

EGI-Engage – Competence Centre for DARIAH-ERIC

Davor Davidović¹, Eveline Wandl-Vogt², Karolj Skala¹, Tibor Kalman³

¹Ruder Bošković Institute, Bijenička cesta 54, Zagreb, Croatia
{davor.davidovic, skala}@irb.hr

²Austrian Academy of Sciences, Sonnenfelsgasse 19, Wien, Austria
Eveline.Wandl-Vogt@oeaw.ac.at

³Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen, Germany
tabor.kalman@gwdg.de

Keywords: DARIAH, EGI, Cloud computing, digital repository, science gateways

1 EGI DARIAH Competence Centre

In the last decade it becomes almost impossible to image excellent science without support of e-Infrastructure. However, although various research areas, such as medicine, chemistry, and physics, mostly depend on the availability of advanced research infrastructure, the area of Arts and Humanities is still not utilizing the available infrastructure at their full potential. As an example, only a few applications coming from the area of Arts and Humanities are exploiting the EGI [1] distributed resources.

To overcome this gap, the EGI DARIAH Competence Centre (CC) project [2] aims to provide a wider and more efficient access to, and use of, research e-Infrastructures at European Grid Infrastructure (EGI) level, including: transnational access, joint research, networking and digital services for user coming from the areas of the Arts and Humanities, represented by the DARIAH-EU [4]. The first step in achieving this goal is by providing end-user support, at both technical and service level. One way of fulfilling this mission is by providing a workflow-based Science Gateway [5]. It is proven in many previous projects that the existence of such gateways significantly improves the usability of the distributed infrastructure (for storage and computational purposes), makes its usage easier, and attracts more users. For a specific research domain, a Science Gateway (based on the generic-purpose gUSE/WS-PGRADE [6,7] and gLibrary [8] technologies) will be adapted and tailored to meet the needs of the user community coming from the field of Arts and Humanities or, in a more general case, from the Social Science and Humanities. The new gateway will provide access and compute services for data residing in distributed grid and cloud storages, generic applications, as well as computational resources. The usefulness and evaluation of the specific Science Gateway will be demonstrated with two specific pilot applications: “Storing and Accessing DARIAH contents on EGI” (SADE), and “Multi-Source Distributed Real-Time Search and Information Retrieval” (SIR). Furthermore, the goal of EGI DARIAH CC is to promote the usefulness of e-Infrastructure and to provide expertise and knowledge to DARIAH users they re-

quired for their work. Therefore, numerous dissemination events including end-user trainings, presentations, and seminars will be organized. Furthermore, the project envisages a proactive roadmap for porting applications to EGI Federated Cloud infrastructure [9,10] and intensifies recognition of the user communities and their needs during the project lifetime. All the outcomes of the EGI DARIAH CC project will be contributed to DARIAH-EU [4] through the DARIAH Virtual Competency Centre 1 “e-Infrastructures”, which will take appropriate means to support and maintain the results.

EGI DARIAH CC is a part of the EGI-Engage Horizon2020 project [3]. The mission of EGI-Engage is to accelerate the implementation of the Open Science Commons vision [11], where researchers from all disciplines have easy and open access to the innovative digital services, data, knowledge and expertise they need for their work. The Open Science Commons is grounded on three pillars: the e-Infrastructure Commons, an ecosystem of key services; the Open Data Commons, where any researcher can access, use and reuse data; and the Knowledge Commons, in which communities have shared ownership of knowledge and participate in the co-development of software and are technically supported to exploit state-of-the-art digital services. EGI-Engage project expands the capabilities offered to scientists (e.g. improved cloud or data services) and the spectrum of its user base by engaging with large Research Infrastructures (RIs), the long-tail of science and industry/SMEs. The main engagement instrument will be a network of eight Competence Centers, one of which is the DARIAH Competence Centre.

2 DARIAH CC and SEMANTIC WEB TECHNOLOGY PROJECTS

At the ESWC2015 we focus on the presentation of the DARIAH CC outline. We are offering insight into DARIAH ERIC (Digital Research Infrastructure for the Arts and Humanities), based on its newly founded working groups along the Virtual Competency Centres and across them. As EGI Engage is just about to start, we are offering good opportunities to connect with the project. On the example of the use case of electronic lexicography (EGI DARIAH CC) and lexical resources (DARIAH ERIC working group) we are demonstrating an example for the liaison with semantic web technologies projects and research infrastructures, which we are looking for and would highly appreciate. Examples of typical digital humanities data are introduced and taken into account, such as paper slips of unique 100-years+ old questionnaires or drawings and manuscripts.

We warmly welcome semantic web partners to further modeling, access and analysis on EGI DARIAH CC content. The Austrian Academy of Sciences, Austrian Centre for Digital Humanities offers internships and jobs in this research area. We invite to contribute and collaborate within the framework of the DARIAH ERIC working groups, e.g. on standardization, lexical resources, and NLP (Natural Language Processing). In general, by being part of the networking session, we envision on the one hand challenging new research topics, on the other exploration of new methodologies

in cross discipline, entry, and cross sectorial teams and look for partners to further this in upcoming Horizon2020 calls.

References

1. European Grid Infrastructure (EGI), <http://www.egi.eu/>
2. DARIAH Competence Centre, https://wiki.egi.eu/wiki/Competence_centre_DARIAH
3. EGI-Engage Horizon2020 project, <https://www.egi.eu/about/egi-engage/>
4. DARIAH-EU – Digital Research Infrastructure for the Arts and Humanities, <https://www.dariah.eu/>
5. Kacsuk, P., et al.: WS-PGRADE/gUSE generic DCI gateway framework for a large variety of user communities. *Journal of Grid Computing*, 10, 4, 601-630 (2012)
6. Akos, B., Farkas, Z., Kacsuk, P.: Building science gateways by utilizing the generic WS-PGRADE/gUSE workflow system. *Computer Science Journal*, 14, 2 307-325 (2013)
7. gUSE – Grid and Cloud User Support Environment, <http://guse.hu/about/architecture>
8. gLibrary – Digital Libraries on the Grid, https://glibrary.ct.infn.it/glibrary_new/index.php
9. European Grid Infrastructure, Federated Cloud, <https://www.egi.eu/infrastructure/cloud/>
10. Petcu, D.: Portability and interoperability between clouds: challenges and case study. *Towards a Service-Based Internet*. Springer Berlin Heidelberg, 62-74 (2011)
11. Open Science Commons, <http://sciencecommons.org/resources/readingroom/principles-for-open-science/>