

19th Conference on Natural Language Processing/Konferenz zur Verarbeitung natürlicher Sprache (KONVENS 2023)

Ingolstadt, Germany
18-23 September

ISBN: 979-8-3313-1645-7

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© (2023) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

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

<i>Is Prompt-Based Finetuning Always Better than Vanilla Finetuning? Insights from Cross-Lingual Language Understanding</i>	
Bolei Ma, Ercong Nie, Helmut Schmid and Hinrich Schuetze	1
<i>Comparing Pre-Training Schemes for Luxembourgish BERT Models</i>	
Cedric Lothritz, Saad Ezzini, Christoph Purschke, Tegawendé Bissyandé, Jacques Klein, Isabella Olariu, Andrey Boytsov, Clément LeFebvre and Anne Goujon	17
<i>LLpro: A Literary Language Processing Pipeline for German Narrative Texts</i>	
Anton Ehrmanntraut, Leonard Konle and Fotis Jannidis	28
<i>Classifying Noun Compounds for Present-Day Compositionality: Contributions of Diachronic Frequency and Productivity Patterns</i>	
Maximilian M. Maurer, Chris Jenkins, Filip Miletic and Sabine Schulte im Walde	40
<i>From Qualitative to Quantitative Research: Semi-Automatic Annotation Scaling in the Digital Humanities</i>	
Fynn Petersen-Frey, Tim Fischer, Florian Schneider, Isabel Eiser, Gertraud Koch and Chris Biemann	52
<i>Data and Approaches for German Text simplification – towards an Accessibility-enhanced Communication</i>	
Thorben Schomacker, Michael Gille, Marina Tropmann-Frick and Jörg von der Hülls	63
<i>Steps towards Addressing Text Classification in Low-Resource Languages</i>	
Maximilian Weißenbacher and Udo Kruschwitz	69
<i>Toward a Multilingual Connective Database: Aligning German/French Concessive Connectives</i>	
Peter Bourgonje, Sophia Rauh and Karolina Zaczynska	77
<i>Factuality Detection using Machine Translation – a Use Case for German Clinical Text</i>	
Mohammed Bin Sumait, Aleksandra Gabryszak, Leonhard Hennig and Roland A. Roller	85
<i>Linking Danish Parser Output to a Central Word Repository - From Morphosemantic Disambiguation to Unique Identifiers</i>	
Eckhard Bick	93
<i>Automatic Dictionary Generation: Could Brothers Grimm Create a Dictionary with BERT?</i>	
Hendryk Weiland, Maïke Behrendt and Stefan Harmeling	102
<i>Towards UkrainianWordNet: Incorporation of an Existing Thesaurus in the Domain of Physics</i>	
Melanie Siegel, Maksym Vakulenko and Jonathan Baum	121
<i>Pointer Networks: A Unified Approach to Extracting German Opinions</i>	
Julia Wunderle, Jan Pfister and Andreas Hotho	127
<i>Exploring Automatic Text Simplification of German Narrative Documents</i>	
Thorben Schomacker, Tillmann Dönicke and Marina Tropmann-Frick	139
<i>Are idioms surprising?</i>	
Jonas Nathanael Phlipp, Michael Richter, Erik Daas and Max Kölbl	149

<i>Answer Candidate Type Selection: Text-To-Text Language Model for Closed Book Question Answering Meets Knowledge Graphs</i>	
Mikhail Salnikov, Maria Lysyuk, Pavel Braslavski, Anton Razzhigaev, Valentin A. Malykh and Alexander Panchenko	155
<i>HS-EMO: Analyzing Emotions in Hate Speech</i>	
Johannes Schäfer and Elina Kistner	165
<i>Evaluating Data Augmentation Techniques for the Training of Luxembourgish Language Models</i>	
Isabella Olariu, Cedric Lothritz, Tegawendé Bissyandé and Jacques Klein	174
<i>Information Type Classification with Contrastive Task-Specialized Sentence Encoders</i>	
Philipp Seeberger, Tobias Bocklet and Korbinian Riedhammer	180
<i>German Text Embedding Clustering Benchmark</i>	
Silvan Wehrli, Bert Arnrich and Christopher Irrgang	187
<i>Aspect-Based Sentiment Analysis as a Multi-Label Classification Task on the Domain of German Hotel Reviews</i>	
Jakob Fehle, Leonie Münster, Thomas Schmidt and Christian Wolff	202
<i>Political claim identification and categorization in a multilingual setting: First experiments</i>	
Urs Zaberer, Sebastian Pado and Gabriella Lapesa	219