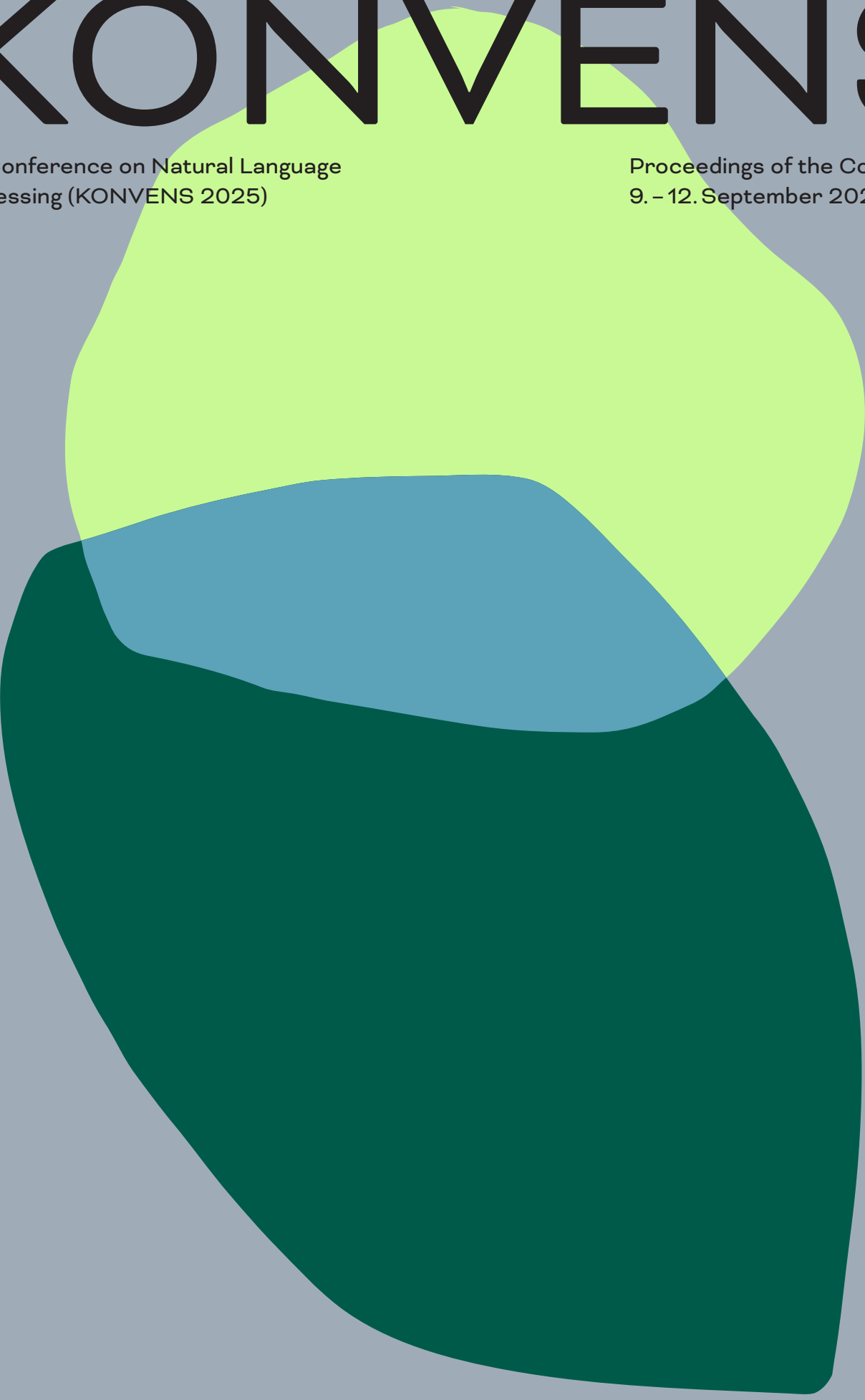


KONVENS



21st Conference on Natural Language
Processing (KONVENS 2025)

Proceedings of the Conference
9. – 12. September 2025

2025

Edited by Christian Wartena
and Ulrich Heid

VOLUME 2 -
WORKSHOPS

Imprint

Bibliographic Information of the German National Library

The German National Library lists this publication in the German National Bibliography. Detailed bibliographic data are available at <https://portal.dnb.de>.

Open Access

This work is published under the CC BY 4.0 license. Please note that certain parts of the work, marked accordingly, may be exempt from this license or subject to other copyright conditions.



HsH Applied Academics | Hannover 2025

Hochschule Hannover – University of Applied Sciences and Arts | Central Library

Ricklinger Stadtweg 118, 30459 Hannover, Germany

E-Mail: serwiss@hs-hannover.de

Website: <https://hs-h.de/bib>

BMFTR Project „DiamOnd@HsH – Towards a Science-Led Open Access Culture via Establishment of a Diamond Open Access University Press for the University of Applied Sciences and Arts Hannover” (16KOA042)



Cover Design: Lina Herschel (Hamburg, Germany), linaherschel.com

eISBN: 978-3-69018-016-0

DOI: <https://doi.org/10.25968/opus-3679>

URN: urn:nbn:de:bsz:960-opus4-36794

KONVENS 2025 is hosted by the University of Hildesheim and chaired by Christian Wartena (DataH Institute for Applied Data Science Hannover, Hochschule Hannover) and Ulrich Heid (Institut für Informationswissenschaft und Sprachtechnologie, Universität Hildesheim).



KONVENS was funded by the zukunft.niedersachsen program, in the program line 'Pro*Niedersachsen – Wissenschaftliche Veranstaltungen der Geistes-, Kultur- und Sozialwissenschaften in Niedersachsen'



Keynote speakers and tutorials were funded by the German Society for Computational Linguistics and Language Technology.



This publication was supported by the NiedersachsenOPEN publication fund, funded by zukunft.niedersachsen.



KONVENS 2025

Conference on Natural Language Processing



KONVENS (Konferenz zur Verarbeitung natürlicher Sprache) is an annual conference series on computational linguistics (biennial until 2018) that started in 1992 and that is organized under the auspices of the German Society for Computational Linguistics and Language Technology, the Special Interest Group on Computational Linguistics of the German Linguistic Society, the Austrian Society for Artificial Intelligence and SwissNLP.

The conference takes place in Hildesheim, Germany, from September 9th to 12th, 2025

Event Website: <https://konvens-2025.hs-hannover.de>

E-Mail: info.konvens2025@gsc1.org

ACL Anthology

The Proceedings of the 21th Conference on Natural Language Processing (KONVENS 2025) as well as the individual papers can also be found on ACL Anthology.

<https://aclanthology.org/events/konvens-2025>

Contents

CPSS	1
Analyzing Polarization in Online Discourse on the 2023-2024 Israel– Hamas War	7
<i>Daniel Miehling, Daniel Dakota, Sandra Kübler</i>	
How phatic is political communication in social media?	17
<i>Aenne Cecilia Kristine Knierim, Jannis Kuck, Ulrich Heid, Thomas Mandl</i>	
TRAVELWARN-Crawler: Constructing longitudinal datasets of government-issued travel warnings for political and social science research	29
<i>Laura Braun, Christian Oswald</i>	
Causally fooled in the name of being honest? Evaluating causal extraction in LLMs for political text	36
<i>Paulina Garcia-Corral, Hannah Dorothy Bechara, Krishnamoorthy Manohara, Slava Jankin</i>	
Heroes, Villains, and Victims: Character Narratives in the WPS Agenda of the UNSC	54
<i>Hannah Mathilde Steinbach, Imge Yüzüncüoğlu, Raluca Rilla, Manfred Stede</i>	
Enhancing Policy Analysis with NLP: A Reproducible Approach to Incentive Classification . .	74
<i>M.A. Waskow, John Philip McCrae</i>	
Narrative Structure Extraction in Disinformation and Trustworthy News: A Comparison of LLM, KG, and KG-Augmented Pipelines	86
<i>Justina Mandravickaitė</i>	
Abstractive Event Analysis of Armed Conflicts: Introducing the UCDP-AE Dataset	104
<i>Étienne Simon, Helene Bøsei Olsen, Ramón Carreño, Rahul Mishra, Nikolay Arefyev, Mert Can Yilmaz, Lilja Øvrelid, Erik Velldal</i>	
Cross-Platform Natural Language Processing Framework for Electoral Misinformation Detection	120
<i>Prasanth Yadla</i>	
Americans are dreamers – Generic statements and stereotyping in political tweets	128
<i>Ines Rehbein</i>	
KlarText	138
Developing A German Document-Level Parallel Dataset For Automatic Text Simplification To Generate Easy Language	142
<i>Vivien Jiranek, Stefan Hillmann</i>	
Modeling Quality of Experience in German Automatic Text Summarization and Machine Translation	169
<i>Dinh Nam Pham, Vivien Macketanz, Shushen Manakhimova, Sebastian Möller</i>	
Evaluation of Machine Translation Errors in German Plain Language Texts in the Domain of Health Information	176
<i>Sarah Ahrens, Silvana Deilen, Sergio Hernandez Garrido, Ekaterina Lapshinova-Koltunski, Christiane Maaß</i>	
A Resource of German Light Verb Constructions Along with Possible Alternative Formulations	186
<i>Ralf Laue</i>	
Automatic Compound Segmentation for Leichte Sprache	191
<i>Jesus Calvillo, Umesh Patil, Johann Seltmann, Anne-Kathrin Schumann</i>	

Automatic Generation of Comprehensible German Summaries for Domain-Specific Scientific Definitional Texts	202
<i>Sophie Decher, Michael Dembach</i>	
German Grammar Profile for Learners: Pedagogical Feature Definition and Automated Extraction	212
<i>Denise Löfflad, Benedikt Beuttler, Detmar Meurers</i>	
SustainEval & NLP4Sustain	224
Overview of the SustainEval 2025 Shared Task: Identifying the Topic and Verifiability of Sustainability Report Excerpts	229
<i>Jakob Prange, Charlott Jakob, Patrick Göttfert, Raphael Huber, Pia Wenzel Neves, An- nemarie Friedrich</i>	
EcoTUB @ SustainEval 2025: Ensembling BERT for German Sustainability Report Classification	239
<i>Sinan Bove, Icondy Kiba-Gassaye, Sirak Tadesse, Lisa Raithel</i>	
Reducing Environmental Costs whilst Maintaining Operational Effectiveness in Large Language Models through Chain Routing	250
<i>Solomon Wheeler, Melissa Torgbi, Harish Tayyar Madabushi</i>	
Contrasting Traditional Models and LLMs: An Evaluation Based on Text Segmentation	274
<i>Robin Jegan, Andreas Henrich</i>	
RAG Based Navigation of the World Ocean Assessment II	282
<i>Zainuddin Saiyed, Colin Orian, Poppy Nicolette Riddle, Gowri Prashanth Kanagaraj, Geoff Krause, Rémi Toupin, Philippe Mongeon, Stephen Brooks</i>	
A German-Language Dataset and Annotation Tool for Tabular Question Answering	291
<i>Staša Karapandžić, Michael Hellwig</i>	
Harmful Content Detection	301
Overview of the GermEval 2025 Shared Task on Harmful Content Detection	306
<i>Jenny Felser, Michael Spranger, Melanie Siegel</i>	
abullardUR@GermEval Shared Task 2025: Fine-tuning ModernGBERT on Highly Imbalanced German Social Media for Harmful Content	320
<i>Samuel Bullard</i>	
FI-CODE@GermEval Shared Task 2025: LLM Prompting for Augmentation of Underrepre- sented Classes	327
<i>Nina Seemann, Yeong Su Lee, Hendrik Bothe, Michaela Geierhos</i>	
Munich_Z@GermEval Shared Task 2025: When Prompting Is Not Enough: The Limits of Large Language Models in GermEval's 2025 Harmful Content Detection Task	337
<i>Florian Ludwig, Stefan Altmann</i>	
NLPeace@GermEval Shared Task 2025: Fine-Tuned BERT vs. Prompted LLMs for German Hate Speech Detection	350
<i>Patrick Göttfert, Raphael Huber, Fabio Mariani</i>	
nymera@GermEval Shared Task 2025: One Ensemble, Many Harms: A Unified Transformer Approach to Harmful Content Detection in German Social Media	357
<i>Hanna Köpcke</i>	
TheDBOs@GermEval Shared Task 2025: Supporting BERT Model Training with LLM-Based Synthetic Tweets for Harmful Content Detection	366
<i>Christoph Papadatos, Claudia Grote, Florian Ludwig</i>	
tweetbusters@GermEval Shared Task 2025: Comprehensive Feature Engineering for Call to Action detection	375
<i>Anastasia Schwichtenberg</i>	

Candy Speech Detection	385
Overview of the GermEval 2025 Shared Task on Candy Speech Detection	388
<i>Yulia Clausen, Tatjana Scheffler, Michael Wiegand</i>	
AIxcellent Vibes at GermEval 2025 Shared Task on Candy Speech Detection: Improving Model Performance by Span-Level Training	398
<i>Christian Rene Thelen, Patrick Gustav Blaneck, Tobias Bornheim, Niklas Grieger, Stephan Bialonski</i>	
Coling-UniA at GermEval 2025 Shared Task on Candy Speech Detection: Retrieval Augmented Generation for Identifying Expressions of Positive Attitudes in German YouTube Comments	404
<i>Georg Hofmann, Annemarie Friedrich</i>	
Flauschgummi at GermEval 2025 Shared Task on Candy Speech Detection: Sentiment Analysis and Classification of Online Comments	411
<i>Björn Kießwetter, Christoph Lukas, Tim Sieber</i>	
HHUflauschig at GermEval 2025 Shared Task on Candy Speech Detection: Hybrid Approaches for Binary Classification and Span Typing	416
<i>Wiebke Petersen, Lara Eulenpesch</i>	
NLP_Augsburg_04 at GermEval 2025 Shared Task on Candy Speech Detection: The Role of Surface Cues in Candy Speech Classification	424
<i>Evren Ataş, Malika Abitova, Fabio Mariani</i>	
NLPSuedwestfalen at GermEval 2025 Shared Task on Candy Speech Detection: Binary Classification of German YouTube Comments using Transformer Models	429
<i>Emir Bradaric, Manuel-Alexander Falk, Anna Hoff</i>	
Quabynar at GermEval 2025 Candy Speech Detection: Zero-shot Approach for Detecting Candy Speech	433
<i>Kwabena Odame Akomeah, Udo Kruschwitz, Bernd Ludwig, Kwame Boateng Akomeah</i>	
TUM NLP Group at GermEval 2025 Shared Task on Candy Speech Detection: Small Dose, Big Effect: Leveraging Synthetic Data for Candy Speech Detection	438
<i>Sophie Luise Opheiden, Elisei Shushpanov, Faeze Ghorbanpour</i>	
 LLM4Subjects	 444
Annif at the GermEval-2025 LLMs4Subjects Task: Traditional XMTC Augmented by Efficient LLMs	447
<i>Osmo Suominen, Juho Inkinen, Mona Lehtinen</i>	
DNB-AI-Project at the GermEval-2025 LLMs4Subjects Task: KIFSPrompt - Knowledge-Injected Few-Shot Prompting	455
<i>Maximilian Kähler, Lisa Kluge, Katja Konermann</i>	
Last Minute at the GermEval-2025 LLMs4Subjects Task: Few-Shot Contrastive Learning for Multilingual Multi-Label Classification	465
<i>Parisa Shiralil, Zahra Sarlak, Ebrahim Ansari</i>	
UBFFM at the GermEval-2025 LLMs4Subjects Task: What if we take "You are an expert in subject indexing" seriously?	471
<i>Clara Wan Ching Ho</i>	
 GermEval Multi-Task Papers	 479
Die SuperGLEBer at GermEval 2025 Shared Tasks: Growing Pains - When More Isn't Always Better	479
<i>Julia Wunderle, Jan Pfister, Andreas Hotho</i>	



VOLUME 2 - WORKSHOPS

KONVENS (Konferenz zur Verarbeitung natürlicher Sprache, Conference on Natural Language Processing) is an annual conference series on computational linguistics first held in 1992. Originally biennial until 2018, the conference is organized under the auspices of the German Society for Computational Linguistics and Language Technology, the Special Interest Group on Computational Linguistics of the German Linguistic Society, the Austrian Society for Artificial Intelligence, and SwissNLP. The 21st edition of KONVENS took place in Hildesheim, Germany, from September 9 – 12, 2025, and was jointly organized by the University of Hildesheim and the University of Applied Sciences and Arts Hannover.

This year's conference centered on artificial intelligence and large language models (LLMs), reflecting the scientific community's efforts to better understand and optimize the properties of LLMs. These topics were addressed in three plenary lectures, a lecture session on methodological issues, and a poster session on LLMs, as well as in sessions on other topics such as methods based on LLMs, discourse and semantics, multilinguality and translation, new language resources, applications of language technology beyond linguistics, and the identification of hate speech on social media.

The proceedings are published in two volumes. The first volume contains all the peer-reviewed papers from the main conference, and the second volume contains the papers from the six workshops and shared tasks organised as part of the conference.