

LAW 2024

**The 18th Linguistic Annotation Workshop (LAW-XVIII)
Co-located with EACL 2024**

Proceedings of the Workshop

March 22, 2024

The LAW organizers gratefully acknowledge the support from the following sponsors.



©2024 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 979-8-89176-073-8

Introduction

Linguistic annotation of natural language corpora is the backbone of supervised methods of statistical natural language processing. The Linguistic Annotation Workshop (LAW) is the annual workshop of the ACL and ELRA Special Interest Group on Annotation (SIGANN), and it provides a forum for the presentation and discussion of innovative research on all aspects of linguistic annotation, including the creation and evaluation of annotation schemes, methods for automatic and manual annotation, use and evaluation of annotation software and frameworks, representation of linguistic data and annotations, semi-supervised “human in the loop” methods of annotation, crowd-sourcing approaches, and more.

As in the past, this year’s LAW provides a forum for annotation researchers to work towards standardization, best practices, and interoperability of annotation information and software.

These proceedings include papers that were presented at the 18th Linguistic Annotation Workshop (LAW-XVIII), co-located with EACL 2024 in St. Julian’s, Malta, on March 22, 2024.

This edition of the workshop is the eighteenth meeting of the ACL and ELRA Special Interest Group for Annotation. The first workshop took place in 2007 at the ACL in Prague. Since then, the LAW has been held every year, consistently drawing substantial participation (both in terms of paper/poster submissions and participation in the actual workshop) providing evidence that the LAW’s overall focus continues to be an important area of interest in the field, a substantial part of which relies on supervised learning from gold standard data sets (despite the advent of LLMs — see below). This year, we received 39 submissions, out of which 20 papers have been accepted to be presented at the workshop, as long or short papers, or as posters.

In addition, LAW-XVIII features an invited talk by Lilja Øvrelid (University of Oslo).

The special theme of LAW-XVIII is *Large Language Models*. In recent years, they have both transformed the field of NLP and alerted the general public to the astonishing powers of new language technologies. Needless to say, they will have an increasing impact also on matters of linguistic annotation; our special session of long papers is meant to stir discussions about prospects and limitations.

Our thanks go to SIGANN, our organizing committee, for their continuing organization of the LAW workshops. Most of all, we would like to thank all the authors for submitting their papers to the workshop and our program committee members for their dedication and their thoughtful reviews.

The LAW-XVIII Program Co-Chairs:
Sophie Henning and Manfred Stede

Organizing Committee

Program Chairs

Sophie Henning, Bosch Center for Artificial Intelligence and LMU Munich
Manfred Stede, University of Potsdam

SIGANN President

Amir Zeldes, Georgetown University

SIGANN Secretary

Ines Rehbein, Mannheim University

SIGANN Officers

Claire Bonial, US Army Research Laboratory
Stefanie Dipper, Ruhr-Universität Bochum
Annemarie Friedrich, University of Augsburg
Chu-Ren Huang, The Hong Kong Polytechnic University
Jena D. Hwang, Allen Institute for AI
Sandra Kübler, Indiana University
Lori Levin, Carnegie Mellon University
Adam Meyers, New York University
Antonio Pareja-Lora, Universidad de Alcalá (UAH) / FITISPos (UAH) / ATLAS (UNED) / DMEG (UdG)
Massimo Poesio, Queen Mary University of London and University of Utrecht
Sameer Pradhan, LDC, Cemantix
Jakob Prange, University of Augsburg
Nancy Ide, Vassar College/Brandeis University
Nathan Schneider, Georgetown University
Manfred Stede, University of Potsdam
Karin Tomanek, Google
Fei Xia, University of Washington
Nianwen Xue, Brandeis University
Deniz Zeyrek, Middle East Technical University

Program Committee

Chairs

Sophie Henning, Bosch Center for Artificial Intelligence and LMU Munich
Manfred Stede, University of Potsdam

Program Committee

Omri Abend, The Hebrew University of Jerusalem
Berfin Aktaş, University of Potsdam
Claire Bonial, US Army Research Lab
Miriam Butt, University of Konstanz
Marie Candito, LLF (Université Paris Cité / CNRS)
Kathryn Conger, University of Colorado, Boulder
Daniel Dakota, Indiana University
Stefanie Dipper, Ruhr University Bochum
Lucia Donatelli, Vrije Universiteit Amsterdam
Jonathan Dunn, University of Illinois Urbana-Champaign
Markus Egg, Humboldt-Universität zu Berlin
Annemarie Friedrich, University of Augsburg
Luke Gessler, University of Colorado, Boulder
Nizar Habash, New York University Abu Dhabi
Andrea Horbach, Universität Hildesheim
Nancy Ide, Vassar College/Brandeis University
René Knaebel, University of Potsdam
Sandra Kübler, Indiana University
Ekaterina Lapshinova-Koltunski, University of Hildesheim
Els Lefever, LT3, Ghent University
Adam Meyers, New York University
Alexis Palmer, University of Colorado Boulder
Antonio Pareja-Lora, Universidad de Alcalá (UAH) / FITISPos (UAH) / ATLAS (UNED) / DMEG (UdG)
Siyao Peng, LMU Munich
Massimo Poesio, Queen Mary University of London and University of Utrecht
Jakob Prange, University of Augsburg
Robin Schaefer, University of Potsdam
Nathan Schneider, Georgetown University
Ludovic Tanguy, CLLE: University of Toulouse / CNRS
Joel Tetreault, Dataminr
Lilian Diana Awuor Wanzare, Maseno University
Michael Wiegand, Alpen-Adria-Universität Klagenfurt
Mohammad Yeghaneh Abkenar, Bundesdruckerei Gruppe GmbH Berlin
Amir Zeldes, Georgetown University
Winnie Huiheng Zeng, Hong Kong Shue Yan University
Deniz Zeyrek, Middle East Technical University
Wei Zhou, Bosch Center for Artificial Intelligence
Heike Zinsmeister, Universität Hamburg

Keynote Talk: Who do you love? Fine-grained sentiment annotation, aggregation and data augmentation

Lilja Øvrelid

University of Oslo

Fri, March 22, 2024 – Time: 11:00 - 12:00 – Room: Radisson Marie Louise Suite 2

Abstract: In this talk I will discuss insights from several large-scale annotation projects aimed at enriching texts with subjective assessments describing their sentiment. In particular, I will describe a trajectory which is shared by many annotation projects aimed at supplying various Information Extraction systems with data: moving from fine-grained annotation of text spans at the sub-sentence level to aggregation of information within or across documents. I will start out by discussing the task of structured sentiment analysis and in particular, the 2022 SemEval shared task devoted to this task. In follow-up work, we assess the usefulness of these annotations for the aggregation of entity-level sentiment annotation for longer documents and provide a study of remaining challenges for the cross-sentence tracking of sentiment. Finally, I will briefly touch on recent work on data augmentation for sentiment analysis evaluating both masked language models and LLM-based prompting in the context of sentiment annotation.

Bio: Lilja Øvrelid is a Professor of Informatics at the University of Oslo, where she leads the Language Technology Group. She has been involved in a number of large annotation projects devoted to a diverse set of NLP tasks, such as UD syntax, NER, negation, sentiment analysis, coreference resolution and question-answering. She has worked on syntactic and semantic parsing, as well as the application of structured prediction approaches to NLP tasks beyond parsing, such as fine-grained sentiment analysis and event extraction.

Table of Contents

<i>TreeForm: End-to-end Annotation and Evaluation for Form Document Parsing</i> Ran Zmigrod, Zhiqiang Ma, Armineh Nourbakhsh and Sameena Shah	1
<i>Annotation Scheme for English Argument Structure Constructions Treebank</i> Hakyung Sung and Kristopher Kyle	12
<i>A Mapping on Current Classifying Categories of Emotions Used in Multimodal Models for Emotion Recognition</i> Ziwei Gong, Muyin Yao, Xinyi Hu, Xiaoning Zhu and Julia Hirschberg	19
<i>Surveying the FAIRness of Annotation Tools: Difficult to find, difficult to reuse</i> Ekaterina Borisova, Raia Abu Ahmad, Leyla Garcia-Castro, Ricardo Usbeck and Georg Rehm	29
<i>Automatic Annotation Elaboration as Feedback to Sign Language Learners</i> Alessia Battisti and Sarah Ebling	46
<i>Towards Better Inclusivity: A Diverse Tweet Corpus of English Varieties</i> Nhi Pham, Lachlan Pham and Adam Meyers	61
<i>Building a corpus for the anonymization of Romanian jurisprudence</i> Vasile Păiș, Dan Tufis, Elena Irimia and Verginica Barbu Mititelu	71
<i>Class Balancing for Efficient Active Learning in Imbalanced Datasets</i> Yaron Fairstein, Oren Kalinsky, Zohar Karnin, Guy Kushilevitz, Alexander Libov and Sofia Tolmach	77
<i>When is a Metaphor Actually Novel? Annotating Metaphor Novelty in the Context of Automatic Metaphor Detection</i> Sebastian Reimann and Tatjana Scheffler	87
<i>Enhancing Text Classification through LLM-Driven Active Learning and Human Annotation</i> Hamidreza Rouzegar and Masoud Makrehchi	98
<i>Using ChatGPT for Annotation of Attitude within the Appraisal Theory: Lessons Learned</i> Mirela Imamovic, Silvana Deilen, Dylan Glynn and Ekaterina Lapshinova-Koltunski	112
<i>Are You Serious? Handling Disagreement When Annotating Conspiracy Theory Texts</i> Ashley Hemm, Sandra Kübler, Michelle Seelig, John Funchion, Manohar Murthi, Kamal Premaratne, Daniel Verdear and Stefan Wuchty	124
<i>A GPT among Annotators: LLM-based Entity-Level Sentiment Annotation</i> Egil Rønningstad, Erik Velldal and Lilja Øvrelid	133
<i>Datasets Creation and Empirical Evaluations of Cross-Lingual Learning on Extremely Low-Resource Languages: A Focus on Comorian Dialects</i> Abdou Mohamed Naira, Benelallam Imade, Bahafid Abdessalam and Erraji Zakarya	140
<i>Prompting Implicit Discourse Relation Annotation</i> Frances Yung, Mansoor Ahmad, Merel Scholman and Vera Demberg	150
<i>PropBank goes Public: Incorporation into Wikidata</i> Elizabeth Spaulding, Kathryn Conger, Anatole Gershman, Mahir Morshed, Susan Windisch Brown, James Pustejovsky, Rosario Uceda-Sosa, Sijia Ge and Martha Palmer	166

<i>Reference and discourse structure annotation of elicited chat continuations in German</i>	
Katja Jasinskaja, Yuting Li, Fahime Same and David Uerlings	176
<i>Dependency Annotation of Ottoman Turkish with Multilingual BERT</i>	
Şaziye Özateş, Tarık Tıraş, Efe Genç and Esmâ Bilgin Tasdemir	188
<i>Donkii: Characterizing and Detecting Errors in Instruction-Tuning Datasets</i>	
Leon Weber, Robert Litschko, Ekaterina Artemova and Barbara Plank	197
<i>EEVEE: An Easy Annotation Tool for Natural Language Processing</i>	
Axel Sorensen, Siyao Peng, Barbara Plank and Rob Van Der Goot	216

Program

Friday, March 22, 2024

08:50 - 09:00 *Opening Remarks*

09:00 - 10:00 *Long Papers 1*

Automatic Annotation Elaboration as Feedback to Sign Language Learners
Alessia Battisti and Sarah Ebling

Enhancing Text Classification through LLM-Driven Active Learning and Human Annotation
Hamidreza Rouzegar and Masoud Makrehchi

Donkii: Characterizing and Detecting Errors in Instruction-Tuning Datasets
Leon Weber, Robert Litschko, Ekaterina Artemova and Barbara Plank

10:00 - 10:30 *Short Papers 1*

TreeForm: End-to-end Annotation and Evaluation for Form Document Parsing
Ran Zmigrod, Zhiqiang Ma, Armineh Nourbakhsh and Sameena Shah

Annotation Scheme for English Argument Structure Constructions Treebank
Hakyung Sung and Kristopher Kyle

10:30 - 11:00 *Coffee Break*

11:00 - 11:50 *Invited Talk – Lilja Øvrelid: Who do you love? Fine-grained sentiment annotation, aggregation and data augmentation*

11:50 - 12:50 *Long Papers 2*

Surveying the FAIRness of Annotation Tools: Difficult to find, difficult to reuse
Ekaterina Borisova, Raia Abu Ahmad, Leyla Garcia-Castro, Ricardo Usbeck and Georg Rehm

Class Balancing for Efficient Active Learning in Imbalanced Datasets
Yaron Fairstein, Oren Kalinsky, Zohar Karnin, Guy Kushilevitz, Alexander Libov and Sofia Tolmach

Friday, March 22, 2024 (continued)

Datasets Creation and Empirical Evaluations of Cross-Lingual Learning on Extremely Low-Resource Languages: A Focus on Comorian Dialects

Abdou Mohamed Naira, Benelallam Imade, Bahafid Abdessalam and Erraji Zakarya

12:50 - 14:00 *Lunch Break*

14:00 - 15:00 *In-Person Poster Session*

A Mapping on Current Classifying Categories of Emotions Used in Multimodal Models for Emotion Recognition

Ziwei Gong, Muyin Yao, Xinyi Hu, Xiaoning Zhu and Julia Hirschberg

Using ChatGPT for Annotation of Attitude within the Appraisal Theory: Lessons Learned

Mirela Imamovic, Silvana Deilen, Dylan Glynn and Ekaterina Lapshinova-Koltunski

A GPT among Annotators: LLM-based Entity-Level Sentiment Annotation

Egil Rønningstad, Erik Velldal and Lilja Øvrelid

PropBank goes Public: Incorporation into Wikidata

Elizabeth Spaulding, Kathryn Conger, Anatole Gershman, Mahir Morshed, Susan Windisch Brown, James Pustejovsky, Rosario Uceda-Sosa, Sijia Ge and Martha Palmer

Dependency Annotation of Ottoman Turkish with Multilingual BERT

Şaziye Özateş, Tarık Tıraş, Efe Genç and Esma Bilgin Tasdemir

EEVEE: An Easy Annotation Tool for Natural Language Processing

Axel Sorensen, Siyao Peng, Barbara Plank and Rob Van Der Goot

15:00 - 15:30 *Short Papers 2*

Building a corpus for the anonymization of Romanian jurisprudence

Vasile Păiș, Dan Tufis, Elena Irimia and Verginica Barbu Mititelu

Are You Serious? Handling Disagreement When Annotating Conspiracy Theory Texts

Ashley Hemm, Sandra Kübler, Michelle Seelig, John Funchion, Manohar Murthi, Kamal Premaratne, Daniel Verdear and Stefan Wuchty

Friday, March 22, 2024 (continued)

15:30 - 16:00 *Coffee Break*

16:00 - 17:00 *Long Papers 3*

When is a Metaphor Actually Novel? Annotating Metaphor Novelty in the Context of Automatic Metaphor Detection

Sebastian Reimann and Tatjana Scheffler

Prompting Implicit Discourse Relation Annotation

Frances Yung, Mansoor Ahmad, Merel Scholman and Vera Demberg

Reference and discourse structure annotation of elicited chat continuations in German

Katja Jasinskaja, Yuting Li, Fahime Same and David Uerlings

17:00 - 17:10 *Short Talk*

Towards Better Inclusivity: A Diverse Tweet Corpus of English Varieties

Nhi Pham, Lachlan Pham and Adam Meyers

17:10 - 17:20 *Closing Remarks*