

Twin Talks 3: Understanding and Facilitating Collaboration in DH

July 23, 2020

Organisers:

Steven Krauwer (steven@clarin.eu)

Darja Fišer (darja.fiser@ff.uni-lj.si)

and Programme Committee

on behalf of:

CLARIN ERIC (www.clarin.eu)

DARIAH ERIC (www.dariah.eu)

SSHOC (sshopencloud.eu)



Opening

Predecessors:

- [Twin Talks 1](#), Copenhagen, DHN2019
- [Twin Talks Panel](#), Liège, DH Benelux 2019
- [Twin Talks 2](#), Riga, DHN2020

Main objective:

- A better understanding of the collaboration between humanities and digital experts when solving humanities research problems and the obstacles encountered

Expected outcome:

- Recommendations for those involved in the education of humanities scholars & professionals and technical experts to develop better training programmes
- Identification of initiatives that CLARIN ERIC, DARIAH ERIC and SSHOC could take in order to support this



What is it?

- Common Language Resources and Technology Infrastructure
- **ESFRI** ERIC status since 2012, Landmark since 2016

Aim

Provide easy and sustainable access:

- for scholars in the **humanities and social sciences** and beyond
- to **digital language data** (in written, spoken, video or multimodal form)
- to **advanced tools** to discover, explore, exploit, annotate, analyse or combine data, wherever they are located

How?

Through a **single sign-on** environment that serves as an ecosystem for **knowledge sharing**

CLARIN & Open Science

- promotion for the **sharing and re-use of data through sustainable data registries**
- all integrated data is available in **open access** for research purposes
- support for **linguistic diversity**
 - data covering more than 1500 languages
 - tools for many languages
 - language resources in all modalities
- adherence to **FAIR data principles**
 - Findable, Accessible, Interoperable, Re-usable
 - promotion of responsible data science

Data Types

- newspaper archives
- literary texts
- social media data
- parliamentary records
- historical letters
- oral history data
- disciplinary libraries
- institutional archival data
- broadcast archives
- see also the info on the CLARIN Resource Families initiative:
<https://www.clarin.eu/resource-families>

Services

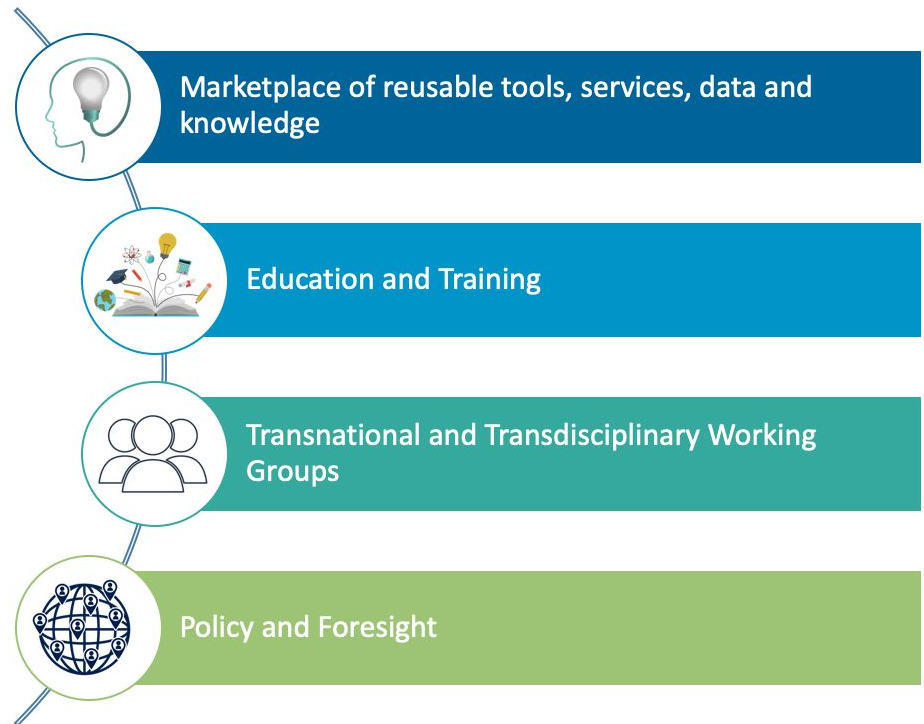


DARIAH is a pan-European Digital Research Infrastructure that enhances and supports digitally-enabled research and teaching across the Arts and Humanities

Launched and added to the ESFRI Roadmap 2006/ Formally constituted as a European Research Infrastructure Consortium in 2014



DARIAH's **mission** is to empower research communities with digital methods to create, connect and share knowledge about culture and society.



Project:



Horizon 2020
European Union Funding
for Research & Innovation

Type of action & funding:
Research and Innovation action
(INFRAEOSC-04-2018)

Partners: 45
(20 beneficiaries + 25 LTPs)

SSH ESFRI Landmarks and Projects
& international SSH data infrastructures

Project budget:
€ 14,455,594.08

Duration: 40 months
(January 2019 – 30 April 2022)

Project website:
www.SSHopencloud.eu



Objectives:

- creating the social sciences and humanities (SSH) part of European Open Science Cloud (EOSC)
- maximising **re-use** through **Open Science** and **FAIR** principles (standards, common catalogue, access control, semantic techniques, training)
- interconnecting existing and new infrastructures (clustered cloud infrastructure)
- establishing appropriate **governance model** for SSH-EOSC

Expected impact



The Social Sciences and Humanities are seamlessly integrated in the European Open Science Cloud



Availability of an EU-wide, easy-to-use SSH Open Marketplace, where tools and data are openly accessible



EU-wide availability of high quality "cloud ready" SSH tools and high quality SSH data



EU-wide availability of trusted and secure access mechanisms for SSH data, conforming to EU legal requirements



State of the art Research Infrastructure in several pilot domains advanced through dedicated SSH data pilots cluster projects



Maximising reuse through Open Science and FAIR principles (standards, common catalogue, access control, semantic techniques, training)

Join our community



<https://www.sshopencloud.eu>



@SSHOpenCloud



info@sshopencloud.eu



/in/sshopencloud

Workshop schedule

15:00 - 15:10	Opening	<u>Steven Krauwer</u> and Darja Fišer
15:10 - 15:20	Talk 1: <i>Deep learning for paleographic analysis of medieval Hebrew manuscripts</i>	Jihad El-Sana, <u>Daria Vasyutinsky Shapira</u> , <u>Irina Rabaev</u> , Berat Kurar Barakat and Ahmad Droby
15:20 - 15:30	Talk 2: <i>Bridging the gap between culture studies and computational linguistics</i>	<u>Inge van de Ven</u> and <u>Menno van Zaanen</u>
15:30 - 15:40	Talk 3: <i>The Whole is Greater Than the Sum of its Parts Analyzing Aristotle Commentaries in Collaboration between Philology and Data Science</i>	<u>Michael Krewet</u> , <u>Danah Tonne</u> , Germaine Götzelmann, Philipp Hegel and Sibylle Söring
15:40 - 15:50	Q&A	All
15:50 - 16:00	Talk 4: <i>Towards improving OCR accuracy with Bulgarian Language Resources</i>	Ivan Kratchanov, Laska Laskova and <u>Kiril Simov</u>
16:00 - 16:10	Talk 5: <i>Bringing Together the Digital with Research, Teaching, and Conservation</i>	Robert M. Ehrenreich, <u>Michael Haley Goldman</u> and Jane E. Klinger
16:10 - 16:20	Talk 6: <i>No room for hate: What research about hate speech taught us about collaboration?</i>	<u>Aida Šulc</u> and <u>Kristina Pahor de Maiti</u>
16:20 - 16:30	Q&A	All
16:30 - 16:50	Discussion	All
16:50 - 17:00	Take-home messages and closing	<u>Steven Krauwer</u> and Darja Fišer