**SCHMUNCK — DARIAH**

<0 images>

**DARIAH-DE: Building a Distributed Research Data Repository for a Digital Humanities Network**

Schmunk, S., Funk, S. and Grupe, N.

DARIAH-DE is the German contribution to the ESFRI project DARIAH-EU—Digital Research Infrastructure for the Arts and Humanities. DARIAH-EU1 is a European Digital Humanities Network and infrastructure funded by the European Commission, which has several national subordinate projects—for example, DARIAH-DE, funded by national research agencies such as the BMBF in Germany.2 DARIAH-DE supports researchers in the humanities working with digital resources and methods in building a digital infrastructure for research tools and data services, as well as coordinating and (if necessary) further developing existing study and training courses in the digital humanities. The central objective of DARIAH-DE is to enable the interoperability of tools and research data. Following internationally valid and accepted standards and policies, DARIAH-DE aims to ensure the long-term preservation and future use of research data. One of DARIAH-DE’s main activities is to build a research data repository for the arts and humanities based on a distributed and generic infrastructure.

This repository will be made available to associated research projects such as EHRI—the European Holocaust Research Infrastructure—as well as individual researchers and projects who want their data to be stored in an easily accessible, persistent, citable long-term archive. A main focus is, therefore, accommodating different project and researcher needs. To achieve this, DARIAH-DE is working together with TextGrid.3 TextGrid is a virtual research environment, consisting of the modular software TextGridLab and the repository TextGridRep. The DARIAH-DE Repository is based on the code used in the TextGrid Repository and will be implemented with a number of services tailored for DARIAH-DE’s specific uses, such as an AAI (Authorisation and Authentification Infrastructure). Recently the DARIAH-DE AAI and the DARIAH-DE Storage API4 for saving data on a bit preservation level were merged such that the storage layer is distributed and synchronized between the participating computing centres. This way the infrastructure is ensured to not only work as storage space for static data—that is, thus be publicly accessible and citable for the long term—but in addition gives the option of storing dynamic data—optionally secured by AAI and constantly updated through regular use.

The data can be accessed through APIs to render it reusable by other tools and services such as the DARIAH-DE Collection Registry, which contains information about arbitrary data repositories and their collections. The Generic Search developed in DARIAH-DE crawls and indexes the collection registry’s collections and provides easy access and retrieval, while the DARIAH-DE Schema Registry provides XML-schemas for mapping and metadata crosswalks.

DARIAH-DE provides not only a repository but a central component of a distributed research infrastructure. The DARIAH-DE project shows how all those services, APIs, dynamic and static repositories, and clients are integrated into a digital humanities network of research data and researchers in the digital humanities, using different workflows and approaches for working with and publishing research data, all based on the data stored in the DARIAH-DE Repository. And all of the developed and used services and tools are reusable in different project contexts.

In addition to the poster presentation at DH 2015 in Sydney, scientific experts and developers will perform a multimedia demonstration of the data ingest, with a specific focus on the humanities scientists’ options to archive, provide, and reuse research data and how to model the necessary workflows and processes. In particular it will be shown how researchers can ingest data into the repository, but also how different research tools get access and ingest them after they analyse or enrich them.

More information about the German project DARIAH-DE and the European project DARIAH-EU can be found at https://de.dariah.eu/ (German) and at https://www.dariah.eu (English).

**Notes**

1. http://www.dariah.eu/.

2. Federal Ministry of Education and Research, http://www.bmbf.de/en/.

3. http://textgrid.de/en/.

4. DARIAH Storage API—A Basic Storage Service API on Bit Preservation Level, http://handle.gwdg.de:8000/11858/00-1734-0000-0009-FEA1-D.