



Universität Stuttgart

**Sibylle Hermann,  
Dorothea Iglezakis,  
Anett Seeland**

# **Make it easy - integration of data description in the research process**

**11. June 2019**

University of Stuttgart



## What the User doesn't like to do

- Publish data because it is not yet common in engineering science
- Spend time with documentation

# What the User Needs

- Manage a lot of data
- Find saved data easily
- Browse data sets
- Change data sets dynamically
- Record metadata easily
- Link results with simulations
- Link data sets from different simulations
- Give controlled access

# Metadata

What our users want to search for (apart from Author, Year)

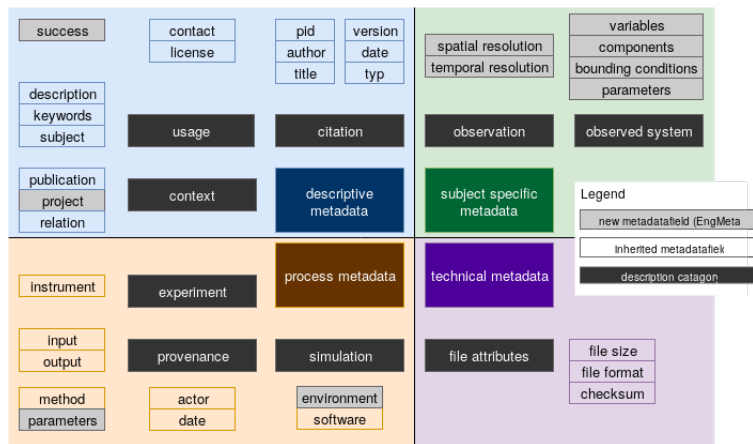
- Variables – measured and controlled
- Parameters of the used method
- Parameters of the observed system

What our users want to document from their research process

- Methods and workflows
- Software and computing environments
- Instruments
- Parameters and assumptions

# EngMeta

## A Metadata Schema for Engineering Science



Schembera & Iglezakis "The Genesis of EngMeta-A Metadata Model for Research Data in Computational Engineering", In: Research Conference on Metadata and Semantics Research, p127–132, 2018, Springer.

# Local Data Management – Prerequisite for Open Data

## Idea

Adding metadata to the data as early in the process and as easy as possible

## Approach

Using a data repository primarily as metadata store and tools around it for smooth interaction

# DaRUS

## Data Repository of the University of Stuttgart

Based on Dataverse

- Open source research data repository software
- Repository hosts multiple virtual archives called Dataverses

The  
**Dataverse**<sup>®</sup>  
Project 

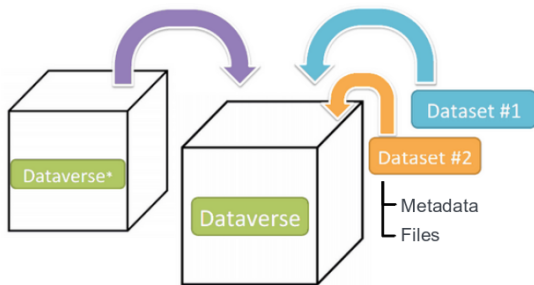
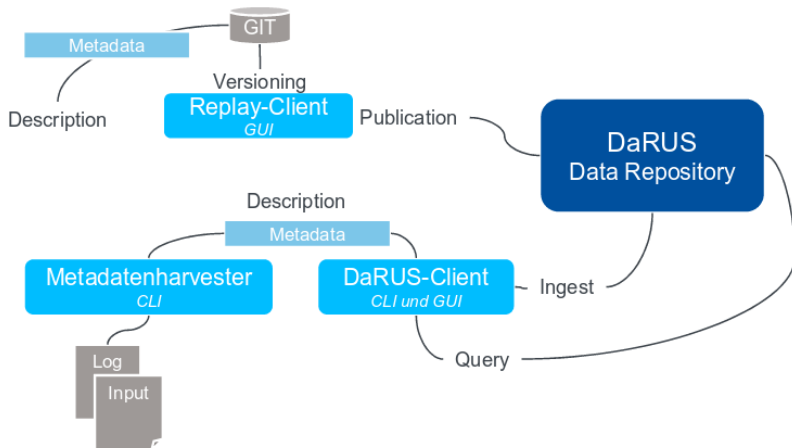


Image: <http://guides.dataverse.org/en/latest/user/dataverse-management.html>, Access: 6/7/2019



# Challenge I: Automation

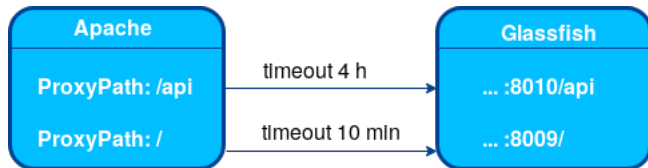
## Ingest of (Meta)data



## Challenge II: Handling of Large Files

Dataverse not designed for large files

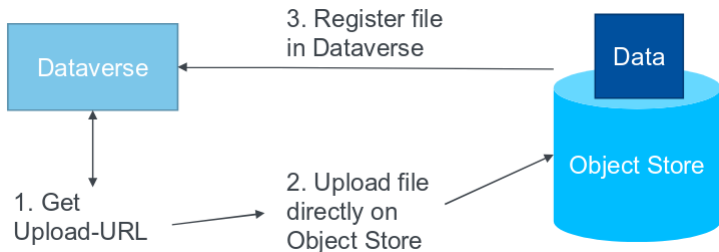
- Users experienced frozen UI and timeouts
- Use REST API for files > 2 GB
- Trade-off between timeout configuration and available threads
- Introduce 2nd thread pool in Glassfish



→ Uploads around 100 GB possible

## Challenge II: Handling of Large Files

- Currently under development













- In planning
  - Connection of object storage to tape library
  - Extend Dataverse to support different storage classes (Download vs Provide-Buttons)

# Outlook: Different Data Overview Needed

1 to 10 of 20 Results

Sort ▾

	<b>turbulent_RFM</b> <span>Draft</span> <span>Unpublished</span>	
May 7, 2019 - Boundary Layers		
Selent, Björn, 2019, "turbulent_RFM", <a href="https://doi.org/10.5072/darus-266">https://doi.org/10.5072/darus-266</a> , DaRUS, DRAFT VERSION		
Simulation of turbulent wall boundary layer		
	<b>turbulent_RFM</b> <span>Draft</span> <span>Unpublished</span>	
May 7, 2019 - Boundary Layers		
Selent, Björn, 2019, "turbulent_RFM", <a href="https://doi.org/10.5072/darus-265">https://doi.org/10.5072/darus-265</a> , DaRUS, DRAFT VERSION		
Simulation of turbulent wall boundary layer		
	<b>turbulent_RFM</b> <span>Draft</span> <span>Unpublished</span>	
May 7, 2019 - Boundary Layers		
Selent, Björn, 2019, "turbulent_RFM", <a href="https://doi.org/10.5072/darus-264">https://doi.org/10.5072/darus-264</a> , DaRUS, DRAFT VERSION		
Simulation of turbulent wall boundary layer		
	<b>turbulent_RFM</b> <span>Draft</span> <span>Unpublished</span>	
May 7, 2019 - Boundary Layers		
Selent, Björn, 2019, "turbulent_RFM", <a href="https://doi.org/10.5072/darus-263">https://doi.org/10.5072/darus-263</a> , DaRUS, DRAFT VERSION		
Simulation of turbulent wall boundary layer		
	<b>turbulent_RFM</b> <span>Draft</span> <span>Unpublished</span>	
May 7, 2019 - Boundary Layers		
Selent, Björn, 2019, "turbulent_RFM", <a href="https://doi.org/10.5072/darus-252">https://doi.org/10.5072/darus-252</a> , DaRUS, DRAFT VERSION		
Simulation of turbulent wall boundary layer		

# Summary

- Starting early in the process means less effort at the end
- To make it easy is still a challenge
- Automation is a key requirement

# Thank you!

## FoKUS

E-Mail: `fokus@izus.uni-stuttgart.de`

URL: `https://www.izus.uni-stuttgart.de/en/fokus/`

## DaRUS

URL.: `https://www.izus.uni-stuttgart.de/en/fokus/darus/`