



# Transcending Scientific Boundaries across SSH Research focusing on Parliamentary Data

## WHAT MAKES PARLIAMENTARY DATA SUITED FOR RESEARCH?

### Parliamentary data is:

- ➔ a major source of socially relevant content
- ➔ a perfect ground for interdisciplinary research as it is interesting for different Social Sciences and Humanities fields, such as political sciences, sociology, history, psychology, linguistics, cultural studies
- ➔ available in ever larger quantities (Freedom of Information Acts that are supported by the United Nations and set in place in over 100 countries worldwide, making the data easily obtainable)
- ➔ well suited for trans-national studies (how similar social issues are tackled by different parliaments) as well as diachronic analyses (how similar social issues are tackled across time)
- ➔ multilingual and multimodal (text, audio, video)
- ➔ enhanced with rich metadata

In the current period of

- increasing social paradigm shifts
- political polarisations
- popular and populist movements

it has become increasingly important to examine

- the political agendas
- the argumentation strategies
- the deeper motivations
- the ultimate goals

of actors on the political stage in general, and in parliamentary institutions in particular.



## PARLIAMENTARY DATA IN CLARIN

**18** parliamentary corpora surveyed

**1** MULTILINGUAL: *Europarl* (21 languages)

**17** MONOLINGUAL: **14** languages:

1 Czech 1 French 2 German 2 Norwegian 1 Slovenian  
1 Danish 1 Estonian 1 Greek 1 Polish 1 Swedish  
2 English 1 Finnish 1 Lithuanian 1 Portuguese

### AVAILABILITY

**3** through a concordancer  
**9** for download  
**5** both

### SIZE

**7** small (<10 million tokens)  
**7** medium (10–100 million tokens)  
**3** large (>100 million tokens)

### LICENCE

**8** CC-BY  
**1** CLARIN ACA  
**1** CLARIN PUB

### ANNOTATION

**8** lemmatised  
**8** PoS-tagged

### CLARIN in a nutshell:

- 2012: ESFRI ERIC status
- 2016: Landmark on the ESFRI Roadmap
- easy and sustainable access for scholars in SSH and beyond to digital language data (in written, spoken, video or multimodal form)
- advanced tools to discover, explore, exploit, annotate, analyse or combine them, wherever they are located
- single sign-on environment
- ecosystem for knowledge sharing

*Data on parliamentary debates can offer opportunities for diachronic studies in SSH and serve as historical source, especially when contextual information is captured in the metadata.*

## PARLIAMENTARY DATA

PROCEEDINGS  
MOTIONS  
DEBATES  
VOTES  
BILLS

### SSH Impact Pillar 1: Valuation

**Themes:** Impact generating processes and methods in SSH research and Added value creation by involving SSH in inter-disciplinary projects driven primarily by other disciplines

**CLARIN contribution:** Collaborative work on parliamentary data carried out by teams of SSH scholars supports smart use of an information type that is of high value for understanding citizenship and democracy as the pillars of Europe's culture.

### SSH Impact Pillar 3: Assessing Impact

**Theme:** Intended and non-intended societal impacts, structuring impacts on policy-making and policies

**CLARIN contribution:** Multidisciplinary SSH research on parliamentary data sets supports the identification of trends using text and data mining algorithms, and their explanation in the context of political sciences.

Franciska de Jong<sup>1</sup> Darja Fišer<sup>1,2</sup>  
Maria Eskevich<sup>1</sup> Karolina Badzmierowska<sup>1</sup>

<sup>1</sup> CLARIN ERIC, Drift 10, 3512 BS Utrecht, the Netherlands  
<sup>2</sup> Faculty of Arts, University of Ljubljana, Aškerčeva 2, 1000 Ljubljana, Slovenia

f.m.g.dejong@uu.nl  
darja.fiser@ff.uni-lj.si  
maria@clarin.eu  
karolina@clarin.eu

Read/download further information about CLARIN Resource Families:



[www.clarin.eu](http://www.clarin.eu)

