

LAW-XIX 2025

**Proceedings of the 19th Linguistic Annotation Workshop  
(LAW-XIX-2025)**

**Proceedings of the Workshop**

July 31, 2025

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



©2025 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
317 Sidney Baker St. S  
Suite 400 - 134  
Kerrville, TX 78028  
USA  
Tel: +1-855-225-1962  
[acl@aclweb.org](mailto:acl@aclweb.org)

ISBN 979-8-89176-262-6

## 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 19th Linguistic Annotation Workshop (LAW-XIX), co-located with ACL 2025 in Vienna, Austria, on July 31, 2025.

This edition of the workshop is the nineteenth 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 and trustworthy evaluation in the era of Large Language Models. This year, we received 66 submissions, out of which 30 papers have been accepted to be presented at the workshop, as long or short papers, or as posters.

In addition, LAW-XIX features two invited talks by Junyi Jessy Li (University of Texas at Austin) and Rotem Dror (Haifa University).

The special theme of LAW-XIX is *Subjectivity and Variation in Linguistic Annotation*. As linguistic annotation increasingly supports diverse NLP applications, questions of annotator subjectivity, inter-annotator variation, and annotation uncertainty have become central to the field. Our special oral sessions aim to stimulate discussions on the challenges, methodological advances, and theoretical implications related to disagreement and subjectivity in annotation practice.

Our thanks go to SIGANN for their financial support and to 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, our program committee members for their dedication and their thoughtful reviews and our keynote speakers for sharing their insights on the topics of subjectivity, variation and social biases in linguistic annotation.

**The LAW-XIX Program Co-Chairs:**  
Siyao Peng and Ines Rehbein

# Organizing Committee

## Program Chairs

Siyao Peng, LMU Munich  
Ines Rehbein, University of Mannheim

## SIGANN President

Amir Zeldes, Georgetown University

## SIGANN Secretary

Ines Rehbein, University of Mannheim

## SIGANN Officers

Claire Bonial, US Army Research Laboratory  
Stefanie Dipper, Ruhr-Universität Bochum  
Annemarie Friedrich, University of Augsburg  
Sophie Henning, Bosch Center for Artificial Intelligence  
Chu-Ren Huang, The Hong Kong Polytechnic University  
Jenna D. Hwang, Allen Institute for AI  
Nancy Ide, Vassar College  
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)  
Siyao Peng, LMU Munich  
Massimo Poesio, Queen Mary University of London and University of Utrecht  
Sameer Pradhan, LDC, Cemantix  
Jakob Prange, University of Augsburg  
Ines Rehbein, University of Mannheim  
Nathan Schneider, Georgetown University  
Manfred Stede, University of Potsdam  
Katrín Tomanek, Google  
Fei Xia, University of Washington  
Nianwen Xue, Brandeis University  
Amir Zeldes, Georgetown University  
Deniz Zeyrek, Middle East Technical University  
Heike Zinsmeister, Hamburg University

# Program Committee

## Chairs

Siyao Peng, LMU Munich  
Ines Rehbein, Mannheim University

## Program Committee

Maria Becker, Heidelberg University  
Verena Blaschke, LMU Munich  
Claire Bonial, Army Research Lab  
Miriam Butt, Konstanz University  
Daniel Dakota, Indiana University  
Marie-Catherine de Marneffe, FNRS - UCLouvain  
Lucia Donatelli, Vrije Universiteit Amsterdam  
Luise Dürlich, Uppsala University  
Jonathan Dunn, University of Illinois, Urbana-Champaign  
Pablo Picasso Feliciano de Faria, University of Campinas  
Annemarie Friedrich, Augsburg University  
Kim Gerdes, Université Paris-Saclay  
Luke Gessler, Indiana University Bloomington  
Ziwei Gong, Columbia University  
Pingjun Hong, LMU Munich  
Chengzhi Martin Hu, LMU Munich  
Nancy Ide, Vassar College  
Maxim Ionov, University of Cologne  
Sandra Kübler, Indiana University  
Ekaterina Lapshinova-Koltunski, University of Hildesheim  
Ji-Ung Lee, Saarland University  
Els Lefever, Ghent University  
Lori Levin, Carnegie Mellon University  
Lauren Levine, Georgetown University  
Sebastian Loftus, LMU Munich  
Abdou Mohamed Naira, SI2M Lab, INSEA  
Anna Nedoluzhko, Charles University  
Ercong Nie, LMU Munich  
Juhyun Oh, KAIST  
Teresa Paccosi, University of Trento  
Alexis Palmer, University of Boulder  
Antonio Pareja-Lora, University of Alcalá  
Massimo Poesio, Utrecht University  
Jakob Prange, Augsburg University  
James Pustejovsky, Brandeis University  
Sebastian Reimann, University of Jena  
Josef Ruppenhofer, University of Hagen  
Egil Rønningstad, University of Oslo  
Tatjana Scheffler, University of Bochum  
Nathan Schneider, Georgetown University  
Merel Scholman, Utrecht University

Wesley Scivetti, Georgetown University  
Indira Sen, Mannheim University  
Soh-Eun Shim, LMU Munich  
Manfred Stede, Potsdam University  
Hakyung Sung, University of Oregon  
Andreas Säuberli, LMU Munich  
Ludovic Tanguy, CNRS, University of Toulouse  
Joel Tetreault, Dataminr  
Lilian Diana Awuor Wanzare, Maseno University  
Bonnie Webber, University of Edinburgh  
Leonie Weissweiler, UT Austin  
Michael Wiegand, University of Vienna  
Mohammad Yeghaneh Abkenar, Potsdam University  
Karolina Zaczynska, Potsdam University  
Amir Zeldes, Georgetown University  
Deniz Zeyrek, Middle East Technical University  
Wei Zhou, RWTH Aachen  
Heike Zinsmeister, Hamburg University  
Longfei Zuo, LMU Munich  
Şaziye Betül Özateş, Boğaziçi University

# Keynote Talk

## Data Annotation in the Era of LLMs - Thoughts and Good Practices

**Rotem Dror**

University of Haifa

**2025-07-31 09:00:00 – Room: Room 1.15-16**

**Abstract:** The rise of large language models (LLMs) presents both opportunities and challenges for data annotation in NLP. In this talk, I will explore the evolving role of LLMs as annotators, particularly in tasks involving subjectivity. I will present recent work that will also be presented in the conference on how to evaluate whether an LLM is a good annotator: "The Alternative Annotator Test for LLM-as-a-Judge: How to Statistically Justify Replacing Human Annotators with LLMs"—highlighting methods introduced in our paper—and discuss how LLMs compare to human annotators in consistency and reliability. I will also introduce new, unpublished research on best practices for identifying when and how LLMs can serve as reliable annotators for subjective NLP tasks. The talk aims to provide both theoretical insights and practical guidance for researchers and practitioners rethinking annotation pipelines in the LLM era.

**Bio:** Dr. Dror is an Assistant Professor (Senior Lecturer) at the Department of Information Systems, University of Haifa. She completed her Postdoctoral Research at the Cognitive Computation Group at the Department of Computer and Information Science, University of Pennsylvania, working with Prof. Dan Roth. She completed her Ph.D. in the Natural Language Processing Group, supervised by Prof. Roi Reichart, at the Faculty of Industrial Engineering and Management at the Technion - Israel Institute of Technology. Her research involves developing statistically sound methodologies for empirical investigation and evaluation for Data Science with a focus on Natural Language Processing applications.



# Keynote Talk

## Engaging experts and LLMs in corpora development

Junyi Jessy Li

University of Texas at Austin

2025-07-31 16:00:00 – Room: Room 1.15-16

**Abstract:** Large language models (LLMs) have become ever more capable, surpassing human performance on a number of tasks. Recent findings showed that LLMs can effectively replace traditional crowdsourcing to a large extent, and model training has increasingly been driven by synthetically generated data. These developments have triggered new questions about corpora development. This talk explores two of them: First, what type of human annotation can still be useful? I discuss our efforts engaging human expertise to effectively capture implicit reasoning in discourse and pragmatics, revealing weaknesses in existing models in those aspects. Second, how can we leverage LLMs to reveal task nuances that may be unknown before annotation? I present Explanation-Based Rescaling (EBR), a method that uses an LLM to rescale coarse-grained human ratings into consistent, fine-grained scores using natural language explanations from annotators, while discerning task subtleties embedded in these explanations.

**Bio:** Jessy Li is an Associate Professor in the Linguistics Department at the University of Texas at Austin. She received her Ph.D. (2017) from the Department of Computer and Information Science at the University of Pennsylvania. Her research interests are in computational linguistics and NLP, specifically discourse and document-level processing, natural language generation, and pragmatics. She is a recipient of an NSF CAREER Award, ACL and EMNLP Outstanding Paper Awards, an ACM SIGSOFT Distinguished Paper Award, among other honors. Jessy is the current Secretary of NAACL.

## Table of Contents

<i>Understanding Disagreement: An Annotation Study of Sentiment and Emotional Language in Environmental Communication</i>	
Christina Barz, Melanie Siegel, Daniel Hanss and Michael Wiegand . . . . .	1
<i>Measuring Label Ambiguity in Subjective Tasks using Predictive Uncertainty Estimation</i>	
Richard Alies, Elena Merdjanovska and Alan Akbik . . . . .	21
<i>Disagreements in analyses of rhetorical text structure: A new dataset and first analyses</i>	
Freya Hewett and Manfred Stede . . . . .	35
<i>Subjectivity in the Annotation of Bridging Anaphora</i>	
Lauren Levine and Amir Zeldes . . . . .	48
<i>The revision of linguistic annotation in the Universal Dependencies framework: a look at the annotators' behavior</i>	
Magali Sanches Duran, Lucelene Lopes and Thiago Alexandre Salgueiro Pardo . . . . .	60
<i>Forbidden FRUIT is the Sweetest: An Annotated Tweets Corpus for French Unfrozen Idioms Identification</i>	
Julien Bezançon, Gaël Lejeune, Antoine Gautier, Marceau Hernandez and Félix Alié . . . . .	70
<i>Another Approach to Agreement Measurement and Prediction with Emotion Annotations</i>	
Quanqi Du and Veronique Hoste . . . . .	87
<i>Harmonizing Divergent Lemmatization and Part-of-Speech Tagging Practices for Latin Participles through the LiLa Knowledge Base</i>	
Marco Passarotti, Federica Iurescia and Paolo Ruffolo . . . . .	103
<i>UD-KSL Treebank v1.3: A semi-automated framework for aligning XPOS-extracted units with UPOS tags</i>	
Hakyung Sung, Gyu-Ho Shin, Chanyoung Lee, You Kyung Sung and Boo Kyung Jung . . . . .	115
<i>Bootstrapping UMRs from Universal Dependencies for Scalable Multilingual Annotation</i>	
Federica Gamba, Alexis Palmer and Daniel Zeman . . . . .	126
<i>Classifying TEI Encoding for DutchDraCor with Transformer Models</i>	
Florian Debaene and Veronique Hoste . . . . .	137
<i>Label Bias in Symbolic Representation of Meaning</i>	
Marie Mikulová, Jan Štěpánek and Jan Hajič . . . . .	142
<i>An Annotation Protocol for Diachronic Evaluation of Semantic Drift in Disability Sources</i>	
Nitisha Jain, Chiara Di Bonaventura, Albert Merono Penuela and Barbara McGillivray . . . . .	160
<i>Pre-annotation Matters: A Comparative Study on POS and Dependency Annotation for an Alsatian Dialect</i>	
Delphine Bernhard, Nathanaël Beiner and Barbara Hoff . . . . .	173
<i>Where it's at: Annotating Verb Placement Types in Learner Language</i>	
Josef Ruppenhofer, Annette Annette Portmann, Christine Renker, Matthias Schwendemann, Katrin Wisniewski and Torsten Zesch . . . . .	187
<i>ICLE-RC: International Corpus of Learner English for Relative Clauses</i>	
Debopam Das, Izabela Czerniak and Peter Bourgonje . . . . .	201

<i>ExpLay: A new Corpus Resource for the Research on Expertise as an Influential Factor on Language Production</i>	
Carmen Schacht and Renate Delucchi Danhier . . . . .	216
<i>Towards Resource-Rich Mizo and Khasi in NLP: Resource Development, Synthetic Data Generation and Model Building</i>	
Soumyadip Ghosh, Henry Lalsiam, Dorothy Marbaniang, Gracious Mary Temsen, Rahul Mishra and Parameswari Krishnamurthy . . . . .	228
<i>Creating Hierarchical Relations in a Multilingual Event-type Ontology</i>	
Zdeňka Urešová, Eva Fučíková and Jan Hajič . . . . .	240
<i>Visual Representations of Temporal Relations between Events and Time Expressions in News Stories</i>	
Evelin Amorim, António Leal, Nana Yu, Purificação Moura Silvano and Alípio Mario Jorge .	250
<i>Annotating candy speech in German YouTube comments</i>	
Yulia Clausen and Tatjana Scheffler . . . . .	264
<i>Variety delights (sometimes) - Annotation differences in morphologically annotated corpora</i>	
Andrea Dömötör, Balázs Indig and Dávid Márk Nemeskey . . . . .	270
<i>Addressing Variability in Interlinear Glossed Texts with Linguistic Linked Data</i>	
Maxim Ionov and Natalia Patiño Mazzotti . . . . .	279
<i>Illuminating Logical Fallacies with the CAMPFIRE Corpus</i>	
Austin Blodgett, Claire Bonial, Taylor A. Pellegrin, Melissa Torgbi and Harish Tayyar Madabushi	285
<i>Cheap Annotation of Complex Information: A Study on the Annotation of Information Status in German TEDx Talks</i>	
Carmen Schacht, Tobias Nischk, Oleksandra Yazdanfar and Stefanie Dipper . . . . .	297
<i>Annotating Spatial Descriptions in Literary and Non-Literary Text</i>	
Emilie Sitter, Omar Momen, Florian Steig, J. Berenike Herrmann and Sina Zarriß . . . . .	308
<i>A GitHub-based Workflow for Annotated Resource Development</i>	
Brandon Waldon and Nathan Schneider . . . . .	326
<i>Enhancing an Annotation Scheme for Clinical Narratives in Portuguese through Human Variation Analysis</i>	
Ana Luisa Fernandes, Purificação Silvano, António Leal, Nuno Guimarães, Rita Rb-Silva, Luís Filipe Cunha and Alípio Jorge . . . . .	332
<i>Expanding the UNSC Conflicts Corpus by Incorporating Domain Expert Annotations and LLM Experiments</i>	
Karolina Zaczynska . . . . .	344
<i>Guidelines for Fine-grained Sentence-level Arabic Readability Annotation</i>	
Nizar Habash, Hanada Taha-Thomure, Khalid Elmadani, Zeina Zeino and Abdallah Abushmaes	359

# Program

**Thursday, July 31, 2025**

- 08:45 - 09:00     *Opening Remarks*
- 09:00 - 09:45     *Keynote 1 – Rotem Dror: Data Annotation in the Era of LLMs - Thoughts and Good Practices*
- 09:45 - 10:30     *Oral Session 1: Disagreement and ambiguity in highly subjective tasks*
- Understanding Disagreement: An Annotation Study of Sentiment and Emotional Language in Environmental Communication*  
Christina Barz, Melanie Siegel, Daniel Hanss and Michael Wiegand
- Measuring Label Ambiguity in Subjective Tasks using Predictive Uncertainty Estimation*  
Richard Alies, Elena Merdjanovska and Alan Akbik
- Disagreements in analyses of rhetorical text structure: A new dataset and first analyses*  
Freya Hewett and Manfred Stede
- 10:30 - 11:00     *Coffee*
- 11:00 - 11:30     *Oral Session 2: Subjectivity in linguistic annotation*
- Subjectivity in the Annotation of Bridging Anaphora*  
Lauren Levine and Amir Zeldes
- The revision of linguistic annotation in the Universal Dependencies framework: a look at the annotators' behavior*  
Magali Sanches Duran, Lucelene Lopes and Thiago Alexandre Salgueiro Pardo
- 11:30 - 12:30     *Poster Session 1*
- Forbidden FRUIT is the Sweetest: An Annotated Tweets Corpus for French Un-frozen Idioms Identification*  
Julien Bezançon, Gaël Lejeune, Antoine Gautier, Marceau Hernandez and Félix Alié
- Another Approach to Agreement Measurement and Prediction with Emotion Annotations*  
Quanqi Du and Veronique Hoste

**Thursday, July 31, 2025 (continued)**

*Harmonizing Divergent Lemmatization and Part-of-Speech Tagging Practices for Latin Participles through the LiLa Knowledge Base*

Marco Passarotti, Federica Iurescia and Paolo Ruffolo

*UD-KSL Treebank v1.3: A semi-automated framework for aligning XPOS-extracted units with UPOS tags*

Hakyung Sung, Gyu-Ho Shin, Chanyoung Lee, You Kyung Sung and Boo Kyung Jung

*Bootstrapping UMRs from Universal Dependencies for Scalable Multilingual Annotation*

Federica Gamba, Alexis Palmer and Daniel Zeman

*Classifying TEI Encoding for DutchDraCor with Transformer Models*

Florian Debaene and Veronique Hoste

*Label Bias in Symbolic Representation of Meaning*

Marie Mikulová, Jan Štěpánek and Jan Hajič

*An Annotation Protocol for Diachronic Evaluation of Semantic Drift in Disability Sources*

Nitisha Jain, Chiara Di Bonaventura, Albert Merono Penuela and Barbara McGillivray

*Pre-annotation Matters: A Comparative Study on POS and Dependency Annotation for an Alsatian Dialect*

Delphine Bernhard, Nathanaël Beiner and Barbara Hoff

*Where it's at: Annotating Verb Placement Types in Learner Language*

Josef Ruppenhofer, Annette Annette Portmann, Christine Renker, Matthias Schwendemann, Katrin Wisniewski and Torsten Zesch

*ICLE-RC: International Corpus of Learner English for Relative Clauses*

Debopam Das, Izabela Czerniak and Peter Bourgonje

*ExpLay: A new Corpus Resource for the Research on Expertise as an Influential Factor on Language Production*

Carmen Schacht and Renate Delucchi Danhier

12:30 - 14:00

*Lunch*

**Thursday, July 31, 2025 (continued)**

14:00 - 15:00 *Poster Session 2*

*Towards Resource-Rich Mizo and Khasi in NLP: Resource Development, Synthetic Data Generation and Model Building*

Soumyadip Ghosh, Henry Lalsiam, Dorothy Marbaniang, Gracious Mary Temsen, Rahul Mishra and Parameswari Krishnamurthy

*Creating Hierarchical Relations in a Multilingual Event-type Ontology*

Zdeňka Urešová, Eva Fučíková and Jan Hajič

*Visual Representations of Temporal Relations between Events and Time Expressions in News Stories*

Evelin Amorim, António Leal, Nana Yu, Purificação Moura Silvano and Alipio Mario Jorge

*Annotating candy speech in German YouTube comments*

Yulia Clausen and Tatjana Scheffler

*Variety delights (sometimes) - Annotation differences in morphologically annotated corpora*

Andrea Dömötör, Balázs Indig and Dávid Márk Nemeskey

*Addressing Variability in Interlinear Glossed Texts with Linguistic Linked Data*

Maxim Ionov and Natalia Patiño Mazzotti

*Illuminating Logical Fallacies with the CAMPFIRE Corpus*

Austin Blodgett, Claire Bonial, Taylor A. Pellegrin, Melissa Torgbi and Harish Tayyar Madabushi

*Cheap Annotation of Complex Information: A Study on the Annotation of Information Status in German TEDx Talks*

Carmen Schacht, Tobias Nischk, Oleksandra Yazdanfar and Stefanie Dipper

*Annotating Spatial Descriptions in Literary and Non-Literary Text*

Emilie Sitter, Omar Momen, Florian Steig, J. Berenike Herrmann and Sina Zariß

*A GitHub-based Workflow for Annotated Resource Development*

Brandon Waldon and Nathan Schneider

**Thursday, July 31, 2025 (continued)**

*Enhancing an Annotation Scheme for Clinical Narratives in Portuguese through Human Variation Analysis*

Ana Luisa Fernandes, Purificação Silvano, António Leal, Nuno Guimarães, Rita Rb-Silva, Luís Filipe Cunha and Alípio Jorge

15:00 - 15:30 *Oral Session 3: Datasets*

*Expanding the UNSC Conflicts Corpus by Incorporating Domain Expert Annotations and LLM Experiments*

Karolina Zaczynska

*Guidelines for Fine-grained Sentence-level Arabic Readability Annotation*

Nizar Habash, Hanada Taha-Thomure, Khalid Elmadani, Zeina Zeino and Abdallah Abushmaes

15:30 - 16:00 *Coffee*

16:00 - 16:45 *Keynote 2 – Junyi Jessy Li: Engaging experts and LLMs in corpora development*

16:45 - 17:00 *Wrap-up*