### **DARIAH-DE**

# Digital Research Infrastructure for the Arts and Humanities

SPONSORED BY THE



Mirjam Blümm Research & Development Department Göttingen State and University Library





### DARIAH-DE – brief overview

- One of 17 members of DARIAH-EU
- Funded by the German Federal Ministry of Education and Research (BMBF)
  - Preliminary phase: 2008 2011
  - Development phase: 2011 2016
  - Operating phase: 2016-2019
  - Fusion with CLARIN-D → CLARIAH-DE: 2019-2021
- DARIAH-DE Cooperation 2019-21
- Used by 80 projects and more than 4.000 scholars



# DARIAH-DE Partner







Niedersächsische Staats- und Universitätsbibliothek Göttingen



Akademie der Wissenschaften und der Literatur Mainz







Herzog August Bibliothek Wolfenbüttel









**MPCDF** 

MAX PLANCK

COMPUTING &

DATA FACILITY









Deutsche Geisteswissenschaftliche Institute im Ausland











**FORSCHUNGSZENTRUM** 



# DARIAH-DE – Objectives

Digital transformation changes research in the Arts and Humanities fundamentally. This concerns:

- Research methods → getting more and more digital
- Research data → are increasingly digital
- Scholars need new abilities
  - > changes of curricula and staffing at universities
- Change of theoretical approaches
  - > must be adapted and modified
- Change in research practices
  - > more interdisciplinary and collaborative





### DARIAH-DE – Mission

Digital transformation and the resulting permanent changes in the Arts and Humanities, can only be solved and realized collectively, interdisciplinary and collaboratively.

In order to support this process and to enable it throughout Germany, DARIAH-DE was founded: to pool competences and skills, to pursue actively research in this area and, above all, to build up a digital research infrastructure on the basis of research requirements by the communities.



# DARIAH-EU Working Groups DARIAH-DE Engagement

SPONSORED BY THE











# Defining cloud infrastructure services for DH

#### Chairs:

Eveline Wandl-Vogt (DARIAH-AT), René van Horik (DARIAH-NL), Tibor Kàlman (DARIAH-DE)

#### Aim:

minimize the gap between the needs of Arts and Humanities research and the cloud- and grid services as offered by different providers

#### **Activities:**

- collecting and analysing requirements (using surveys, interviews),
- promoting existing services and opportunities (by prodividing demonstrators and by providing guidance and training)
- liaising (by discussing conditions / guidelines for DARIAH)



# Federated Identity Management for DARIAH (FIM4D)

#### Chairs:

Peter Gietz (DARIAH-DE), Lars Wieneke (DARIAH-LU)

#### Aims:

- Dissemination of the DARIAH AAI
- pilots to integrate services from different DARIAH countries

#### Activities/Challenges:

- Better integration of the campus Identity Provider
- Better consistency in the usage of Attributes
- Better transparency of the levels of trust
- Interoperability with other AAI infrastructures in Arts, humanities and social sciences



# Text and Data Analytics

#### Chairs:

Fotis Jannidis (DARIAH-DE), Mike Kestemont (DARIAH-BE)

#### Aims:

 developing and demonstrating methodologies for applying the state of the art in NLP to Humanities research questions

#### **Activities:**

- provide guides to specific best practice
- inventory successful applications of prominent tools
- demonstrate specific examples of services which can be used in Digital Humanities scholarship



# Impact factors and success criteria

Chair: Juliane Stiller (DARIAH-DE)

#### Aims:

explore the impact of digital research infrastructures on different areas ranging from research practices to societal implications

#### **Activities:**

- meassure impact of the work of DARIAH and its affiliated projects in the Digital Humanities domain and beyond.
- Collect and analyse qualitative and quantitative impact measures and key performance indicators for Digital Humanities, focusing on the implications for humanities, the research process and society
- Link: Impactomatrix on impact areas, influencing factors and criteria to measure change: <a href="https://dariah-de.github.io/lmpactomatrix/">https://dariah-de.github.io/lmpactomatrix/</a>



### **DARIAH-DE**

Activities, Tools & Services

SPONSORED BY THE













# By the community For the community



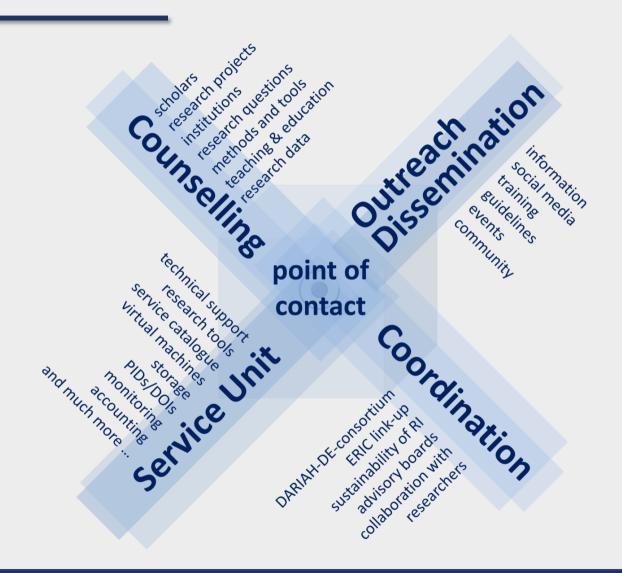


Architecture of participation





# **DARIAH-DE** Coordination Office







# DARIAH-DE – four main tasks

#### **TEACHING**

- Workshops on methods, expert colloquia, Summer Schools
- Coordination of national and international curricular developments

#### **RESEARCH DATA**

- Best practices for metadata, standardised exchange of data, ontologies
- Development of generic search, collection and schema registry, repository

#### **RESEACH**

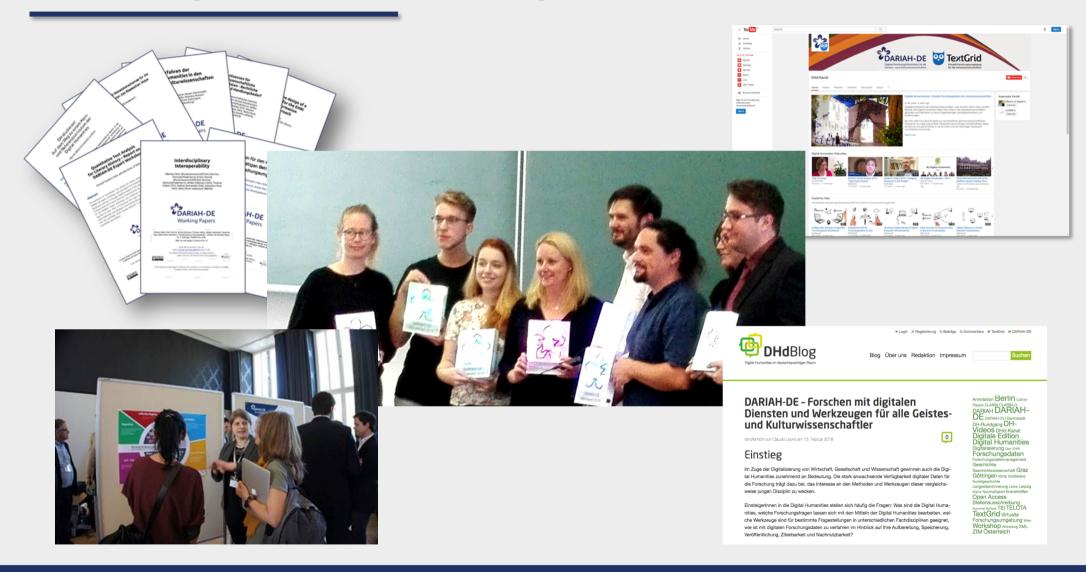
- DH methods and practices
- Use Cases: Annotation, Big Data
- Tools and services
- **Bibliography Doing Digital Humanities**

### **TECHNICAL INFRASTRUCTRUE**

- collaborative research environments
- virtual machines
- Monitoring
- authentication and authorization infrastructure etc.



# Teaching and Community

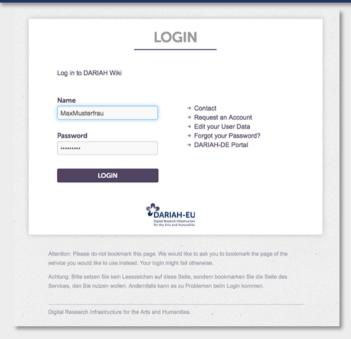


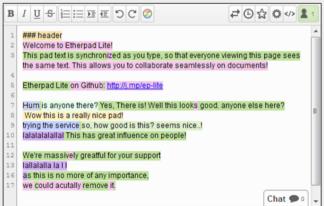




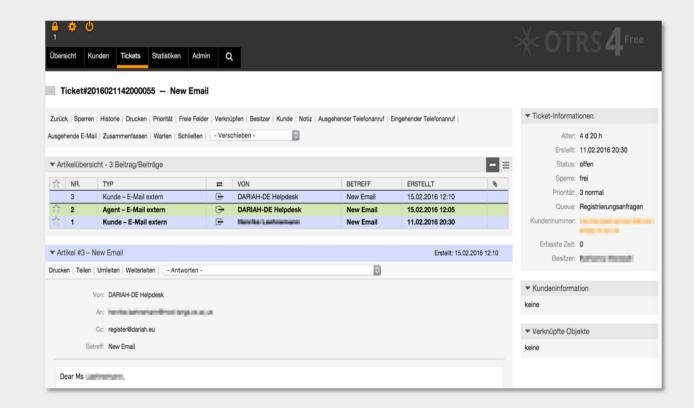


### Technical Infrastructure







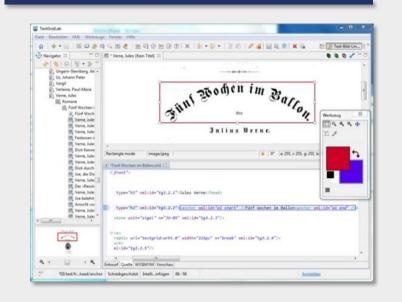


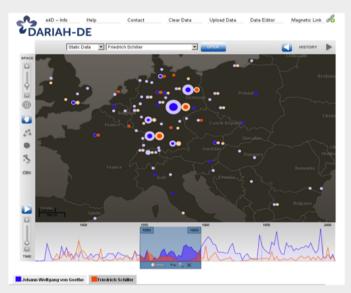


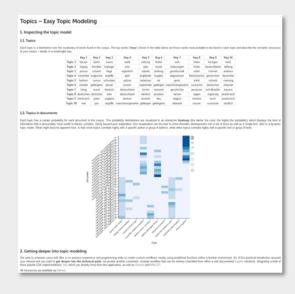


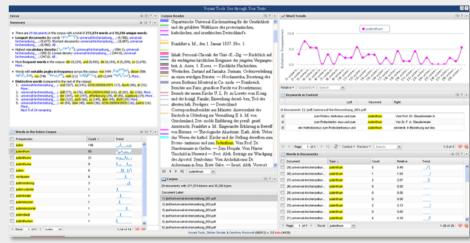


# Research – Scholary Tools









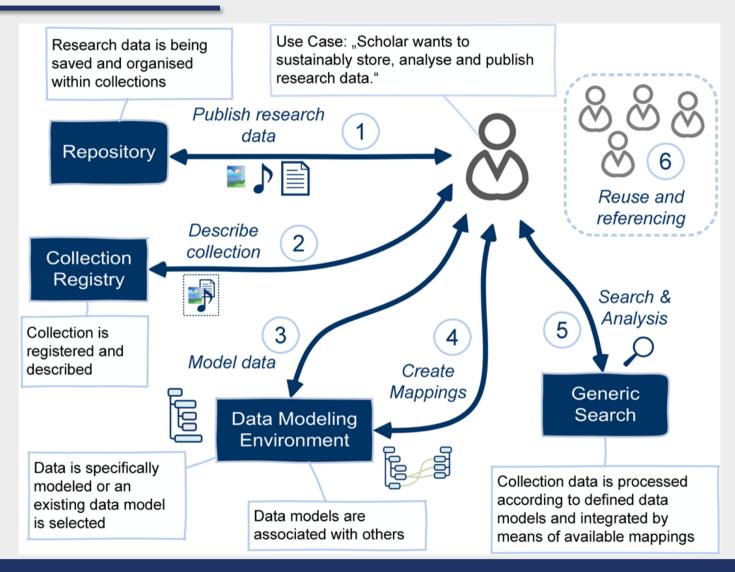








### Data Federation Architecture







### Research Data – Arts and Humanities

#### Definition Research Data

By research data in the Arts and Humanities is meant all those sources and outputs that are **based on a research question**, collected either by scholars, libraries, archives or any other institution, which may be described, evaluated and/or generated and stored in **machine-readable form** for **analyzing**, **archiving** purposes, **citability**, and for further **processing**.



# Research Data – Arts and Humanities I

Research data in the arts and humanities are:

All sources and outputs which are **based on a research question, collected** either by scholars, libraries, archives, or any other institution and which may be **described, evaluated** and/or **generated**, and **stored** in **machine-readable** form for analyzing, archiving purposes, citability, and for further processing.





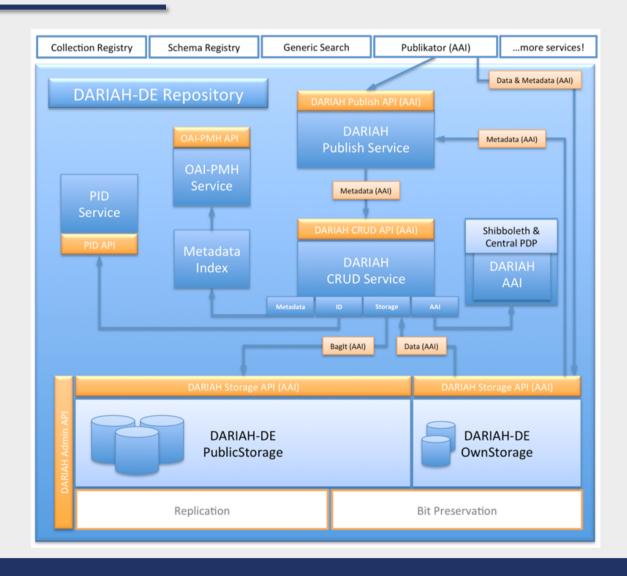
### Research Data – Arts and Humanities II

- Metadata, bibliographical data, finding aids
- Digital and/or digitized data and/or digital representation of analog data
- Digital objects
- Full text, transcripts
- Enriched full text
- Images, movies, music & notes
- Authority files, controlled vocabularies, ontologies
- etc.





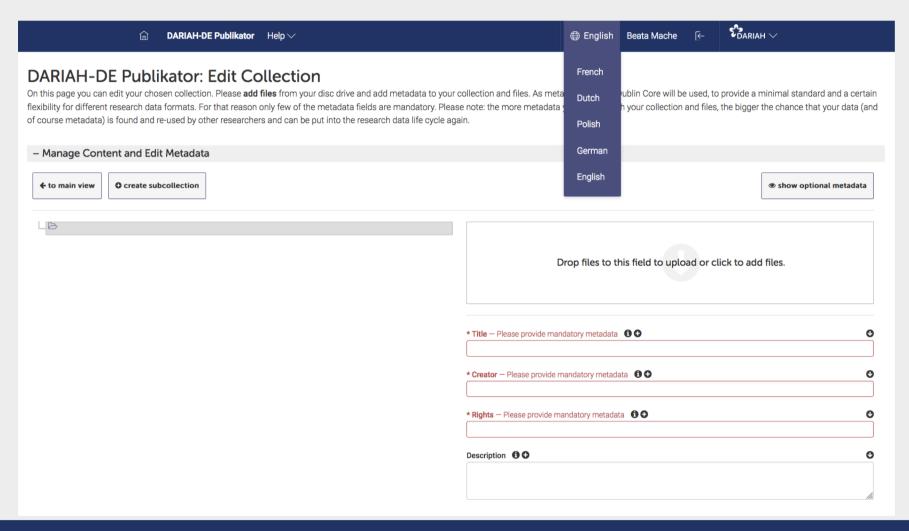
# DARIAH-DE Repository







# DARIAH-DE Publikator: data capturing









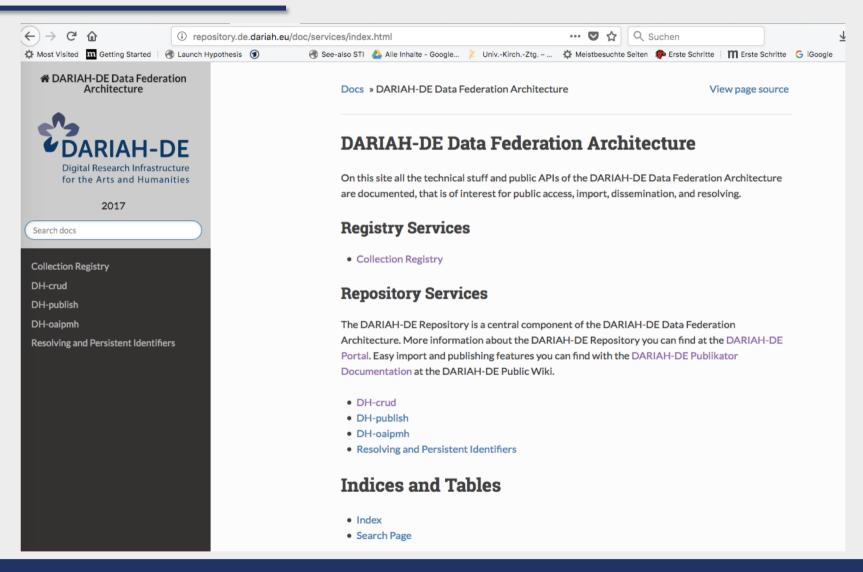
### Your research data...

- is safely stored in the repository,
- gets a Persistent Identifier (DOI),
- is so permanently referenceable and quotable,
- is publicly accessible,
- can be described in the Collection Registry as a collection and
- is finally searchable in the Generic Search.

This means that your research data is included in the research data lifecycle and thus is available for subsequent use.



### **Documentation**

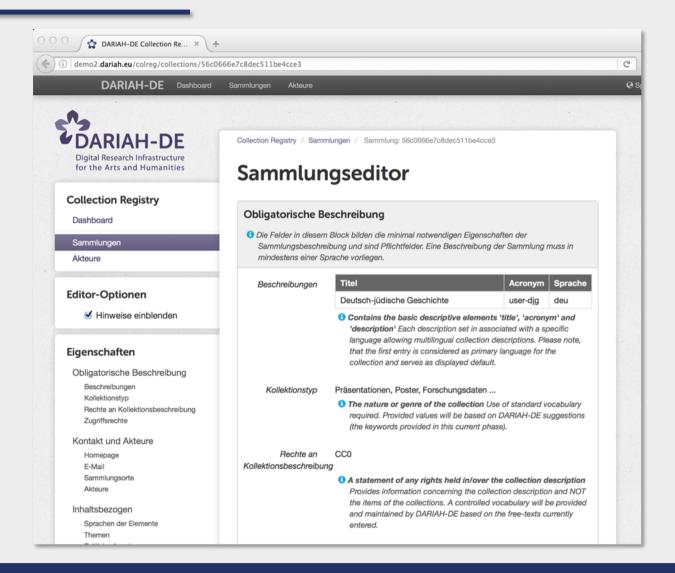






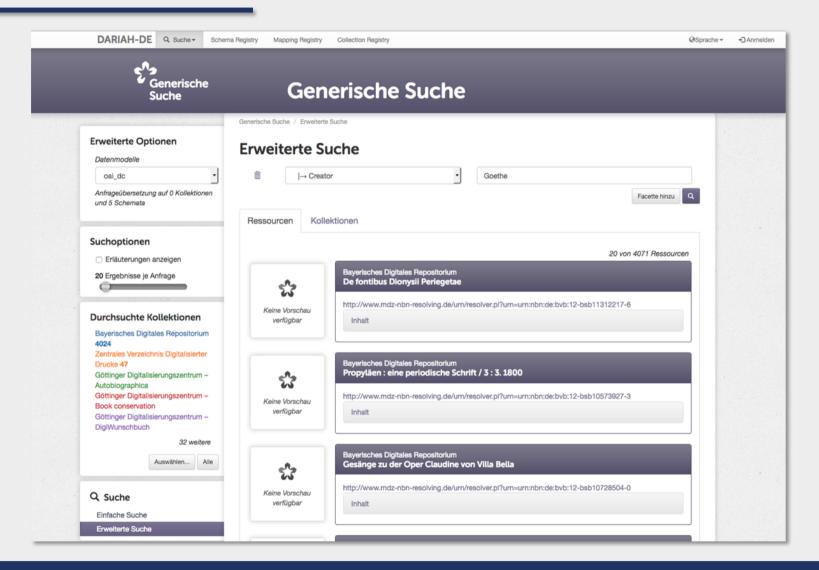


# DARIAH-DE Collection Registry





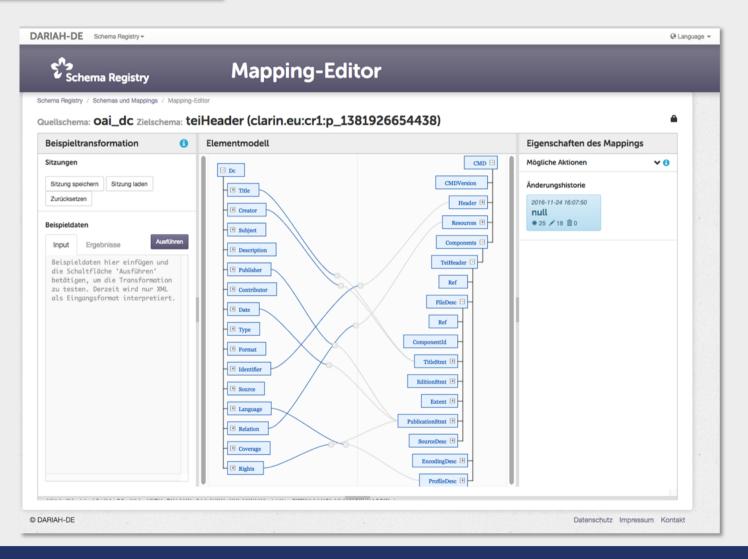
# **DARIAH-DE** Generic Search







# Data Modeling Environment







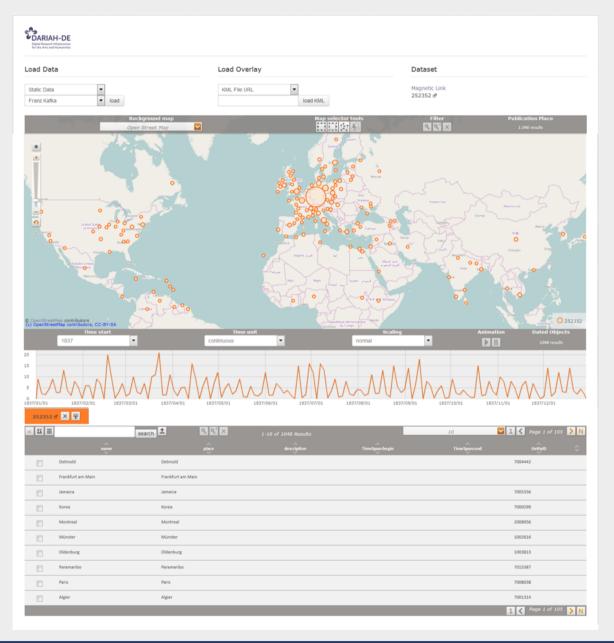
### **DARIAH-DE Geo-Browser**

A tool for spatio-temporal visualisation

Map(s)

**Timeline** 

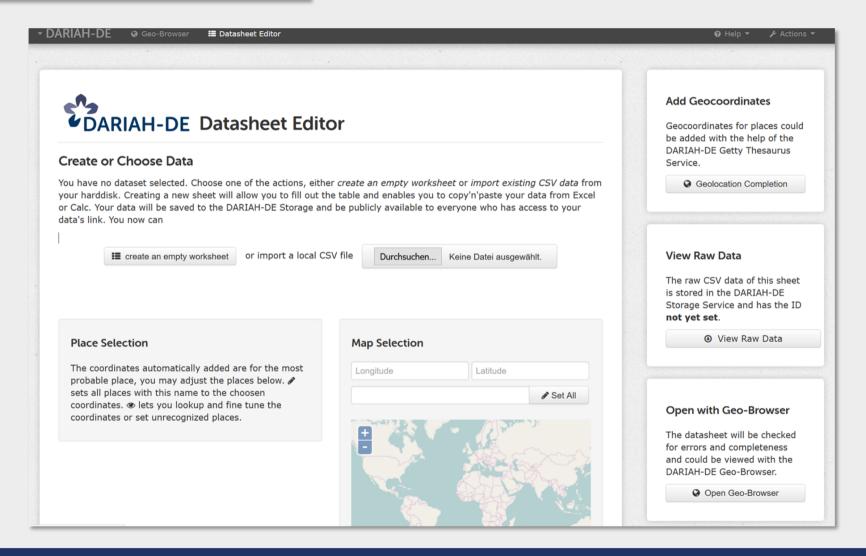
**Documentation** of data





### **DARIAH-DE** Datasheet Editor

A tool for creation of spatio-temporal datasets





# Thank you for your attention!

Questions?

SPONSORED BY THE



bluemm@sub.uni-goettingen.de





