CPSS 2023

The 3rd Workshop on Computational Linguistics for the Political and Social Sciences

Proceedings of the Workshop

September 22, 2023

©2023 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 979-8-89176-032-5

Foreword

This workshop is the *3rd edition of Computational Linguistics for the Political and Social Sciences (CPSS)*. The main goal of the workshop is to bring together researchers and ideas from computational linguistics/NLP and the text-as-data community from political and social science, in order to foster collaboration and catalyze further interdisciplinary research efforts between these communities.

Submission format: We had two types of *workshop submissions*:

- archival papers describing original and unpublished work
- non-archival papers (abstracts) that present already published research or ongoing work.

This dichotomy met the different needs of researchers from different communities, allowing them to come together and exchange ideas in a "get to know each other" environment which was likely to foster collaborations.

Potential topics:

- Modeling political communication with NLP (e.g. topic classification, position measurement)
- Mining policy debates from heterogeneous textual sources
- Modeling complex social constructs (e.g. populism, polarization, identity) with NLP methods
- Political and social bias in language models
- Methodological insights in interdisciplinary collaboration: workflows, challenges, best practices
- NLP support to understand and support democratic decision making
- Resources and tools for Political/Social Science research
- ... and many more

Message from the Program Chairs

A special thank you goes to our amazing program committee for their invaluable input and dedication to the review process:

Ronja Sczepanski, ETH Zurich

Tobias Widman, Aarhus University

Indira Sen, GESIS

Robert Huber, USalzburg

Valerie Hase, LMU

Christian Rauh, WZB

Lisa Zehnter, HU Berlin/ Manifesto

Moritz Osnabrügge, Durham University

Agnieszka Falenska, UStuttgart

Dominik Stammbach, ETH Zurich

Ines Reinig, UMannheim

Maximilian Splithöver, UHannover

Tornike Tsereteli, UMannheim

Manfred Stede, UPotsdam

Eva Maria Vecchi, UStuttgart

Invited Speakers:

- **Lucie Flek**, University of Marburg and Bonn-Aachen International Center for Information Technology
  Title: On "fixing" the framing: Is bias detection a pair-wise comparison task?
  Abstract: NLP techniques for framing analysis can shed new light on fundamental questions in social science research, helping to understand biases towards entities and concepts e.g. in educational textbooks, news media or social networks. In this talk, I will present our studies on discovering indoctrination frames in historical event descriptions, showcasing their validity for contemporary media studies. I will further discuss the case for perspectivist annotation schemes and its relevance for the social grounding of LLMs.
- **Sebastian Padó**, University of Stuttgart
  Title: Computational construction of discourse networks for political debates.
  Abstract: Political debates form a crucial component of democratic decision processes and their structure and dynamics are highly interesting for political scientists. In this talk, I will present a series of interdisciplinary studies bringing together political scientists and NLP researchers with the goal of integrating manual and automatic approaches to constructing discourse networks. Topics include semi-automatic annotation, fairness, and hierarchical classification.

Organizing Committee

Christopher Klamm is an interdisciplinary PhD student at the University of Mannheim. His research efforts and interests are in the areas of Natural Language Processing and Computational Political Science, with a focus on automatic rhetoric and framing analysis.

Gabriella Lapesa is an interdisciplinary Computational Linguist working at the Institute for Natural Language Processing at the, currently leading the BMBF-funded independent research group E-DELIB, which works towards the development of NLP methods which would support deliberative discourse via (semi)automatic moderation.

Valentin Gold is a postdoctoral researcher at the Institute of Methods and Methodological Foundations in the Social Sciences at the University of Göttingen. He is currently coordinating the Deliberation Laboratory – an interdisciplinary project funded by the Volkswagen Foundation bringing together social science, computational linguistics and argument and ethos mining.

Theresa Gessler is an Assistant Professor for Comparative Politics at European University Viadrina in Frankfurt (Oder). Her research interests include the use of NLP methods to study political conflict around democracy, immigration and gender.

Simone Paolo Ponzetto holds the chair of Information Systems III (Enterprise Data Analysis) at the University of Mannheim, where he leads the Natural Language Processing and Information Retrieval group. His research interests include research on text understanding and its interdisciplinary application in the Social Sciences and Humanities.

Table of Contents

Automatic Emotion Experiencer Recognition
Maximilian Wegge and Roman Klinger 1

Personalized Intended and Perceived Sarcasm Detection on Twitter
Joan Plepi, Magdalena Buski and Lucie Flek 8

SpeakGer: A meta-data enriched speech corpus of German state and federal parliaments
Kai-Robin Lange and Carsten Jentsch 19

Multilabel Legal Element Classification on German Parliamentary Debates in a Low-Ressource Setting
Martin Hock and Christopher Klamm 29

Bubble up – A Fine-tuning Approach for Style Transfer to Community-specific Subreddit Language
Alessandra Zarcone and Fabian Kopf 46

According to BERTopic, what do Danish Parties Debate on when they Address Energy and Environment?
Costanza Navarretta and Dorte H. Hansen 59

The UNSC-Graph: An Extensible Knowledge Graph for the UNSC Corpus
Stian Rødven-Eide, Karolina Zaczynska, Antonio Pires, Ronny Patz and Manfred Stede 69

Deep Dive into the Language of International Relations: NLP-based Analysis of UNESCO's Summary Records
Joanna Wojciechowska, Mateusz Odrowaz-Sypniewski, Maria W. Smigielska, Igor Kaminski, Emilia Wiśnios, Bartosz Pieliński and Hanna Schreiber 75

Evaluating the Quality of the GermaParl Corpus of Plenary Protocols (v2.0.0)
Christoph Leonhardt and Andreas Blätte 88

Conference Program

Friday, September 22, 2023

9:00–9:15 ***Opening***

9:15–10:30 *Invited Talk by Sebastian Pado: Computational construction of discourse networks for political debates*

10:30–10:45 Coffee Break

10:45–12:00 Poster Session

Automatic Emotion Experiencer Recognition
Maximilian Wegge and Roman Klinger

Personalized Intended and Perceived Sarcasm Detection on Twitter
Joan Plepi, Magdalena Buski and Lucie Flek

SpeakGer: A meta-data enriched speech corpus of German state and federal parliaments
Kai-Robin Lange and Carsten Jentsch

Multilabel Legal Element Classification on German Parliamentary Debates in a Low-Ressource Setting
Martin Hock and Christopher Klamm

Bubble up – A Fine-tuning Approach for Style Transfer to Community-specific Subreddit Language
Alessandra Zarcone and Fabian Kopf

According to BERTopic, what do Danish Parties Debate on when they Address Energy and Environment?
Costanza Navarretta and Dorte H. Hansen

Friday, September 22, 2023 (continued)

12:00–13:30 Lunch Break

13:30–15:00 Panel Session

The UNSC-Graph: An Extensible Knowledge Graph for the UNSC Corpus
Stian Rødven-Eide, Karolina Zaczynska, Antonio Pires, Ronny Patz and Manfred Stede

Deep Dive into the Language of International Relations: NLP-based Analysis of UNESCO's Summary Records
Joanna Wojciechowska, Mateusz Odrowaz-Sypniewski, Maria W. Smigielska, Igor Kaminski, Emilia Wiśnios, Bartosz Pieliński and Hanna Schreiber

Evaluating the Quality of the GermaParl Corpus of Plenary Protocols (v2.0.0)
Christoph Leonhardt and Andreas Blätte

15:00–15:15 Coffee Break

15:15–16:30 Poster Session

16:30–17:45 *Invited Talk by Lucie Flek: On "fixing" the framing: Is bias detection a pair-wise comparison task?*

17:45–18:00 Closing

Friday, September 22, 2023 (continued)

18:30 Dinner