1 Allgemeine Angaben

DFG-Geschäftszeichen: *KL 2869/1–1 und KL 2869/1–2*
Projektnummer: *667374*
Titel des Projekts: *Strukturierte Emotionsanalyse in Text in verschiedenen Domänen (Phase 1) /*
Komputationelle Ereignisbewertung auf Basis von Appraisaltheorien für Emotionsanalyse (Phase 2)
Name(n) des/r Antragstellenden: *Prof. Dr. Roman Klinger*
Dienstanschrift/en: *Otto-Friedrich-Universität Bamberg, Lehrstuhl für Grundlagen der Sprachverarbeitung, 96045 Bamberg, Germany*
Name(n) der Mitverantwortlichen: Sebastian Padó (erste Phase)
Name(n) der Kooperationspartnerinnen und –partner:
Berichtszeitraum (gesamte Förderdauer): 01.01.2018–31.12.2025

2 Summary

Zusammenfassung

Der Begriff „Emotionsanalyse“ in Texten umfasst verschiedene Aufgaben der natürlichen Sprachverarbeitung, die das gemeinsame Ziel verfolgen, Computern das Verständnis von Emotionen zu ermöglichen. Am verbreitetsten ist die Emotionsklassifikation, bei der einer vordefinierten Texteinheit eine oder mehrere Emotionen zugeordnet werden. Während diese Formulierung geeignet ist, die Emotion der Lesenden oder der Autorin bzw. des Autors zu identifizieren, erweitert das Emotion Role Labeling die Perspektive um die im Text erwähnten Entitäten und extrahiert Textspannen, die der Emotionsursache entsprechen.
Die zugrunde liegenden Emotionstheorien stimmen in einem wichtigen Punkt überein: Eine Emotion wird durch ein internes oder externes Ereignis verursacht und umfasst mehrere Teilkomponenten, darunter das subjektive Gefühl und eine kognitive Bewertung. Emotionen und Ereignisse stehen daher in zweifacher Weise miteinander in Beziehung. (1) Emotionen sind Ereignisse; diese Perspektive bildet die Grundlage der natürlichen Sprachverarbeitung für das Emotion Role Labeling. (2) Emotionen werden durch Ereignisse verursacht; diese Perspektive wird durch Forschung explizit gemacht, die untersucht, wie psychologische Bewertungstheorien (Appraisal-Theorien) in NLP-Modelle integriert werden können, um Ereignisse zu interpretieren.
In der ersten Phase dieses Projekts „Structured Multi-Domain Emotion Analysis from Text“ (SEAT) untersuchten wir (1). Hier entwickelten wir Textkorpora aus verschiedenen Domänen, die mit Rollen von Emotionserlebenden, Emotionsauslösern und im Text beschriebenen Zielen bzw. Stimuli annotiert wurden. Darüber hinaus führten wir Modellierungsexperimente zum Emotion Role Labeling durch und untersuchten dessen Mehrwert für die Emotionsklassifikation. Schließlich analysierten wir die Robustheit der Modelle über verschiedene Domänen hinweg.

In der zweiten Phase dieses Projekts „Computational Event Analysis based on Appraisal Theories for Emotion Analysis“ (CEAT) wandten wir uns (2) zu – der Untersuchung textueller Ereignisbeschreibungen und ihrer Bewertung durch Menschen, die diese Ereignisse erleben, sowie der Frage, ob menschliche Annotatorinnen und Annotatoren solche Bewertungen rekonstruieren können. Zu diesem Zweck entwickelten wir neuartige Annotierungsverfahren und darauf aufbauende Emotionsklassifikationsmodelle, die von der Vorhersage und Annotation von Bewertungsvariablen profitieren.

Summary

The term emotion analysis in text subsumes various natural language processing tasks which have in common the goal to enable computers to understand emotions. Most popular is emotion classification in which one or multiple emotions are assigned to a predefined textual unit. While such setting is appropriate for identifying the reader's or author's emotion, emotion role labeling adds the perspective of mentioned entities and extracts text spans that correspond to the emotion cause. The underlying emotion theories agree on one important point; that an emotion is caused by some internal or external event and comprises several subcomponents, including the subjective feeling and a cognitive evaluation. Emotions and events are therefore related in two ways. (1) Emotions are events; and this perspective is the fundament in natural language processing for emotion role labeling. (2) Emotions are caused by events; a perspective that is made explicit with research how to incorporate psychological appraisal theories in NLP models to interpret events.
In the first phase of this project, "Structured Multi- Domain Emotion Analysis from Text" (SEAT), we studied (1). Here, we developed text corpora from various domains, annotated with emotion experiencer roles, emotion cues, and targets/stimuli described in the text. We further conducted emotion role labeling modeling experiments and its value for emotion classification. Finally, we investigated how robust models are across domains.
In the second phase of this project, "Computational Event Analysis based on Appraisal Theories for Emotion Analysis (CEAT)", we moved to (2) – to study textual descriptions of events and how they are appraised by people living through them, and if human annotators can reconstruct such appraisals. To do so, we developed novel annotation procedures and developed emotion classification models on top of them which benefit from appraisal variable prediction and annotations.

3 Wissenschaftlicher Arbeits- und Ergebnisbericht

Starting Point and Goals of the Project

The starting point of the SEAT project (Phase 1) was the observation that sentiment analysis is well-defined and includes structured prediction setups such as aspect-based and targeted sentiment analysis. Emotion analysis has mostly been tackled as classification and regression. The goal of the SEAT

project was therefore, to transfer existing methods for structured sentiment analysis to emotion analysis, and do so in a cross-domain scenario.
As part of this work, we developed methods for emotion cause detection. We found that these causes often correspond to events, which are challenging for automatic emotion assessment. Therefore, we identified the starting point for CEAT (Phase 2), in which we aimed at developing event interpretation methods regarding their emotional connotation, with the help of appraisal theories. Here, we needed to understand if such variables can be reliably annotated and subsequently modeled with statistical methods.

Project-specific Results

In the first phase, we developed corpora for structured emotion analysis in literature (Kim/Klinger 2018, Kim/Klinger 2019) and news articles (Bostan/Kim/Klinger 2020), including German data (Doan Dang/Oberländer/Klinger 2021) and conducted a corpus comparison study (Bostan/Klinger 2018). We developed a generic modeling framework for structured text analysis, that has not only been applied in SEAT/CEAT, but also in the QUOTE project, which started in parallel (Adel et al. 2018). In line with the project goals, we focused in modeling on role labeling. In Oberländer/Klinger 2020, we compared clause classification with segmentation methods. With Doan Dang/Oberländer/Klinger 2021, we proposed methods for emotion stimulus detection in German. Oberländer/Reich/Klinger 2020 analyzed if role labeling is beneficial for emotion classification. Due to the paradigm shift to large language models and prompting during the project, we studied zero-shot emotion predictions (Plaza del Arco/Martín Valdivia/Klinger 2022; Bareiß/Klinger/Barnes 2024) and if fine-tuned models learn mostly topic distributions or actual emotion expressions (Wegge/Klinger 2024). Schäfer et al. 2025 investigated if LLM predictions are more appropriate for particular demographics in a society than others. As part of the modeling, we paid particular attention to emotion intensities (Strohm/Klinger 2018; Bostan/Klinger 2019; Sabbatino et al. 2022, Troiano et al. 2021).
During the work on structured prediction, we recognized the need to better understand how emotions are expressed. We therefore conducted studies on emotion communication channels (Kim/Klinger 2019/Casel/Heindl/Klinger 2011) and a comparison between German and English (Troiano/Padó/Klinger 2019).
This analysis led to the recognition of a need to look closer at Scherer's emotion component process model and emotion appraisals. We studied those first with an annotation and modeling publication (Hofmann et al. 2020/Hofmann et al. 2021), where we used external annotators to reconstruct the text author's event appraisals. This work motivated a central publication in our project, in which we analyzed if text actually carries sufficient information for appraisal annotations (Troiano/Oberländer/Klinger 2021). This work led to a set of follow-up studies in and outside our group, including novel annotation setups (Troiano et al., 2024).
We further combined appraisal-based approaches with structured emotion analysis, mostly to make the appraisal predictions for the correct emotion experiencer mentioned in the text (Troiano et al., 2022; Wegge et al. 2022, Wegge/Klinger 2023). Finally, we showed that appraisal predictions alone may not be sufficient, and subsequent emotion categorization also benefits from the analysis of appraisal trajectories (Wemmer/Labat/Klinger 2024; Schäfer/Wagner/Klinger 2026; Schäfer/Klinger 2026).

During the course of the project, we further published a set of papers that provide a big picture perspective, including a survey on emotion analysis in literature (Kim/Klinger 2019), Roman Klinger's habilitation thesis (Klinger 2020), and a position paper describing open tasks in event-centric emotion analysis (Klinger, 2023).

Changes in the project plan

In SEAT, we by and large followed the original project plan. We adapted to the use of Transformer models in addition to feature-based models. Further, we did conduct multilingual studies, but did spend less time than originally intended on cross-lingual modeling and emotion trajectories. This topic was moved to the second phase of the project.
In CEAT, the project followed the original plan with the adaptation of originally intended feature-based modeling to the use of large language models.

4 Veröffentliche Projektergebnisse

4.1. Publikationen mit wissenschaftlicher Qualitätssicherung

Florian Strohm and Roman Klinger. An empirical analysis of the role of amplifiers, downtoners, and negations in emotion classification in microblogs. In The 5th IEEE International Conference on Data Science and Advanced Analytics, Special Track on Sentiment, Emotion, and Credibility of Information in Social Data, DSAA, Turin, Italy, October 2018. IEEE. https://ieeexplore.ieee.org/document/8631476

Laura Ana Maria Bostan and Roman Klinger. An analysis of annotated corpora for emotion classification in text. In Proceedings of the 27th International Conference on Computational Linguistics, pages 2104--2119. Association for Computational Linguistics, 2018. https://aclanthology.org/C18-1179/

Evgeny Kim and Roman Klinger. Who feels what and why? annotation of a literature corpus with semantic roles of emotions. In Proceedings of the 27th International Conference on Computational Linguistics, pages 1345--1359. Association for Computational Linguistics, 2018. https://aclanthology.org/C18-1114/

Heike Adel, Laura Ana Maria Bostan, Sean Papay, Sebastian Padó, and Roman Klinger. DERE: A task and domain-independent slot filling framework for declarative relation extraction. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Brussels, Belgium, October, November 2018. Association for Computational Linguistics. https://aclanthology.org/D18-2008/

Laura Ana Maria Bostan and Roman Klinger. Exploring fine-tuned embeddings that model intensifiers for emotion analysis. In Proceedings of the Tenth Workshop on Computational

Approaches to Subjectivity, Sentiment and Social Media Analysis, Minneapolis, USA, June 2019. Association for Computational Linguistics. https://aclanthology.org/W19-1304/

Evgeny Kim and Roman Klinger. Frowning Frodo, wincing Leia, and a seriously great friendship: Learning to classify emotional relationships of fictional characters. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 647--653, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. https://aclanthology.org/N19-1067/

Enrica Troiano, Sebastian Padó, and Roman Klinger. Crowdsourcing and validating event-focused emotion corpora for German and English. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 4005--4011, Florence, Italy, July 2019. Association for Computational Linguistics. https://aclanthology.org/P19-1391/

Evgeny Kim and Roman Klinger. An analysis of emotion communication channels in fan-fiction: Towards emotional storytelling. In Proceedings of the Second Workshop on Storytelling, pages 56--64, Florence, Italy, 2019. Association for Computational Linguistics. https://aclanthology.org/W19-3406/

Evgeny Kim and Roman Klinger. A survey on sentiment and emotion analysis for computational literary studies. Zeitschrift fuer Digitale Geisteswissenschaften, 4, 2019. https://zfdg.de/2019_008

Laura Ana Maria Bostan, Evgeny Kim, and Roman Klinger. GoodNewsEveryone: A corpus of news headlines annotated with emotions, semantic roles, and reader perception. In Proceedings of The 12th Language Resources and Evaluation Conference, pages 1554--1566, Marseille, France, May 2020. European Language Resources Association. https://aclanthology.org/2020.lrec-1.194/

Roman Klinger. Structured Modelling of Affect in Text. Cumulative habilitation, University of Stuttgart, 2020. http://dx.doi.org/10.18419/opus-10994

Jan Hofmann, Enrica Troiano, Kai Sassenberg, and Roman Klinger. Appraisal theories for emotion classification in text. In Proceedings of the 28th International Conference on Computational Linguistics, pages 125--138, Barcelona, Spain (Online), December 2020. International Committee on Computational Linguistics. https://aclanthology.org/2020.coling-main.11/

Laura Ana Maria Oberländer and Roman Klinger. Token sequence labeling vs. clause classification for English emotion stimulus detection. In Proceedings of the Ninth Joint Conference on Lexical and Computational Semantics, pages 58--70, Barcelona, Spain (Online), December 2020. Association for Computational Linguistics. https://aclanthology.org/2020.starsem-1.7/

Laura Oberländer, Kevin Reich, and Roman Klinger. Experiencers, stimuli, or targets: Which semantic roles enable machine learning to infer the emotions? In Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media, Barcelona, Spain, December 2020. Association for Computational Linguistics. https://aclanthology.org/2020.peoples-1.12/

Enrica Troiano, Sebastian Padó, and Roman Klinger. Emotion ratings: How intensity, annotation confidence and agreements are entangled. In Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 40--49, Online, April 2021. Association for Computational Linguistics. https://aclanthology.org/2021.wassa-1.5/

Jan Hofmann, Enrica Troiano, and Roman Klinger. Emotion-aware, emotion-agnostic, or automatic: Corpus creation strategies to obtain cognitive event appraisal annotations. In Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 160--170, Online, April 2021. Association for Computational Linguistics. https://aclanthology.org/2021.wassa-1.17/

Felix Casel, Amelie Heindl, and Roman Klinger. Emotion recognition under consideration of the emotion component process model. In Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021), pages 49--61, Düsseldorf, Germany, 6--9 September 2021. KONVENS 2021 Organizers. https://aclanthology.org/2021.konvens-1.5/

Bao Minh Doan Dang, Laura Oberländer, and Roman Klinger. Emotion stimulus detection in German news headlines. In Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021), pages 73--85, Düsseldorf, Germany, 6--9 September 2021. KONVENS 2021 Organizers. https://aclanthology.org/2021.konvens-1.7/

Valentino Sabbatino, Enrica Troiano, Antje Schweitzer, and Roman Klinger. “splink” is happy and “phrouth” is scary: Emotion intensity analysis for nonsense words. In Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, pages 37--50, Dublin, Ireland, May 2022. Association for Computational Linguistics. https://aclanthology.org/2022.wassa-1.4/

Enrica Troiano, Laura Ana Maria Oberlaender, Maximilian Wegge, and Roman Klinger. x-envent: A corpus of event descriptions with experiencer-specific emotion and appraisal annotations. In Proceedings of the Language Resources and Evaluation Conference, pages 1365--1375, Marseille, France, June 2022. European Language Resources Association. https://aclanthology.org/2022.lrec-1.146/

Flor Miriam Plaza-del Arco, María-Teresa Martín-Valdivia, and Roman Klinger. Natural language inference prompts for zero-shot emotion classification in text across corpora. In Proceedings of the 29th International Conference on Computational Linguistics, pages 6805--6817, Gyeongju, Republic of Korea, October 2022. International Committee on Computational Linguistics. https://aclanthology.org/2022.coling-1.592/

Maximilian Wegge, Enrica Troiano, Laura Ana Maria Oberlaender, and Roman Klinger. Experiencer-specific emotion and appraisal prediction. In Proceedings of the Fifth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS), pages 25--32, Abu Dhabi, UAE, November 2022. Association for Computational Linguistics. https://aclanthology.org/2022.nlpcss-1.3/

Maximilian Wegge and Roman Klinger. Automatic emotion experiencer recognition. In 3rd Workshop on Computational Linguistics for the Political and Social Sciences (CPSS), May 2023. https://aclanthology.org/2023.cpss-1.1/

Enrica Troiano, Laura Oberländer, and Roman Klinger. Dimensional modeling of emotions in text with appraisal theories: Corpus creation, annotation reliability, and prediction. Computational Linguistics, 49(1), March 2023. https://direct.mit.edu/coli/article/49/1/1/113064/

Roman Klinger. Where are we in event-centric emotion analysis? bridging emotion role labeling and appraisal-based approaches. In Proceedings of the Big Picture Workshop: Crafting a Research Narrative, Singapore, December 2023. EMNLP, Association for Computational Linguistics. https://aclanthology.org/2023.bigpicture-1.1/

Patrick Bareiß, Roman Klinger, and Jeremy Barnes. English prompts are better for NLI-based zero-shot emotion classification than target-language prompts. In Companion Proceedings of the ACM on Web Conference 2024, WWW '24, page 1318–1326, New York, NY, USA, 2024. Association for Computing Machinery. https://dl.acm.org/doi/10.1145/3589335.3651902

Maximilian Wegge and Roman Klinger. Topic bias in emotion classification. In Rob van der Goot, JinYeong Bak, Max Müller-Eberstein, Wei Xu, Alan Ritter, and Tim Baldwin, editors, Proceedings of the Ninth Workshop on Noisy and User-generated Text (W-NUT 2024), pages 89--103, San Ġiljan, Malta, March 2024. Association for Computational Linguistics. https://aclanthology.org/2024.wnut-1.9/

Eileen Wemmer, Sofie Labat, and Roman Klinger. EmoProgress: Cumulated emotion progression analysis in dreams and customer service dialogues. In Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), pages 5660--5677, Torino, Italy, May 2024. ELRA and ICCL. https://aclanthology.org/2024.lrec-main.503/

Enrica Troiano, Sofie Labat, Marco Stranisci, Rossana Damiano, Viviana Patti, and Roman Klinger. Dealing with controversy: An emotion and coping strategy corpus based on role playing. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, Findings of the Association for Computational Linguistics: EMNLP 2024, pages 1634--1658, Miami, Florida, USA, November 2024. Association for Computational Linguistics. https://aclanthology.org/2024.findings-emnlp.89/

Johannes Schäfer, Aidan Combs, Christopher Bagdon, Jiahui Li, Nadine Probol, Lynn Greschner, Sean Papay, Yarik Menchaca Resendiz, Aswathy Velutharambath, Amelie Wuehrl, Sabine Weber, and Roman Klinger. Which demographics do LLMs default to during annotation? In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 17331--17348, Vienna, Austria, July 2025. Association for Computational Linguistics. https://aclanthology.org/2025.acl-long.848/

Johannes Schäfer and Roman Klinger. Disambiguation of emotion annotations by contextualizing events in plausible narratives. In Proceedings of the Language Resources and Evaluation Conference, Palma de Mallorca, Spain, May 2026. European Language Resources Association. https://arxiv.org/pdf/2508.09954

Johannes Schäfer, Janne Wagner, and Roman Klinger. Appraisal trajectories in narratives reveal distinct patterns of emotion evocation. In Proceedings of the 15th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, Rabat, Marocco, 2026. Association for Computational Linguistics. https://aclanthology.org/2026.wassa-1.7/

4.2 Weitere Publikationen und öffentlich gemachte Ergebnisse

- Unified Emotion Corpus:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/unifyemotion/
- Fan-Fiction Corpus:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/emotionchannels/
- Corpora on appraisal based emotion analysis:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/emotionappraisal/
- Emotion Component Corpus:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/cpmemotion/
- Emotion Relations in Literature:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/relationalemotions/
- REMAN:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/reman/
- Emotion Intensifier Corpus:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/modifier-emotion/
- Event-centered Emotion Corpora:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/deisear/
- GoodNewsEveryone:
  https://www.ims.uni-stuttgart.de/en/research/resources/corpora/goodnewseveryone/
- Emotion Intensity Data for Nonsense words:
  https://www.romanklinger.de/data-sets/nonsense-words-emotion-intensities.zip

- Gersti: https://www.ims.uni-stuttgart.de/forschung/ressourcen/korpora/gersti/
- Emotion coping: https://www.uni-bamberg.de/en/nlproc/resources/emotioncoping/
- Cumulative emotion progressions: https://www.uni-bamberg.de/en/nlproc/resources/emoprogress/
- Appraisal trajectories: https://www.uni-bamberg.de/en/nlproc/resources/emotional-backstories/