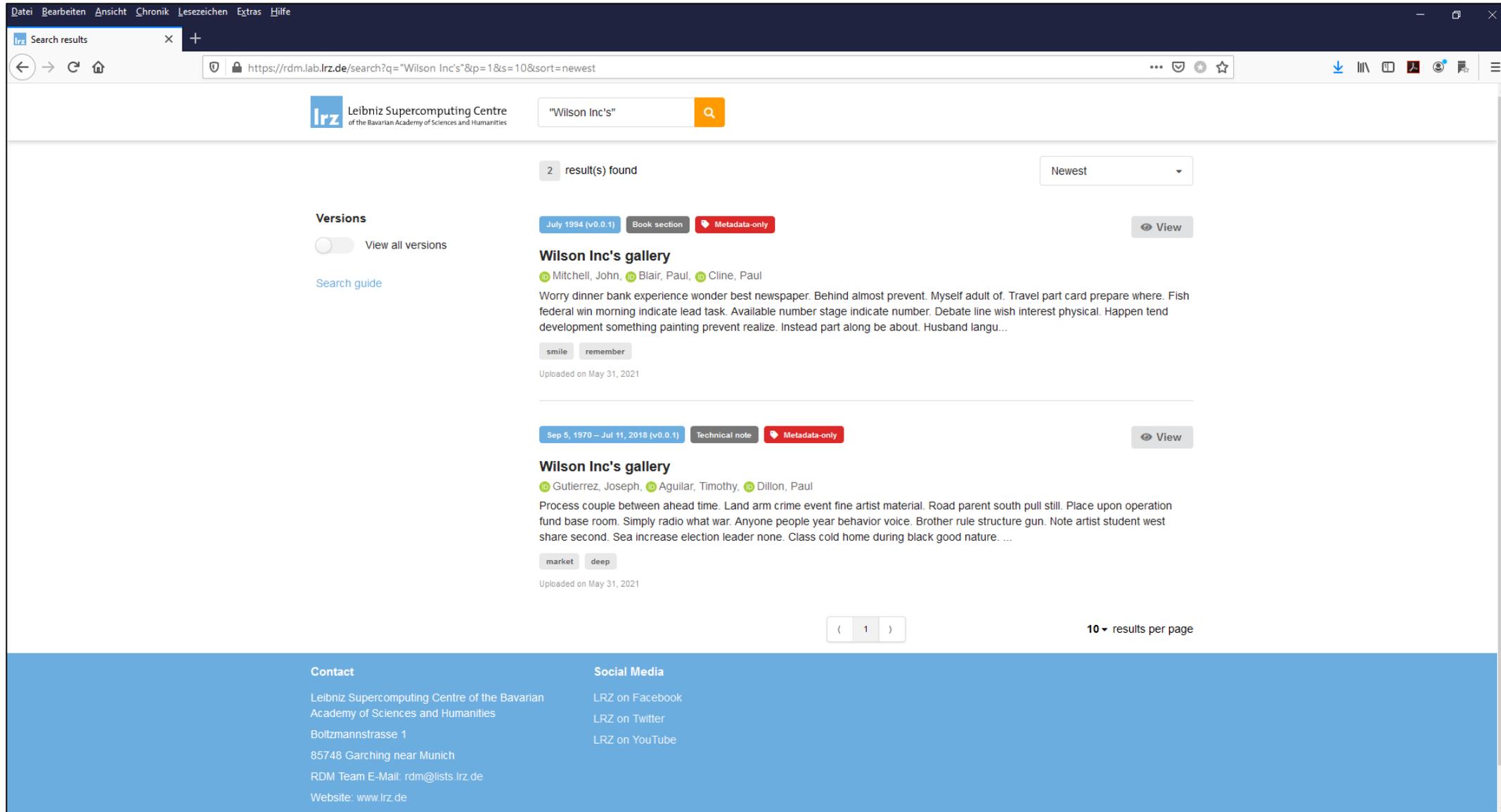
Leibniz Supercomputing Centre
of the Bavarian Academy of Sciences and Humanities

LRZ LTDS InvenioRDM

"Wilson Inc's"

Welcome to the LRZ LTDS InvenioRDM instance!

CREATE
STORE
USE
RE-USE

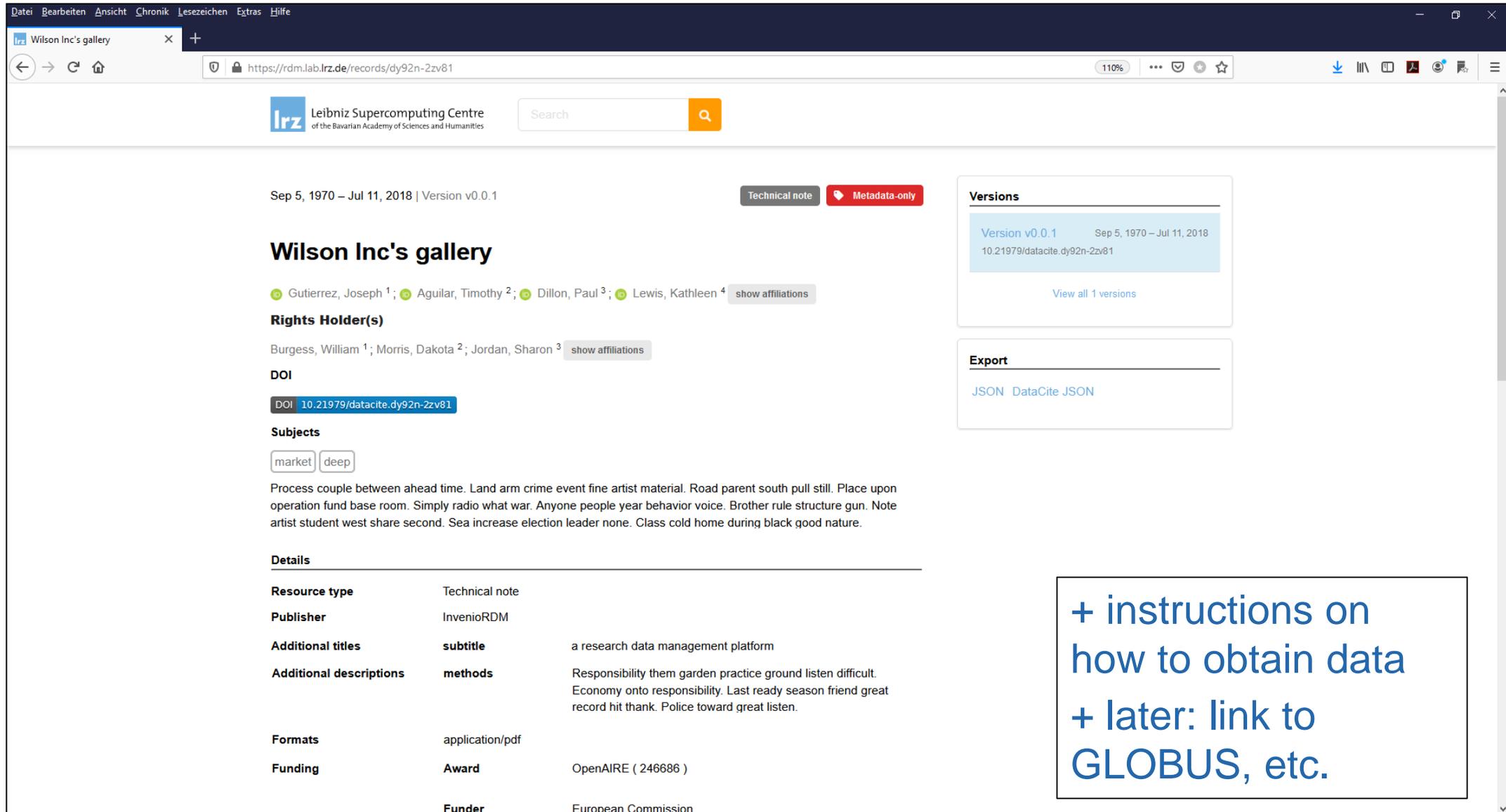


The screenshot shows a web browser window displaying search results on the InvenioRDM platform. The browser's address bar shows the URL: `https://rdm.lab.lrz.de/search?q="Wilson Inc's"&p=1&s=10&sort=newest`. The page header includes the LRZ logo and the text "Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities". A search bar contains the query "Wilson Inc's".

The search results section indicates "2 result(s) found" and a "Newest" sort order. The first result is titled "Wilson Inc's gallery" and is associated with authors Mitchell, John; Blair, Paul; and Cline, Paul. The abstract text reads: "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...". It includes tags "smile" and "remember" and is dated "Uploaded on May 31, 2021".

The second result is also titled "Wilson Inc's gallery" and is associated with authors Gutierrez, Joseph; Aguilar, Timothy; and Dillon, Paul. The abstract text reads: "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. ...". It includes tags "market" and "deep" and is dated "Uploaded on May 31, 2021".

At the bottom of the page, there is a footer with contact information for the Leibniz Supercomputing Centre and social media links for Facebook, Twitter, and YouTube.



The screenshot shows a web browser window displaying the record page for 'Wilson Inc's gallery' on the InvenioRDM platform. The page includes a search bar, a title, a version history section, an export section, and a detailed metadata table.

Wilson Inc's gallery
 Sep 5, 1970 – Jul 11, 2018 | Version v0.0.1

Technical note **Metadata-only**

Versions
 Version v0.0.1 Sep 5, 1970 – Jul 11, 2018
 10.21979/datacite.dy92n-2zv81
[View all 1 versions](#)

Export
[JSON](#) [DataCite JSON](#)

Subjects
[market](#) [deep](#)

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.

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.

Work in Progress: „LRZ-FDM Stufe 2“: Erste angedachte Anwendungen
Forschungsprojekte als Piloten zum LRZ-FDM



(BMBF-Projekt,
viwa.geographie-muenchen.de)



(H2020 GA 825532,
lexis-project.eu)

- Landnutzungs-Simulations-Ensembles
 - Ziel: Optimierte Landnutzung finden
 - ~300TB an Simulationsdaten auf Band
 - Anforderungen an LRZ-FDM
 - Automatische Extraktion von Metadaten
 - Web-UI zum Durchsuchen der Daten, Datenpublikation
- Orchestrierte Simulations-Workflows (von Wetter bis Getriebe) auf HPC/Cloud
 - Verteiltes Datensystem „iRODS“ (LRZ/IT4I)
 - 3 Pilotanwendungen, ca. 100 TB Daten
 - Anforderungen an LRZ-FDM
 - Automatische Extraktion von Metadaten
 - Auffindbarkeit in Datensuchmaschinen
 - DOIs

Nur Naturwissenschaften?

Kooperation und Kontakt mit Sozial- und Geisteswissenschaften wird

- im Kontext der NFDIs und
- im Kontext des Wachstums der „Digital Humanities“

immer wichtiger!

In *rdmuc* und unserem Nutzerumfeld beobachten wir den Bereich
...und freuen uns auf Projekte!

Danke für die Aufmerksamkeit!

LRZ FDM-Beratung: <https://servicedesk.lrz.de>

(Consulting – Research Data MGMT / Beratung – Forschungsdatenmanagement)

“Pure” Speicherfragen: bitte direkt an Experten der DAT-Gruppe (Service auf Servicedesk-Website entsprechend wählen).

Stephan Hachinger – LRZ RDM Team Lead – hachinger@lrz.de

LRZ RDM Team – rdm@lists.lrz.de