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

PARLIAMENTARY DATA IN CLARIN
- parlimentary corpora surveyed
- multilingual: 21 languages
- 17 monolingual
- 24 languages:
  - Czech
  - Danish
  - Estonian
  - German
  - English
  - French
  - Greek
  - Hebrew
  - Hungarian
  - Indonesian
  - Italian
  - Polish
  - Portuguese
  - Slovak
  - Slovenian
  - Spanish
  - Swedish
  - Turkish
- availability
  - through a concordancer
  - for download
  - both
- size
  - small (<10 million tokens)
  - medium (10-100 million tokens)
  - large (>100 million tokens)
- licence
  - CC-BY
  - CLARIN-ALC
  - CLARIN-PUD
- annotation
  - lemmanised
  - 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

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
Themes: 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 Jong1, Darja Fiser1, Maria Eskевич1, Karolina Badzmirowska1
1 CLARIN ERIC, Drift 10, 2528 BS, Utrecht, the Netherlands
1 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