Combining Metadata Infrastructures: Working on a Metadata Federation between DARIAH-DE and CLARIN

Over the past years great effort went into the establishment of research infrastructures for the Humanities and Social Sciences. As a consequence of the diversity of their targeted research fields and communities, miscellaneous infrastructure projects have developed unique solutions for overlapping target groups. This especially holds for describing available resources by structured metadata and providing them to a wider audience in a user-friendly fashion. For this purpose, DARIAH-DE and the CLARIN project have developed metadata infrastructures that support users in all relevant tasks of the typical lifecycle of a metadata record. To overcome the gap between both approaches, an evaluation of potentially helpful procedures was carried out and concrete solutions for cross-project integration of metadata records were implemented. As a result, the visibility of resources is extended to tools and workflows of both projects, giving users more options and more choices when working with metadata records.

Motivation and Goal

Over the past years great effort went into the establishment of research infrastructures for the Humanities and Social Sciences. As a consequence of the diversity of their targeted research fields and communities, miscellaneous infrastructure projects have developed unique solutions for overlapping target groups. This especially holds for describing available resources by structured metadata and providing them to a wider audience in a user-friendly fashion. For this purpose, DARIAH-DE and the CLARIN project have developed metadata infrastructures that support users in all relevant tasks of the typical lifecycle of a metadata record. To overcome the gap between both approaches, an evaluation of potentially helpful procedures was carried out and concrete solutions for cross-project integration of metadata records were implemented. As a result, the visibility of resources is extended to tools and workflows of both projects, giving users more options and more choices when working with metadata records.

Research Infrastructures: Different Structures for Comparable Purposes

![Research Infrastructures Diagram]

Implementation

- Handling of and access to resources & metadata affected by focus groups and primary use cases
- Central interface for access to MD records: OAI-PMH
- Integration models based on DME metamodel
- Services for data publishing, registration and modeling
- Optimized for domain- and project-specific views
- Access to CLARIN MD:
  - Registration of CMDI profiles (subset) in Schema Registry
  - Creation of mappings to Dublin Core in Schema Registry
  - Harvesting and Import in Generic Search
- Access to DARIAH-DE MD:
  - Dublin Core closest to global data model in DARIAH-DE & mandatory in OAI-PMH
  - Integration in CLARIN’s harvesting and import workflow based on dedicated CMDI MD profile
  - First unification of used vocabulary for relevant search facets

Want to know more?

- CLARIN Annual Conference 2017 (CLARIN 2017), Budapest, Hungary.

Outcomes and Further Work

- Future plans:
  - Extended support of CMDI profiles
  - Automatic import of metadata records
  - Stronger connection between DARIAH’s data modeling environment (DME) and CLARIN’s component registry
  - Mutual usage of tools for metadata curation/quality assessment

![Outcome and Further Work Diagram]