

Tour de CLARIN

<https://www.clarin.eu/tour-de-clarin>

Darja Fišer^{1,2}, Jakob Lenardič¹, Karolina Badzmierowska³

¹ University of Ljubljana, ² Jožef Stefan Institute, ³ CLARIN ERIC
darja.fiser@ff.uni-lj.si, jakob.lenardic@ff.uni-lj.si, karolina@clarin.eu



General information

- A bi-monthly initiative that highlights User Involvement activities in CLARIN national consortia
- Helps to increase the visibility of the CLARIN national consortia
- Reveals the richness of the CLARIN landscape
- Displays the full range of activities throughout the CLARIN network

Finland – April & May 2017

- Tool: **AaltoASR**
 - Automatic speech recognition, can handle spontaneous speech and agglutinative languages
- Resource: **Suomi24**
 - Corpus of Finland's largest discussion forums, 2.6 million tokens, used in the multidisciplinary Citizen Mindscapes project to develop new ways of looking at social and political interaction
- Interview: **Dr Tommi Jantunen**
 - A linguist specialising in the Finnish Sign Language, developing and annotating the Finnish Sign Language corpus
- Event: **Kielipankki roadshows**
 - Series of talks to celebrate Kielipankki's 20th anniversary

Sweden – June & July 2017

- Tool: **Korp**
 - Versatile corpus concordancer used by several consortia
- Resource: **Riksdag's open data**
 - Swedish parliamentary corpus, 1.25 billion tokens, complex annotation, used in DH research (e.g. analysis of film policies by P. Snickars and F. Norén)
- Interview: **Professor Maria Ågren**
 - A historian studying professional roles of men and women in early modern Swedish history, developed the GaW database
- Event: **Tutorial and workshop on automatic sentence selection from corpora**
 - Knowledge sharing between SWE and EST, aimed at lexicographers and L2 researchers

Austria – August & September 2017

- The tool: **Viennese Lexicographic Editor**
 - XML editing system for dictionaries, successfully used in international projects (e.g. Viennese Corpus of Arabic Varieties)
- The resource: **ABaC:us Corpus**
 - Corpus of religious texts from the Baroque period, rich and accurate manual annotations
- Interview: **Dr Stephan Prochazka**
 - Historian working on the Viennese Corpus of Arabic varieties and within the TUNICO project
- Event: **ACDH Tool Library**
 - Developers and experienced professionals share their theoretical and practical knowledge on tools that are designed for Digital Humanities users

Future stops

- **The Dutch Language Union – October and November 2017**
- **Poland – December 2017 and January 2018**
- **Other slots still open, welcome to apply!**



CLARIN ERIC was established in 2012; it is a landmark in the 2016 ESRI roadmap.

www.clarin.eu

 clarin@clarin.eu

 [@CLARINERIC](https://twitter.com/CLARINERIC)

 facebook.com/CLARINERIC

 github.com/clarin-eric

Acknowledgements

We would like to thank all the representatives who have helped shape Tour de CLARIN: Mieta Lennes, Lars Borin and Tanja Wissik