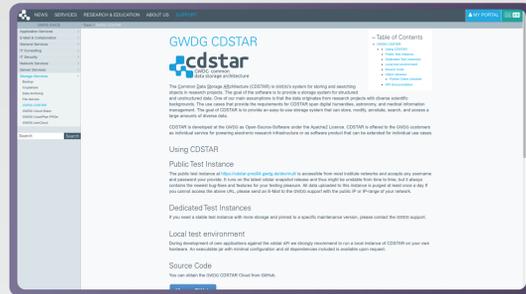


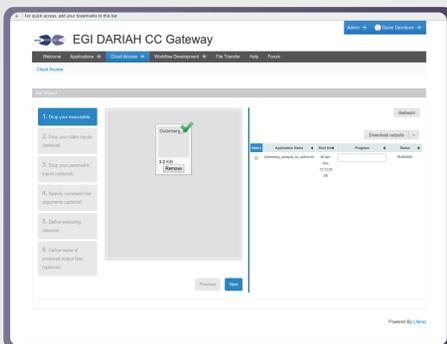
DARIAH CC Services for Digital Humanities

SIR (Multi-Source Distributed Real-Time Search and Information Retrieval)



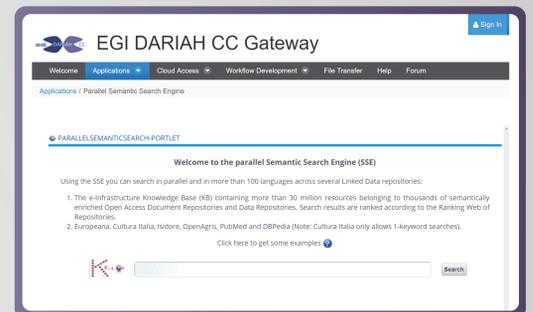
- Offer big data search and analytics capabilities for large data sets
- A combination of configurable pipelines and an interface for manual interaction allows to implement workflows that benefit from state-of-the-art technologies like elastic search or Hadoop
- The first use case for SIR -> an Optical Character Recognition (OCR) post-processing pipeline for data masking. It can help users to automatically find and mask terms, e.g. in scenarios where documents have to be anonymised. The service is currently only available for German texts

Cloud access



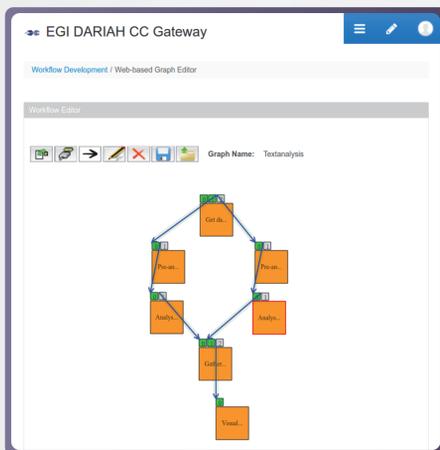
- Provides an easy to use 'Job Wizard' environment for users
- Through the Job Wizard one can define with a few clicks computational jobs that will execute on EGI cloud resources
- The environment takes care of all aspects of job management in the cloud, including instantiation of Virtual Machine images and retrieval of outputs
- Jobs can be single applications or 'parameter study' type applications

PSSE/SSE (Parallel Semantic Search Engine / Semantic Search Engine)



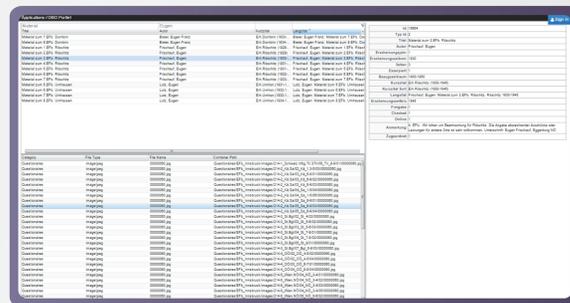
- Enables simultaneous search across several databases: Europeana, CulturalItalia, Isidore, OpenAgris, PubMed, and DBpedia
- Search results are ranked according to the RankingWeb of Repositories and connected to Google Scholar and Altmetric
- LodLive allows to navigate/explore the Linked Data graph of each resource found by a search
- Benefits for Cultural Heritage community:
 - Discover connections between several types of resources, e.g. documents, data, images, etc.
 - Discover connections and correlations between Cultural Heritage topics and topics belonging to other disciplines/domains
 - Connect the SSE with specific vocabularies and thesauri to make the results of searches more relevant for a given Cultural Heritage sub-domain

Workflow development



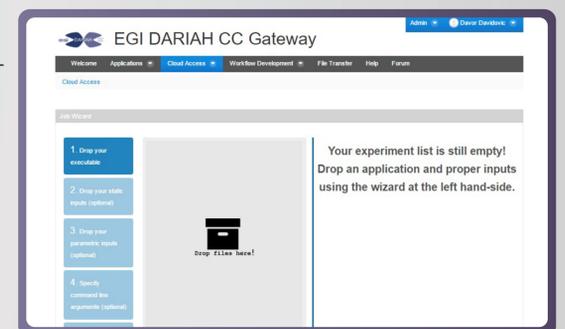
- Service that enables development and execution of complex workflow-based applications
- The workflows are executed on the EGI FedCloud sites, members of the DARIAH Virtual Organization

DBO@Cloud



- Cloud-based repository presenting the work of a 100+ years old collection of Bavarian dialects within the Austrian-Hungarian monarchy from the beginning of German language to nowadays
- The datasets are provided by the Austrian Academy of Science with 50,000+ records
- Hierarchical, "à la iTunes" searching functionality

File transfer



- Enables easy data transfer from, to and between different storage services
- Supporting protocols: HTTP, HTTPS, SFTP, GSIFTP, SRM, iRODS and S3

DARIAH Science Gateway

Welcome to the EGI DARIAH Gateway!

This gateway was developed within the EGI-Engage H2020 project and operated by the EGI DARIAH Competence Centre. This gateway displays state-of-the-art capabilities of what is available in digital science with the use of EGI. The gateway provides various web-based applications and services for the Digital Humanities researchers, institutes and communities.

The gateway includes a Help page for giving more information on the applications and services it provides.

The gateway includes the following applications:

- Simple Semantic Search Engine (SSE): Allows users to search in the e-Infrastructure Knowledge Base (Open Access Document Repositories and Data Repositories). No authentication is required to access the service.
- Parallel Semantic Search Engine (PSSE): PSSE is a parallelised version of SSE enabling simultaneous search across the e-Infrastructure Knowledge Bases: Europeana, Cultural Italia, Isidore, OpenAgris, PubMed and DBpedia platforms. No authentication is required to access the service.
- DBO@Cloud: A Cloud-based repository presenting 100+ years old collection of Bavarian dialects. No authentication is required to access the service.

The gateway includes the following services:

- Cloud Access: Single job applications and parameter-sweep applications can be run on the DARIAH VO clouds without porting efforts. Authentication is required to access the gateway service and then robot certificate provides access to the cloud.
- Workflow Development: Complex workflow applications can be developed and run on all the resources of the DARIAH VO. Authentication is required to access the gateway service and then robot certificate provides access to the cloud.
- File transfer: Enables transferring data from, to and between storage services providing HTTP, HTTPS, SFTP, GSIFTP, SRM, iRODS and S3 protocols. Authentication is required to access the gateway service and then additional authentication is needed to access the selected storage service.

DARIAH Science Gateway is an EGI service provided by DARIAH Competence Centre, co-funded by EGI-Engage <http://egi.eu/en/>

<https://dariah-gateway.lpds.sztaki.hu>

Contact info:

Karolj Skala, Davor Davidovic (skala@irb.hr, ddavid@irb.hr)
 Mailing list: cc-dariah@mailman.egi.eu
 Wiki pages: https://wiki.egi.eu/wiki/Competence_centre_DARIAH

Supported by H2020 EGI ENGAGE (DARIAH CC):