OSCARS
Open Science Clusters' Action for Research & Society

ParlaCAP:
Comparing agenda settings across parliaments via the ParlaMint dataset

Nikola Ljubešić
Jožef Stefan Institute
Slovenia

Funded by the European Union

Summary

Parliaments are the cornerstone of democracy in Europe.
They ensure the political representation of citizens.

Work of parliaments can be followed through
textual transcripts of parliamentary proceedings.

Upstream project **ParlaMint**
- FAIR availability of proceedings of 29 European parliaments
- unified encoding and comparable metadata
- linguistic corpora intended primarily for language research

ParlaCAP's mission is to open this dataset for political scientists and other researchers that rely on tabular data rather than on textual collections.

CLARIN ERIC flagship project
(2020-2023)

26 national and 3 regional parliaments

8 million speeches, 1.2 billion words

29 languages

Linguistically annotated, translated into English

Rich speaker and party metadata

OSCARS

Most scientific disciplines lack the ability to work with **more than a billion words** of text in **29 different languages**

“Text as data” paradigm
- each speech represented through variables of **topic discussed** and **sentiment expressed**

Recent improvements in **natural language processing** - high-accuracy automatic **multilingual text classification**

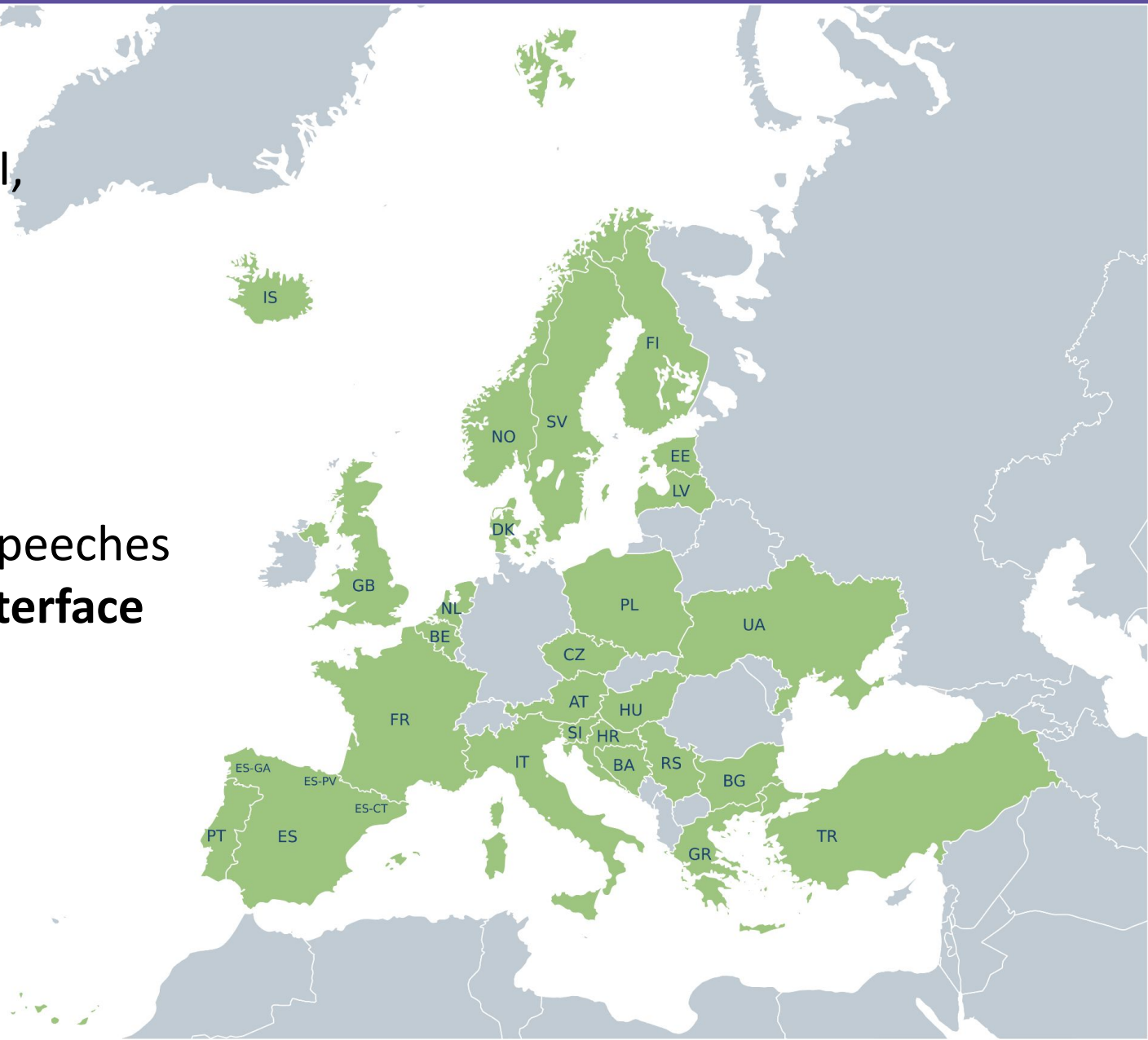
Data Distribution

Linguistic corpus through CLARIN.SI,
part of CLARIN ERIC

Tabular dataset through CROSSDA,
part of CESSDA ERIC

Access to subsets of the 8-million-speeches dataset through an **API and web interface**

Dissemination and training events
to ensure data uptake



Scientific Impact

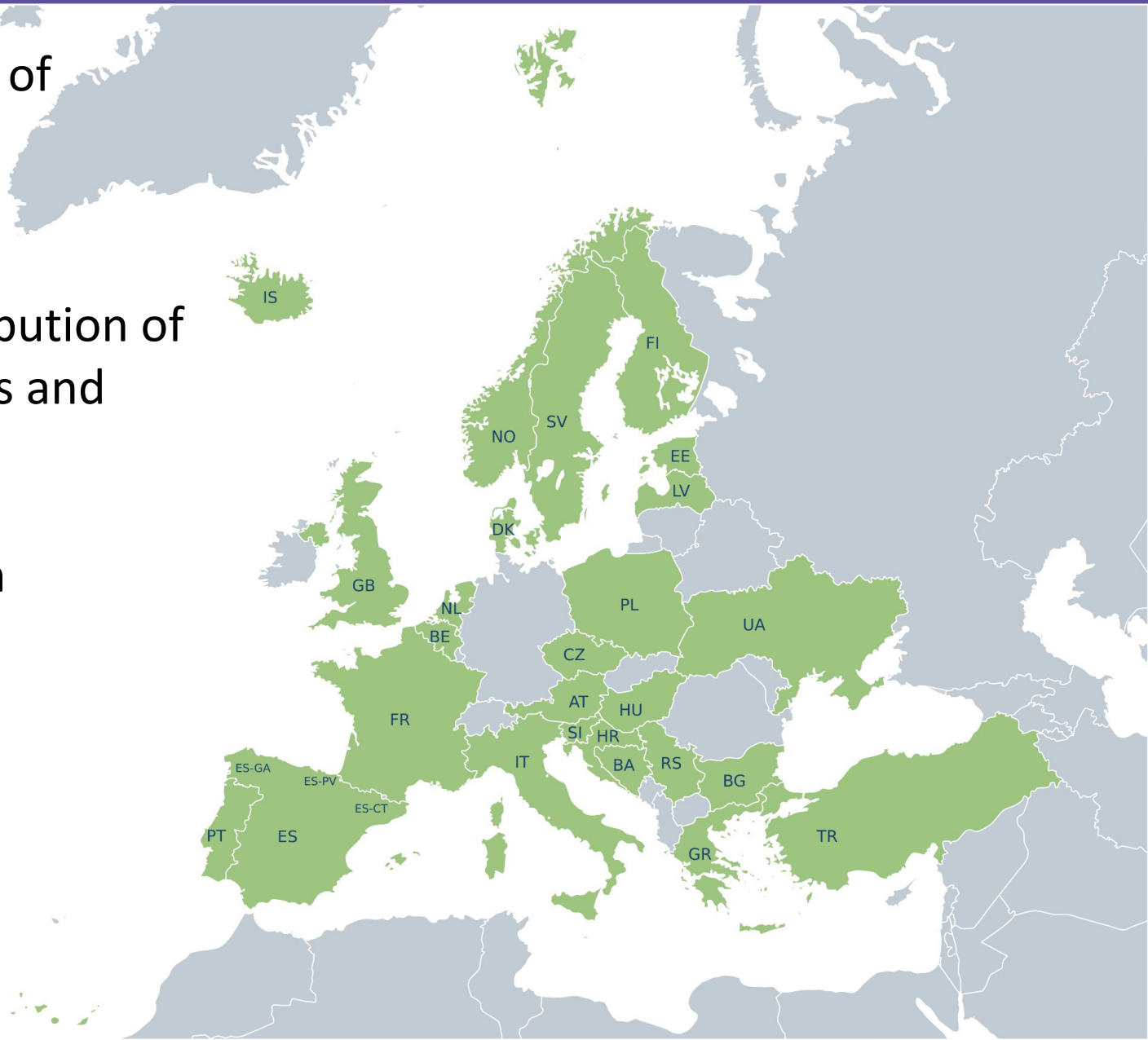
OSCARS

Comparative parliamentary studies of **political agenda-setting** across European parliaments

Unprecedented insight in the distribution of **political focus** across the 26 nations and its dependence on

- state of democracy
- position on the political spectrum
- expression of sentiment

Many additional insights to follow by us and fellow researchers



OSCARS
Open Science Clusters' Action for Research & Society

ParlaCAP:

Comparing agenda settings across parliaments via the ParlaMint dataset

nikola.ljubesic@ijs.si
Jožef Stefan Institute, Ljubljana, Slovenia

Institute of Contemporary History, Ljubljana, Slovenia
Faculty of Political Science, Zagreb, Croatia
Polish Academy of Sciences, Warsaw, Poland
Bulgarian Academy of Sciences, Sofia, Bulgaria

Funded by the European Union