ParlaCLARIN: Creating and Using Parliamentary Corpora

Workshop @ LREC2018, Miyazaki, 7 May 2018

www.clarin.eu
CLARIN in five bullets

• **CLARIN** is the Common Language Resources and Technology Infrastructure

• that provides 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)

• and **advanced tools** to discover, explore, exploit, annotate, analyse or combine them, wherever they are located

• through a **single sign-on** environment
CLARIN in resource types

- Parliamentary records
- Literary texts
- Social Media data
- Historical letters
- Oral History data
- Disciplinary libraries
- Institutional archival data
- Broadcast archives
- Newspaper archives
- …
worksheets / tutorials
- ParlaCLARIN
- MLP-MOMENT2018
- Tools for handling Speech Databases

talks/posters
- Meet CLARIN’s Key Resource Families (Wednesday, Area 1, 16:45-17:55)
  Darja Fišer, Jakob Lenardič and Tomaž Erjavec
- CLARIN: Towards FAIR and Responsible Data Science Using Language Resources (Friday, O29, 9:45-10:05)
  Franciska de Jong, Bente Maegaard, Koenraad De Smedt, Darja Fišer and Dieter Van Uytvanck

booth during conference
- overview of talks and posters linked to CLARIN
- documentation
- meeting point
Logistics

• joint dinner
  – @ Pine terrace
  – welcome at 19:30
  – drinks on CLARIN

• slides will be collected and shared via the event page for the workshop on the CLARIN website: https://www.clarin.eu/ParlaCLARIN

• …
This workshop was prepared with financial support from the following H2020 projects: CLARIN-PLUS PARTHENOS

* see you @ www.clarin.eu or f.m.g.dejong@uu.nl
Darja Fišer & Jakob Lenardič
University of Ljubljana and Jožef Stefan Institute, Slovenia

CLARIN Corpora for Parliamentary Discourse Research