DARIAH-DE
Digital Research Infrastructure for the Arts and Humanities

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)
  - 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
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
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:
- measure 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: https://dariah-de.github.io/Impactomatrix/
DARIAH-DE
Activities, Tools & Services
Architectures of participation

By the community
For the community
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

RESEARCH

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

TECHNICAL INFRASTRUCTURE

- Collaborative research environments
- Virtual machines
- Monitoring
- Authentication and authorization infrastructure etc.
Teaching and Community
Technical Infrastructure
Research – Scholary Tools
Data Federation Architecture

Research data is being saved and organised within collections.

Use Case: "Scholar wants to sustainably store, analyse and publish research data."

1. Publish research data
2. Describe collection
3. Collect data
4. Create mappings
5. Search & Analysis
6. Reuse and referencing

Repository

Collection Registry

Collection is registered and described

Model data

Data Modeling Environment

Data is specifically modeled or an existing data model is selected

Data models are associated with others

Collection data is processed according to defined data models and integrated by means of available mappings
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 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

DARIAH-DE Publikator: Edit Collection

On this page you can edit your chosen collection. Please add files from your disc drive and add metadata to your collection and files. As metadata adds flexibility for different research data formats. For that reason only few of the metadata fields are mandatory. Please note: the more metadata (of course metadata) is found and re-used by other researchers and can be put into the research data life cycle again.

- Manage Content and Edit Metadata

Show optional metadata

Drop files to this field to upload or click to add files.

Title — Please provide mandatory metadata

Creator — Please provide mandatory metadata

Rights — Please provide mandatory metadata

Description

To main view | Create subcollection
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 Data Federation Architecture

On this site all the technical stuff and public APIs of the DARIAH-DE Data Federation Architecture are documented, that is of interest for public access, import, dissemination, and resolving.

Registry Services

- Collection Registry

Repository Services

The DARIAH-DE Repository is a central component of the DARIAH-DE Data Federation Architecture. More information about the DARIAH-DE Repository you can find at the DARIAH-DE Portal. Easy import and publishing features you can find with the DARIAH-DE Publikator Documentation at the DARIAH-DE Public Wiki.

- DH-crud
- DH-publish
- DH-oalpmh
- Resolving and Persistent Identifiers

Indices and Tables

- Index
- Search Page
DARIAH-DE Collection Registry

Collection Registry
Dashboard
Sammlungen
Akteure

Editor-Optionen
- Hinweise einblenden

Eigenschaften
Obligatorische Beschreibung
Beschreibungen
Kollektionsname
Rechte an Kollektionsbeschreibung
Zugriffsberechtigungen

Obligatorische Beschreibung

<table>
<thead>
<tr>
<th>Title</th>
<th>Acronym</th>
<th>Sprache</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deutsch-jüdische Geschichte</td>
<td>user-djk</td>
<td>deu</td>
</tr>
</tbody>
</table>

Contains the basic descriptive elements 'title', 'acronym' and 'description'. Each description set is associated with a specific language allowing multilingual collection descriptions. Please note, that the first entry is considered as primary language for the collection and serves as displayed default.

Kollektionsname
Präsentationen, Poster, Forschungsdaten ...

Rechte an Kollektionsbeschreibung
CCD

A statement of any rights held in/over the collection description. Provides information concerning the collection description and NOT the items of the collections. A controlled vocabulary will be provided and maintained by DARIAH-DE based on the free-texts currently entered.
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

Create or Choose Data
You have no dataset selected. Choose one of the actions, either create an empty worksheet or import existing CSV data from your hardisk. Creating a new sheet will allow you to fill out the table and enables you to copy’n’paste your data from Excel or Calc. Your data will be saved to the DARIAH-DE Storage and be publicly available to everyone who has access to your data’s link. You now can

- create an empty worksheet
- import a local CSV file

Place Selection
The coordinates automatically added are for the most probable place, you may adjust the places below. sets all places with this name to the choosen coordinates. lets you lookup and fine tune the coordinates or set unrecognized places.

Map Selection

Open with Geo-Browser
The datasheet will be checked for errors and completeness and could be viewed with the DARIAH-DE Geo-Browser.

Add Geocoordinates
Geocoordinates for places could be added with the help of the DARIAH-DE Getty Thesaurus Service.

View Raw Data
The raw CSV data of this sheet is stored in the DARIAH-DE Storage Service and has the ID not yet set.
Thank you for your attention!

Questions?

bluemm@sub.uni-goettingen.de