

Briding the GAP: 100 Jahre Dialektlexikographie als Cloud Service. Der SADE Use Case im DARIAH Competence Centre

,
eveline.wandl-vogt@oeaw.ac.at
Österreichische Akademie der Wissenschaften, Austrian
Centre for Digital Humanities; AT

,
roberto.barbera@ct.infn.it
Istituto Nazionale de Fisica Nucleare; IT

,
guiseppe.larocca@ct.infn.it
Istituto Nazionale de Fisica Nucleare; IT

,
antonio.calanducci@ct.infn.it
Istituto Nazionale de Fisica Nucleare; IT

,
tkalman@gwdg.de
Gesellschaft für wissenschaftliche Datenverarbeitung
mgH Göttingen; DE

Intro

In den Naturwissenschaften ist es seit mehreren Jahrzehnten undenkbar, exzellente Forschung nicht auf Basis elektronischer Forschungsinfrastrukturen zu betreiben bzw. sich diese zu Nutze zu machen. Anders ist dies in den Geisteswissenschaften, wo es nach wie vor zu den Ausnahmen gehört, sich elektronische Forschungsinfrastrukturen für die eigene Forschungsarbeit einzusetzen.

Um diese Lücke zu schließen, wurde das Projekt EGI ENGAGE – DARIAH Competence Centre gestartet (Näheres s.u.).

In diesem Paper diskutieren die Autor:innen beispielhaft den Einsatz und die Zusammenarbeit zweier sehr unterschiedlicher Anbieter von europäischen Forschungsinfrastrukturen, nämlich Digital Research Infrastructures for the Arts and Humanities European Research Infrastructure Consortium (DARIAH-ERIC (cf. DARIAH) und European Grid Infrastructure (EGI 2010-*).

Die Ausgangsprojekte: EGI ENGAGE, DARIAH CC, exploreAT! und COST ENEL

Das Projekt EGI ENGAGE – Engaging the Research Community towards an Open Science Commons (EGI 2015-*) – startete am 1. März 2015 unter Mitfinanzierung der Europäischen Kommission (Horizon2020) mit einer Projektlaufzeit von 30 Monaten und in Zusammenarbeit von mehr als 70 Institutionen in über 30 Ländern. Die Projektmission ist, Open Science Commons für die Forschungsgemeinschaften zu erschließen, und Zugang zu bzw. Nutzung von elektronischen Infrastrukturen für die Forschung zu schaffen, und somit innovative Forschung zu fördern.

Zu diesem Zweck wird mit Forschungsinfrastrukturanbietern unterschiedlicher Disziplinen zusammengearbeitet, unter anderem mit DARIAH-ERIC.

Im Workpackage DARIAH Competence Centre (cf. EGI-DARIAH Competence Center) werden im besonderen Forscher_innen aus den Arts und Humanities angesprochen und eingeladen, elektronische Infrastrukturen zu nutzen.

Übergeordnete Ziele des Competence Centre sind

(1) die Stärkung der Zusammenarbeit zwischen DARIAH-EU und EGI, basierend auf Workflow-orientierten Gateway-Applikationen und der extensiven Nutzung von Applikationen aus den Arts und Humanities in der EGI Cloud (EGI FedCloud).

(2) die Vermehrung der zugänglichen e-Science Services und Applikationen für die Forscherinnen der Arts und Humanities sowie die Integration existierender NGI Ressourcen in EGI

(3) das Bewusstsein von möglichen Vorteilen (Stichwort „exzellente Forschung“) durch elektronische Infrastrukturen und e-Science Technologien für Forscher:innen der Arts und Humanities zu wecken bzw. zu stärken, indem man Konditionen schafft für nachhaltige Steigerung einer wachsenden Community von den Arts und Humanities und in den Social Sciences

(4) die Arbeit, die mit anderen Initiativen gestartet wurde wie zB. DC-NET (2009-2012), DCH-RP (2012-2014) auszuweiten und Ergebnisse für die Arts und Humanities konkret nutzbar zu machen und zu implementieren.

Für das DARIAH-CC werden Miniprojekte pilotmäßig umgesetzt. Im konkreten Paper stellen die Autor:innen den SADE Use Case (Näheres s.u.) vor, in welchem die „Datenbank der bairischen Mundarten in Österreich (DBÖ)“ basierend auf unterschiedlichen Formaten und Arbeitsschwerpunkten – Archivierung, Digitalisierung (cf. Wandl-Vogt 2007-*), technische Professionalisierung und Georeferenzierung (cf. explore.AT (2015-*), semantisch-kulturelle Vernetzung (Näheres s.u.; cf. COST-ENEL 2013-*)- weiterentwickelt wird.

Im Rahmen des laufenden Nationalstiftungsprojekts exploreAT! (1. April 2015; 48-60 Monate) werden

methodisch neue Wege der Lexikographie erprobt. Schwerpunkte der aktuellen Arbeit liegen in der semantisch-kulturellen Vernetzung (Linked Open Data), visuellen Analyse, der Entwicklung von Serious Games sowie einer Erarbeitung und Umsetzung von Bürgerbeteiligungsmodellen (Citizen Science) und der Implementierung von Open Science. Methodisch werden damit Schritte in Richtung eines neuen digitalen Wörterbuchs (WBÖ 3.0) gesetzt.

Die Ergebnisse werden mittels bestehender Netzwerke wie vor allem der DARIAH Arbeitsgruppe Lexical Resources und des COST Netzwerks IS 1305 zur elektronischen Lexikographie (ENeL 2013-2017; Davidovic 2015 et al.) im europäischen Kontext eingebettet, reflektiert und diskutiert. Für den SADE Use Case stellen diese Communities erste – aber nicht ausschließliche – Anlaufstellen für Zusammenarbeit dar.

Der SADE Use Case

SADE – kurz für: Storing and Accessing DARIAH content on EGI – sieht die Entwicklung eines kompletten Workflows für Forscher:innen im Bereich lexikalische Ressourcen / Lexikographie dar.

Ausgehend der Beispielsammlung der ÖAW – zunächst der Daten aus dem Projekt (2007-) – und in Kontext mit Kooperationspartnern dieses Projekts, beispielsweise dem Naturhistorischen Museum Wien, Wikimedia.AT, Open Knowledge Foundation Österreich, Europeana, sowie der Partner aus DARIAH und COST ENeL, wird ein lexikalisches Netzwerk für kulturelle Fragestellungen erschlossen.

Ein erster Schritt für die Nutzung von Science Commons ist die Entwicklung von Science Gateways. Diese tragen nachweislich zur vermehrten Nutzung von Infrastrukturen in den jeweiligen Forschungsbereichen / Disziplinen bei (cf. Balasko et al. 2013). Ob ein Science Gateway für DARIAH CC oder für einen Forschungsbereich eingerichtet wird, ist derzeit nicht endgültig entschieden. Das Gateway beruht auf Adaptierungen und fallspezifischen Erweiterungen von gUSE/WS-PGRADE (gUSE 2009-*) sowie gLibrary (cf. gLibrary) Technologien. Es wird gemäß den Erwartungen und Bedürfnissen der Lexikograph:innen angepasst und weiterentwickelt.

Um eine Verbindung zwischen den sehr unterschiedlichen Forscher:innengruppen herzustellen, finden unter anderem Vernetzungstreffen statt, z. B. Februar 2015: COST-ENeL-meeting in Wien; Dezember 2015: DARIAH-COST-meeting zu Common Names in Wien.

Die Ergebnisse des SADE Use Case werden umgekehrt im COST-ENeL-Netzwerk zur Diskussion gestellt, um User Interessen auf breiter Basis einzuholen und in die Weiterentwicklung der Infrastruktur einzubringen.

Für die Nutzung der Cloud von essentieller Bedeutung ist das Vorhandensein einer inzwischen eingerichteten

Virtuellen Organisation für die Arts und Humanities – vo.dariah.eu (vgl. EGI-VO (joined EGI 01072011)).

Forschungsparadimenwechsel im Blickwinkel der Zusammenarbeit

Zusammenfassend wird das vorgestellte Projekt unter dem breiten Blickwinkel der Workplace-Innovation (Kesselring 2014 et al.) diskutiert. Dabei wird der Schwerpunkt auf folgende Punkte gelegt:

(1) Produktinnovation

Durch die Zusammenarbeit entstehen neuen „Produkte“ wie beispielsweise ein Science Gateway für die Geisteswissenschaften.

Für die Arts und Humanities wird beabsichtigt, Produkte zu entwickeln, die Folge-innovationen bedingen und den Forschungsprozess stimulieren.

(2) Prozessinnovation

Die Zusammenarbeit im interdisziplinären, internationalen, interkulturellen, multilingualen Kontext bedingt Veränderungen laufender Forschungsprozesse bzw. zielt sogar teilweise von Beginn an auf eine derartige Veränderung und Innovation ab.

Durch die Implementierung in bestehende Forschungsnetzwerke und die Anbindung an aktuelle Forschungsfragen auf europäischem Niveau soll sichergestellt werden, dass es sich nicht nur um Projektinnovationen – z. B. eine neue Datenspeicherung für die handelt – sondern dass dadurch neue Forschungs- und Entwicklungsprozesse angestoßen werden.

(3) Organisationsinnovation

Die Einbettung in das Forschungsparadigma digitaler Infrastrukturen und Cloud Services bedingt Organisationsinnovationen, wie beispielsweise die Gründung der virtuellen Organisation für DARIAH, der Working Group Cloud Services in DARIAH u. ä.

Nach der Implementierung der Services aus dem SADE Use Case soll eine Impactmessung der Workplaceinnovation entwickelt werden.

An einer Maximierung der positiven Wirkungen der technischen Neuerungen flankierend durch soziale Innovationen am Beispielfall der digitalen Lexikographie wird gearbeitet.

Bibliographie

Balasko, Akos / Farkas, Zoltan / Kacsuk, Peter (2013): „Building science gateways by utilizing the generic WS-PGRADE/gUSE workflow system.“, in: *Computer Science Journal* 14, 2: 307-325.

COST-ENeL (2013-*): *European CO-operation in Science and Technology: European Network of e-Lexicography (ENeL)*. COST Association. http://www.cost.eu/COST_Actions/isch/IS1305 [letzter Zugriff 15. Oktober 2015].

DARIAH: *Digital Research Infrastructures for the Arts and Humanities* (DARIAH). <http://www.dariah.eu> [letzter Zugriff 15. Oktober 2015].

Davidovic, David / Wandl-Vogt, Eveline / Skala, Karolj / Kalman, Tibor (2015): „EGI Engage – Competence Centre for DARIAH.“, in: *12th European Semantic Web Conference 2015 Project Networking Session*. http://2015.eswc-conferences.org/sites/default/files/PN-ESWC-2015_num5.pdf [letzter Zugriff 15. Oktober 2015].

DC-NET (2009-2012): *Digital Cultural heritage NETwork*. ERA-NET (European Research Area Network). <http://www.dc-net.org/> [letzter Zugriff 15. Oktober 2015].

DCH-RP (2012-2014): *Digital Cultural Heritage Roadmap for Preservation*. <http://www.dch-rp.eu/> [letzter Zugriff 15. Oktober 2015].

EGI (2010-*): *European Grid Infrastrucutre* (EGI). <http://www.egi.eu> [letzter Zugriff 15. Oktober 2015].

EGI: Dariah Competence Centre: https://wiki.egi.eu/wiki/Competence_centre_DARIAH [letzter Zugriff 15. Oktober 2015].

EGI-Engage (2015-*): *Engaging the Research Community towards an Open Science Commons* (EGI-ENGAGE). <https://www.egi.eu/about/egi-engage/> [letzter Zugriff 15. Oktober 2015].

EGI-VO (joined EGI 01072011): *EGI: Virtual Organisation for arts and humanities*: vo.dariah.eu. EGI-Engage Competence Centre https://wiki.egi.eu/wiki/EGI_Virtual_Organisation_for_arts_and_humanities:_vo.dariah.eu [letzter Zugriff 15. Oktober 2015].

ENeL (2013-2017): *European Network of e-Lexicography* (ENeL). COST <http://www.elexicography.eu> [letzter Zugriff 15. Oktober 2015].

exploreAT! (2015-*): *exploring austria's culture through the language glass*. <http://www.oeaw.ac.at/acdh/de/node/187> [letzter Zugriff 15. Oktober 2015].

gLibrary (o.J.): *Digital Libraries on the Grid* (gLibrary). INFN http://glibrary.ct.infn.it/glibrary_new/index.php [letzter Zugriff 15. Oktober 2015].

gUSE (2009-*): *Grid and Cloud User Support Environment* (gUSE). Laboratory of Parallel and Distributed Systems (LPDS) <http://guse.hu/about/architecture> [letzter Zugriff 15. Oktober 2015].

Kesselring, Alexander / Blasy, Cosima / Scopetta, Anette (2014): *Workplace Innovation. Concepts and indicators*. European Commission Report <http://ec.europa.eu/DocsRoom/documents/8250/attachments/1/translations/en/renditions/native> [letzter Zugriff 15. Oktober 2015].

Wandl-Vogt, Eveline (ed.) (2007-*): *"Datenbank der bairischen Mundarten in Österreich"* (DBÖ), Wien 1993-2010, electronically mapped. Institut für Österreichische Dialekt- und Namenlexika, Zentrum für Sprachwissenschaften, Bild- und Tondokumentation, Österreichische Akademie der Wissenschaften (publiziert: Wien: 1. Juli 2010). <http://wboe.oeaw.ac.at> [letzter Zugriff 15. Oktober 2015].