17th Conference on Natural Language Processing/Konferenz zur Verarbeitung naturlicher Sprache (KONVENS 2021)

Dusseldorf, Germany
6-9 September 2021

ISBN: 978-1-7138-3697-1

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2021) by the Association for Computational Linguistics
All rights reserved.
Copyright for individual papers remains with the authors.

Printed with permission by Curran Associates, Inc. (2021)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Long Papers

The Impact of Word Embeddings on Neural Dependency Parsing 1
Benedikt Adelmann, Wolfgang Menzel and Heike Zinsmeister

Benchmarking down-scaled (not so large) pre-trained language models 14
Matthias Aßenmacher, Patrick Schulze and Christian Heumann

ArgueBERT: How To Improve BERT Embeddings for Measuring the Similarity of Arguments . . 28
Maike Behrendt and Stefan Harmeling

How Hateful are Movies? A Study and Prediction on Movie Subtitles 37
Niklas von Boguszewski, Sana Moin, Anirban Bhowmick, Seid Muhie Yimam and Chris Biemann

Emotion Recognition under Consideration of the Emotion Component Process Model 49
Felix Casel, Amelie Heindl and Roman Klinger

Identifikation von Vorkommensformen der Lemmata in Quellenzitaten frühneuhochdeutscher Lexikoneinträge 62
Stefanie Dipper and Jan Christian Schaffert

Emotion Stimulus Detection in German News Headlines 73
Bao Minh Doan Dang, Laura Oberländer and Roman Klinger

Lexicon-based Sentiment Analysis in German: Systematic Evaluation of Resources and Preprocessing Techniques 86
Jakob Fehle, Thomas Schmidt and Christian Wolff

Definition Extraction from Mathematical Texts on Graph Theory in German and English 104
Theresa Kruse and Fritz Kliche

Extraction and Normalization of Vague Time Expressions in German 114
Ulrike May, Karolina Zaczynska, Julián Moreno-Schneider and Georg Rehm

Automatic Phrase Recognition in Historical German 127
Katrin Ortmann

Automatically Identifying Online Grooming Chats Using CNN-based Feature Extraction 137
Svenja Preuß, Tabea Bayha, Luna Pia Bley, Vivien Dehne, Alessa Jordan, Sophie Reimann, Fina Roberto, Josephine Romy Zahm, Hanna Siewerts, Dirk Labudde and Michael Spranger

Who is we? Disambiguating the referents of first person plural pronouns in parliamentary debates 147
Ines Rehbein, Josef Ruppenhofer and Julian Bernauer

Examining the Effects of Preprocessing on the Detection of Offensive Language in German Tweets 159
Sebastian Reimann and Daniel Dakota

Neural End-to-end Coreference Resolution for German in Different Domains 170
Fynn Schröder, Hans Ole Hatzel and Chris Biemann

How to Estimate Continuous Sentiments From Texts Using Binary Training Data 182
Sandra Wankmüller and Christian Heumann

forumBERT: Topic Adaptation and Classification of Contextualized Forum Comments in German 193
Ayush Yadav and Benjamin Milde

Short Papers

Robustness of end-to-end Automatic Speech Recognition Models – A Case Study using Mozilla DeepSpeech 203
Aashish Agarwal and Torsten Zesch

Effects of Layer Freezing on Transferring a Speech Recognition System to Under-resourced Languages 208
Onno Eberhard and Torsten Zesch

DeInStance: Creating and Evaluating a German Corpus for Fine-Grained Inferred Stance Detection 213
Anne Göhring, Manfred Klenner and Sophia Conrad

Combining text and vision in compound semantics: Towards a cognitively plausible multimodal model 218
Abhijeet Gupta, Fritz Günther, Ingo Plag, Laura Kallmeyer and Stefan Conrad

MobIE: A German Dataset for Named Entity Recognition, Entity Linking and Relation Extraction in the Mobility Domain 223
Leonhard Hennig, Phuc Tran Truong and Aleksandra Gabryszak

Automatically evaluating the conceptual complexity of German texts 228
Freya Hewett and Manfred Stede

WordGuess: Using Associations for Guessing, Learning and Exploring Related Words 235
Cennet Oguz, André Blessing, Jonas Kuhn and Sabine Schulte Im Walde

Towards a balanced annotated Low Saxon dataset for diachronic investigation of dialectal variation 242
Janine Siewert, Yves Scherrer and Jörg Tiedemann

German Abusive Language Dataset with Focus on COVID-19 247
Maximilian Wich, Svenja Räther and Georg Groh

Comparing Contextual and Static Word Embeddings with Small Philosophical Data 253
Wei Zhou and Jelke Bloem