

# Die HeFDI-Repositoryn am Beispiel des TUDatalib

Vorstellung von TUDatalib, dem institutionellen Forschungsdatenrepositorium der TU Darmstadt und des Betriebsmodells für weitere Hochschulen innerhalb von HeFDI

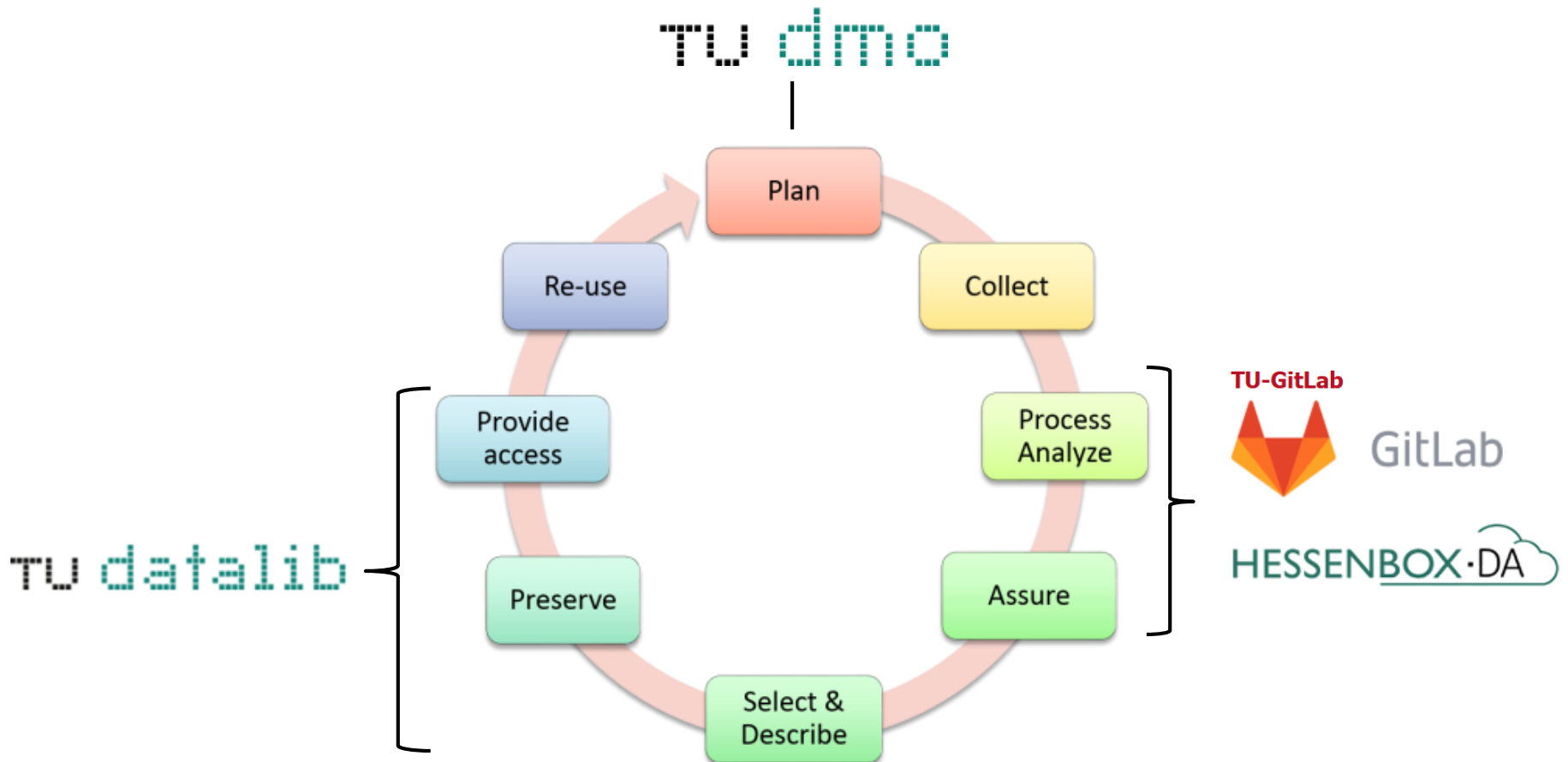
*Gerald Jagusch, ULB Darmstadt*



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



# FDM-Services an der TU Darmstadt



## Präsidiumsbeschluss Ende 2017

„Einführung einer Infrastruktur zur Archivierung und Veröffentlichung von Forschungsdaten an der TU Darmstadt: Beschluss zu Projektressourcen (Personal- und Sachmittel)“

Beinhaltet:

- Beschluss eines von der „AG Forschungsdaten an der TU“ erarbeiteten FDM-Finanzierungsmodells
- Bewilligung von Personal- und Sachmittel zum Auf- und Ausbau von technischer Infrastruktur gemeinsam durch ULB und HRZ (=TUdata) für die Jahre 2018 - 2022

## Eigenschaften



Wahlweise können Fachgebiete mit TUdatalib **Datensätze mit Forschungsdaten** und beschreibende **Metadaten**...

- ... **archivieren** (10 Jahre und länger),
- ... mit beliebigen anderen **teilen**,
- ... zitierfähig **veröffentlichen** inkl. DOI-Vergabe und dauerhafter Speicherung.

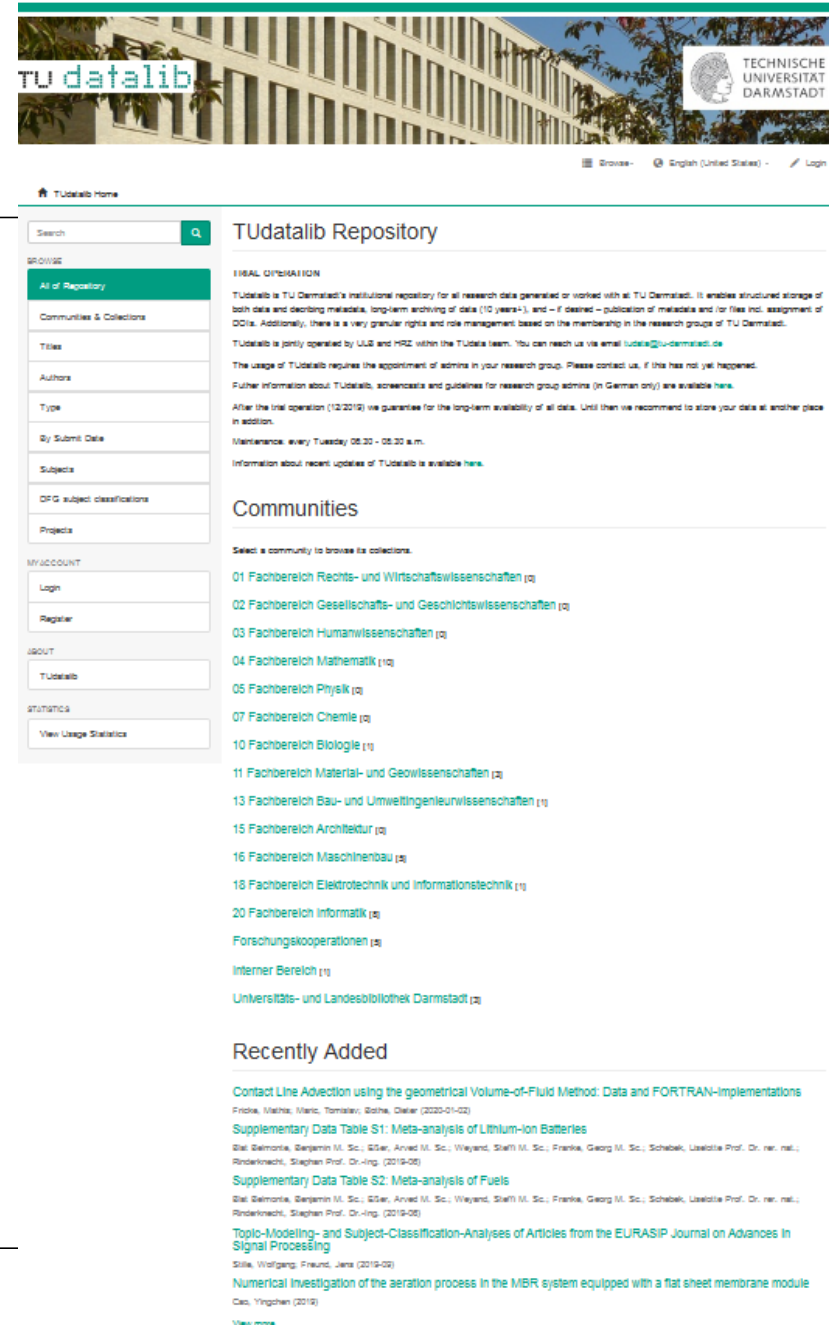
TUdatalib ermöglicht die Integration von reichhaltigen Metadaten, die **Versionierung** von Datensätzen und die **Integration** mit anderen Systemen:

- Zuordnung von Datensätzen zu **Drittmittelprojekten**
- Verlinkung zu **Publikationen** (per DOI)
- Verlinkung von **Personen** mit [ORCID](#)
- Schnittstellen zu Datenspeichern in den Fachbereichen per **REST-API**
- **Automatischer Nachweis** der veröffentlichten Daten in Datensuchmaschinen, OpenAIRE etc.

# Repository TUdatalib

## Eigenschaften

- TUdatalib ermöglicht und erfordert ein detailliertes **Rechte- und Rollenmanagement** basierend auf der Zugehörigkeit zu Fachgebieten der TU
  - je Fachgebiet sind dazu mindestens 2 Admins zu benennen
  - Sie können auch **Review-Prozesse** für Datensätze konfigurieren.
- TUdatalib ist v.a. geeignet für Dateien auf die **selten zugegriffen** wird.
- DOI-Registrierung für Datensatz möglich, wenn Metadaten öffentlich sichtbar sind. Dateien selbst können immer noch zugriffsbeschränkt sein.



The screenshot displays the TUdatalib website interface. At the top, there is a navigation bar with 'TUdatalib Home', a search bar, and language options (English (United States)). The main content area is titled 'TUdatalib Repository' and includes a search bar, a navigation menu, and a list of communities. The 'Communities' section lists various departments like '01 Fachbereich Rechts- und Wirtschaftswissenschaften' through '20 Fachbereich Informatik'. The 'Recently Added' section lists recent publications with titles like 'Contact Line Advection using the geometrical Volume-of-Fluid Method'.

# Forschungsdaten-Repository TUdatalib



Wer kann TUdatalib verwenden?  
Wie funktioniert die Anmeldung?



Der Service steht allen Angehörigen der TU Darmstadt zur Verfügung und kann über den Aufruf der URL [tudatalib.ulb.tu-darmstadt.de](https://tudatalib.ulb.tu-darmstadt.de) direkt verwendet werden.

**Login mit TU-ID** (oder für externe Nutzer mit Registrierung)

- In der Regel werden Sie automatisch Ihrem Fachgebiet zugeordnet und können dann TUdatalib entsprechend der Ihnen vom Fachgebietsadmin eingeräumten Rechte nutzen.
- Externe Nutzende müssen von einem Fachgebietsadmin freigeschaltet werden.
- Nicht angemeldete Nutzende können ausschließlich auf frei verfügbare Datensätze zugreifen.

# Forschungsdaten-Repository TUdatalib

## Entgelte



- Die Nutzung von TUdatalib ist **kostenfrei** für **bis zu 2 TB** Gesamtvolumen neuer Daten **pro Jahr und Fachgebiet / Professur**.
  - Leichtes Monitoring der eigenen archivierten Datenvolumina
- Bei größeren Mengen neuer Daten ist für das **über 2 TB** hinausgehende Datenvolumen eine **einmalige Kostenbeteiligung** durch das Fachgebiet erforderlich.
  - Diese liegt derzeit bei **250 € / TB für die 10-jährige Archivierung**.
  - Die Rechnungsstellung wird vom HRZ abgewickelt.

# Forschungsdaten-Repository TUdatalib

## Struktur in **TUdatalib** und Rollen

### Bereiche = Fachgebiete/Professuren

- Bereichs-Admins (ideal. 2 pro Fachgebiet) verwalten Rollen & Rechte für ihr Fachgebiet

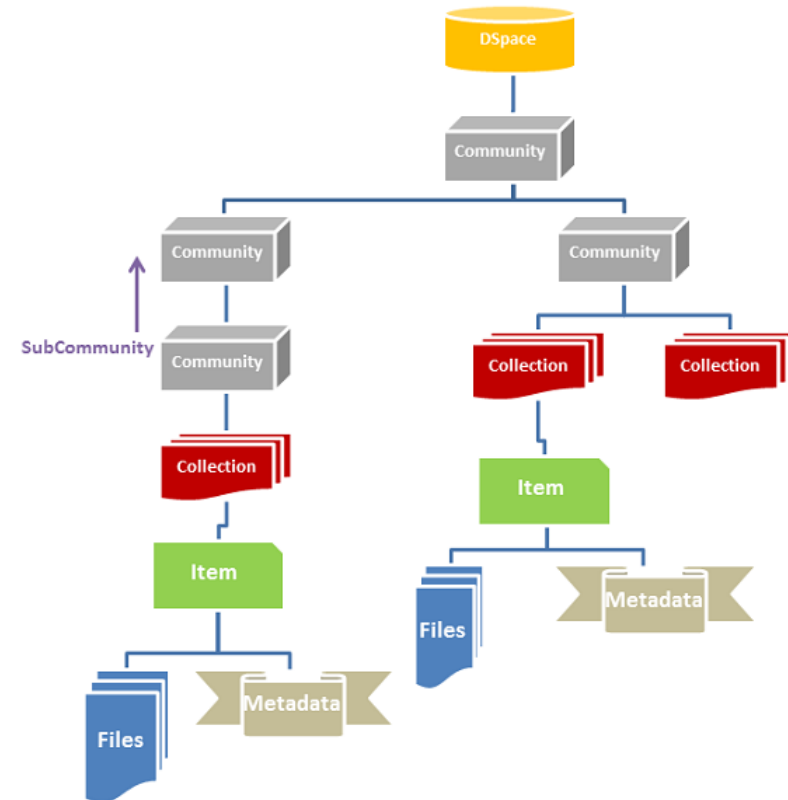
### Unterbereiche und **Sammlungen**

*frei anlegbar durch Bereichs-Admins*

- Sammlungs-Admins verwalten Rollen & Rechte für eine Sammlung und schalten Datensätze frei

### Datensatz (Item)

- Identifier (z.B. DOI)
- Metadaten
- „beliebig“ viele **Dateien (Files)**





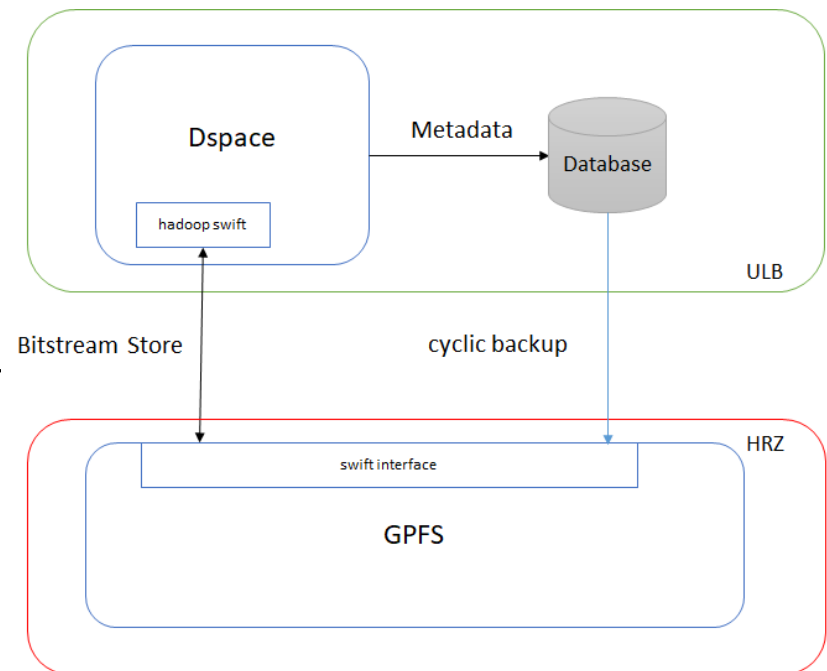
# Forschungsdaten-Repository TUdatalib

## Technischer Hintergrund

### Betrieb durch TUdata

- Software-Betrieb und User Support durch ULB
  - beruhend auf der weit verbreiteten Open-Source-Software DSpace
- Gesicherter Storage durch HRZ
- Die Daten in TUdatalib werden in zwei Teilen gespeichert.
  - Metadaten, Struktur und Berechtigungen in PostgreSQL-Datenbank
  - Bitstream der eigtl. Forschungsdaten in Object Storage (IBM Spectrum Scale / GPFS)

TU datalib



## REST-Schnittstelle



- REST = REpresentational State Transfer
- [https://en.wikipedia.org/wiki/Representational\\_state\\_transfer](https://en.wikipedia.org/wiki/Representational_state_transfer)
- Schnittstelle für die direkte Kommunikation Maschine-Maschine ohne Web-Interface
- <https://tudatalib.ulb.tu-darmstadt.de/rest>
- Bei TUdatalib (derzeit) nur 4 Befehle nötig / möglich:  
GET, POST, PUT, DELETE
- Über Attribute (fast) alles möglich, was in der Weboberfläche von TUdatalib auch geht
- Hohes Automatisierungs- und Standardisierungspotential, v.a. für die Archivierung vieler und großer Daten sowie hoher Datenraten
- Beispielskripte in Python verfügbar unter <https://git.rwth-aachen.de/gerald.langhanke/tudatalib>

# Forschungsdaten-Repositoryn in HeFDI Betriebsmodell



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

## Hintergrund

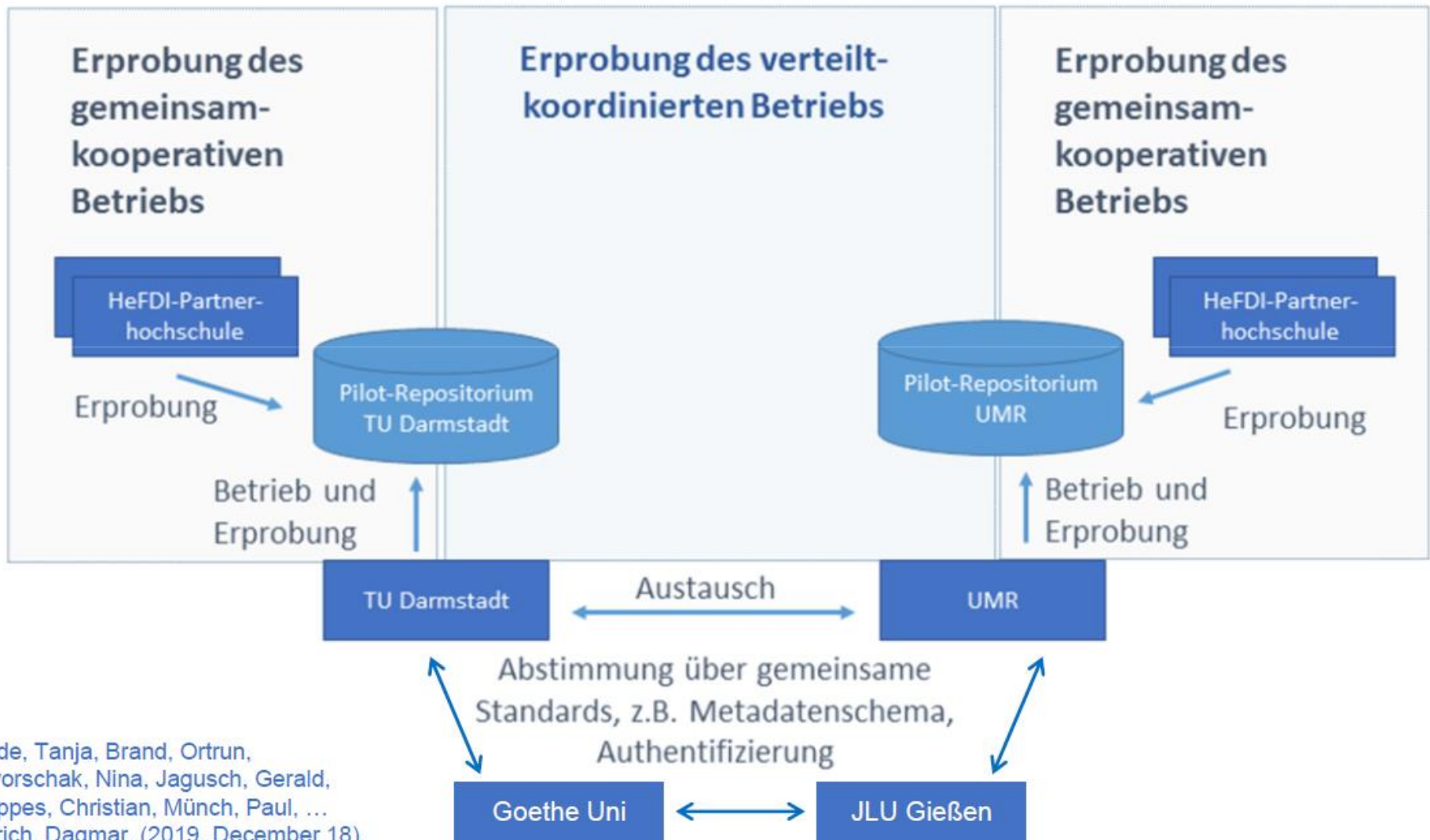
Ziel Forschungsdatenrepositorium für alle HeFDI-Partnerhochschulen

- Allen Standorten eine effiziente Lösung anbieten unter Berücksichtigung lokaler Interessen
- Innerhalb Projekts DSpace als Repositoriumssoftware evaluiert und ausgewählt
- zu Beginn an Uni Marburg und TU Darmstadt erprobt
- mittlerweile auch JLU Gießen (produktiv) und Goethe-Uni Frankfurt (im Aufbau)

Die folgenden Folien basieren auf denen der Breakout-Session #2 des DSpace-Anwendertreffens 2021: „Betriebskonzepte für Repositorienlösungen im Verbund am Beispiel von HeFDI“ von Paul Münch, Birte Cordes, Alexander Vielhauer (UB Marburg):

[https://wiki.lyrasis.org/display/DSPACE/DSpace+Anwendertreffen+2021?preview=/204865807/208340339/20210415\\_DSpace-AT\\_ThementischBetriebskonzept.pdf](https://wiki.lyrasis.org/display/DSPACE/DSpace+Anwendertreffen+2021?preview=/204865807/208340339/20210415_DSpace-AT_ThementischBetriebskonzept.pdf) (abgerufen am 05.07.2021)

# Forschungsdaten-Repositoryen in HeFDI Betriebsmodell



# Forschungsdaten-Repositoryn in HeFDI Betriebsmodell



## Dokument

<https://doi.org/10.5281/zenodo.3582996>

enthält

### **Anforderungsanalyse**

- allgemeine Anforderungen an Repository und Software

### **Ausführungen zum gemeinsam-kooperativen Betrieb**

- konkrete Anforderungen an Partnerhochschulen

### **Ziele und Zusammenarbeit im verteilt-koordinierten Betrieb**

- Leitlinien zur HeFDI-weiten Zusammenarbeit im Bereich Repositoryn

### **Anhang mit Vorlagen zur Kooperationsvereinbarung,**

- Kostenberechnung (Stand 2019), Exit-Strategie

# Forschungsdaten-Repositoryn in HeFDI

## Betriebsmodell



### Gemeinsam-kooperativer Betrieb

#### Betriebskonzept

- Übersicht zentraler Leistungen, Anbindung IDM, Anpassung lokale Bedarfe, Support-/Service-Leistungen

#### Betriebsmodell (mögliche Umsetzung des Betriebskonzepts)

- rechtliche Rahmenbedingungen (siehe Dokumente im Anhang)
- Verteilung der Support/Service-Aufgaben
- Kooperation zw. Hochschulen, z.B. Kommunikationswege, etc.

#### Aktueller Stand der Umsetzung

##### Uni Marburg:

- Konzept bildet die Grundlage für Kooperation Uni Marburg /Uni Kassel
- Einzelne Instanzen / selbe Code-Basis
- Unterschiedliche Anforderungen an Struktur, Rollenmanagement, etc. via Konfiguration

##### Umsetzung TU Darmstadt:

- Für Großteil der hessischen HAWn (h\_da, FUAS, HSRM, HS Fulda, HS Geisenheim)
- Alle in einer Instanz (TUdatalib)

# Forschungsdaten-Repositoryen in HeFDI Betriebsmodell

## Verteilt-koordinierter Betrieb

### Ziele: Synergieeffekte nutzen

- möglichst einheitliche Standards schaffen
  - einheitliche Implementierung Einreichungsprozess:
    - Mindeststandard Metadaten Abfrage
    - Lizenzauswahl-Schritt
    - Upload-Schritt
  - einheitlicher Umgang mit DFG-Fächern:
    - gemeinsam entwickeltes Mapping DFG -> DNB-Sachgruppen
    - gleiche Implementierung des automatischen Mappings in DSpace
  - einheitliche Regelung zur Aufbewahrungsfrist
- möglichst einheitliches Angebot für die Publikation von Forschungsdaten in Hessen schaffen

### Zusammenarbeit durch monatliche Treffen

- gemeinsame Anforderungen an Aufbau/Betrieb Repositoryen:
  - FAIR-Prinzipien, Verwendung Metadaten nach DC/DataCite, etc.