

NAACL 2025

**The 2025 Conference of the Nations of the Americas Chapter  
of the Association for Computational Linguistics**

**Proceedings of the Student Research Workshop**

April 30 - May 1, 2025

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

## **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 979-8-89176-192-6

## Introduction

Welcome to the NAACL 2025 Student Research Workshop.

The Student Research Workshop (SRW) is a workshop for student researchers in computational linguistics and natural language processing, and provides a unique opportunity for student participants to present their work and receive valuable feedback from the research community.

Continuing the tradition of previous student research workshops, we offer archival and non-archival tracks, and accept both research papers as well as thesis proposals in each track. The research paper track welcomes submissions from Ph.D. students, Masters students, and advanced undergraduate or high school students. Additionally, the thesis proposal submissions caters to advanced Masters and Ph.D. students who have identified their thesis topic, offering them a platform to receive feedback on their proposal and guidance on potential future avenues for their research.

This year, we received a record 169 submissions in total. Of the 145 valid submissions, we accepted 89 total, resulting in an acceptance rate of 61%. Out of the 89 accepted papers, 48 were archival research papers, 29 were non-archival research papers, 6 were archival thesis proposals, and 6 were non-archival thesis proposals.

Another core aspect of the SRW is mentoring. In line with previous years, we had a pre-submission mentoring program before the submission deadline. A total of 28 papers participated in the pre-submission mentoring program. This program offered students the opportunity to receive comments from an experienced researcher to improve the writing style and presentation of their submissions. We are incredibly grateful to all researchers who volunteered as mentors, particularly due the considerable increase in student requests this year.

We are immensely grateful to the Association for Computational Linguistics for their sponsorship. Their support has played a significant role in ensuring the success of the conference and has allowed a large number of students to publish their work and attend the conference. We also express our sincere gratitude to the program committee members for their thorough reviews of each paper. We are also deeply appreciative of the NAACL 2025 organizing committee for their ongoing support, and our faculty advisors Maria Pacheco and Shira Wein, for their valuable guidance which was invaluable to organizing this year's workshop. Lastly, we thank all the student authors for submitting their work and participating in the 2025 edition of the NAACL SRW.

# **Organizing Committee**

## **Student Chairs**

Abteen Ebrahimi, University of Colorado Boulder  
Emmy Liu, Carnegie Mellon University  
Samar Haider, University of Pennsylvania

## **Faculty Advisors**

Maria Leonor Pacheco, University of Colorado Boulder  
Shira Wein, Amherst College

## Program Committee

### Mentors

Valerio Basile  
David Chiang  
Brian Davis  
Chris Develder  
Luis Espinosa-Anke  
Dipesh Gautam  
Daphne Ippolito  
Najoung Kim  
Mascha Kurpicz  
Kaixin Ma  
Archita Pathak  
Philip Resnik  
Carolyn Rose  
Richard Sproat  
Andrea Varga  
Bonnie Webber

### Reviewers

Hadi Abdi Ghavidel, Noor Abo Mokh, Rodrigo Agerri, Dareen Safar Alharthi, Miguel A. Alonso, Tatiana Anikina, Stéphane Aroca-Ouellette

S R Balasundaram, Tadesse Destaw Belay, Peter Belcak, Samuel Belkadi, Lamia Benhiba, Himanshu Beniwal, Gabriel Bernier-Colborne, Steven Bethard, Henrike Beyer, Atharva Bhargude, Gagan Bhatia, Abari Bhattacharya, Chris Biemann, Nimet Beyza Bozdog, Sabur Butt

Allen Chang, Khaoula Chehbouni, David Chiang, Young Min Cho, Javier Conde

Zhongdongming Dai, Samvit Dammalapati, Brian Davis, Jacob Devasier, Chris Develder, Kaustubh Dhole, Cecilia Domingo, Ritam Dutt, Upasana Dutta

Carlos Escolano, Luis Espinosa-Anke, Shaun Esua-Mensah

Martin Fajčík, Amany Fashwan, Virginia K. Felkner, Dayne Freitag, Yicheng Fu, Richard Futrell

Manas Gaur, Dipesh Gautam, Kshitish Ghate, Kripabandhu Ghosh, Dhiman Goswami, Nidhi Goyal, Qiming Guo, Abhay Gupta

Kevin Han, Lifeng Han, Peitao Han, Chandler Haney, Youssef Al Hariri, Truong-Son Hy, Truong-Son Hy, Mika Härmäläinen

Helen Jin, Ashvini Kumar Jindal, Abhinav Joshi

Pritam Kadasi, Tomoyuki Kajiware, Katikapalli Subramanyam Kalyan, Hidetaka Kamigaito, Sudipta Kar, Daisuke Kawahara, Lemlem Eyob Kawo, Danush Khanna, Lev Kharlashkin, Grigori

Khvatskii, Philipp Koehn, Zhanibek Kozhimbayev, Mascha Kurpicz-Briki

Alexandra Lavrentovich, Yongjia Lei, Bryan Li, Lei Li, Lixiang Li, Austen Liao, Jasy Suet Yan Liew, Constantine Lignos, Jinhyeong Lim, Peerat Limkonchotiwat, Xuye Liu, Yijun Liu, Zhu Liu, Josh Magnus Ludan

Bolei Ma, Wolfgang Maass, Yunho Maeng, Valentin Malykh, Ali Marashian, John Philip McCrae, Nicolo Micheletti, Filip Miletić, Filip Miletić, Negar Mokhberian

Sachin Narayan Nagargoje, Atharva Naik, Shoichi Naito, Nihal V. Nayak, Hamada Nayel, Lilian Ngweta

Yusuke Oda, Atul Kr. Ojha, Şaziye Betül Özates

Artemis Panagopoulou, Alexander Panchenko, Tanmay Parekh, Hyeyoung Park, Namyoung Park, Ajay S Patil, Xin Peng, Lis Pereira, Vladia Pinheiro, Adithya Pratapa, Gabriele Prato, Priyanshu Priya, Rifki Afina Putri

Sree Harsha Ramesh, Philip Resnik, Enora Rice, Anthony Rios, João Victor Pessoa Rocha, Elisei Rykov

Ahmed Sabir, Parisa Safikhani, Harshita Sahijwani, Yusuke Sakai, Jonathan Sakunkoo, Ahnaf Mozib Samin, Prajvi Saxena, Michael Saxon, Rajiv Ratn Shah, Vishal Shah, Wan Jou She, Sina Sheikholeslami, Andrii Shportko, Chenglei Si, Afroz Ahamad Siddiqui, Harman Singh, Aryan Singhal, Tamar Solorio, Yejin Son, Yueqi Song, Richard Sproat, Yirong Sun, Sathya Krishnan Suresh, Nikita Sushko

Savannah Jennifer Thais, Uthayasanker Thayasivam, Nicholas Tomlin

Takehito Utsuro

Sowmya Vajjala, Andrea Varga, Prasoon Varshney, Ishwara Vasista, Supriti Vijay

Wenlu Wang, Yimu Wang, Yu Wang, Taro Watanabe, Bonnie Webber, Adam Wiemerslage, Steven R Wilson, Zach Wood-Doughty, Jian Wu, Yulong Wu

Yiqing Xie, Xi Xu

Ivory Yang, Ziyu Yao, Haotian Ye, Aditya Yedetore

Haiqi Zhang, Haoran Zhang, Chenyang Zhao, Xingmeng Zhao, Yang Zhong, Andrew Zhu, Yuqicheng Zhu, Zhengyuan Zhu

## Keynote Talk

**Philip Resnik**

University of Maryland, College Park



**Bio:** Philip Resnik is MPower Professor at University of Maryland with joint appointments in the Department of Linguistics and the Institute for Advanced Computer Studies. He earned his bachelor's in Computer Science at Harvard and his PhD in Computer and Information Science at the University of Pennsylvania, and does research in computational linguistics. Prior to joining UMD, he was an associate scientist at BBN, a graduate summer intern at IBM T.J. Watson Research Center (subsequently awarded an IBM Graduate Fellowship) while at UPenn, and a research scientist at Sun Microsystems Laboratories. In 2020 he was designated a Fellow of the Association for Computational Linguistics. Philip's most recent research has focused in two areas. One is the computational cognitive neuroscience of language, where he has been using computational modeling in connection with brain imaging to look at the role of context and predictive processing during online language comprehension. The other is computational social science, with an emphasis on connecting the signal available in people's language use with underlying mental state – this has applications in computational political science, particularly in connection with ideology, framing, and beliefs, and in mental health, focusing on the ways that linguistic behavior may help to identify and monitor depression, schizophrenia, and suicidality. Philip is a scientific advisor for NORC at the University of Chicago (a non-partisan, independent social research organization). In entrepreneurial life he was a technical co-founder of CodeRyte (NLP for electronic health records, acquired by 3M in 2012), and is an advisor to FiscalNote (machine learning and analytics for government relations, went public in 2022), and Trustible (a leading technology provider of responsible AI governance).



# Table of Contents

<i>Fine-Grained and Multi-Dimensional Metrics for Document-Level Machine Translation</i>	
Yirong Sun, Dawei Zhu, Yanjun Chen, Erjia Xiao, Xinghao Chen and Xiaoyu Shen . . . . .	1
<i>INSIGHTBUDDY-AI: Medication Extraction and Entity Linking using Pre-Trained Language Models and Ensemble Learning</i>	
Pablo Romero, Lifeng Han and Goran Nenadic . . . . .	18
<i>Linguistic Features in German BERT: The Role of Morphology, Syntax, and Semantics in Multi-Class Text Classification</i>	
Henrike Beyer and Diego Frassinelli . . . . .	28
<i>Thesis Proposal: Uncertainty in Knowledge Graph Embeddings</i>	
Yuqicheng Zhu . . . . .	40
<i>Detecting Sexism in Tweets: A Sentiment Analysis and Graph Neural Network Approach</i>	
Diana P. Madera-Espíndola, Zoe Caballero-Domínguez, Valeria J. Ramírez-Macías, Sabur Butt and Hector Ceballos . . . . .	48
<i>Towards Codec-LM Co-design for Neural Codec Language Models</i>	
Shih-Lun Wu, Aakash Lahoti, Arjun D Desai, Karan Goel, Chris Donahue and Albert Gu . . . .	55
<i>Low-resource Machine Translation for Code-switched Kazakh-Russian Language Pair</i>	
Maksim Borisov, Zhanibek Kozhirbayev and Valentin Malykh . . . . .	66
<i>Generative Product Recommendations for Implicit Superlative Queries</i>	
Kaustubh Dhole, Nikhita Vedula, Saar Kuzi, Giuseppe Castellucci, Eugene Agichtein and Shervin Malmasi . . . . .	77
<i>ConQuer: A Framework for Concept-Based Quiz Generation</i>	
Yicheng Fu, Zikui Wang, Liuxin Yang, Meiqing Huo and Zhongdongming Dai . . . . .	92
<i>What is it? Towards a Generalizable Native American Language Identification System</i>	
Ivory Yang, Weicheng Ma, Carlos Guerrero Alvarez, William Dinauer and Soroush Vosoughi	105
<i>Med-CoDE: Medical Critique based Disagreement Evaluation Framework</i>	
Mohit Gupta, Akiko Aizawa and Rajiv Ratn Shah . . . . .	112
<i>Sentimatic: Sentiment-guided Automatic Generation of Preference Datasets for Customer Support Dialogue System</i>	
Suhyun Lee and ChangHeon Han . . . . .	120
<i>Privacy-Preserving Federated Learning for Hate Speech Detection</i>	
Ivo De Souza Bueno Júnior, Haotian Ye, Axel Wisiolek and Hinrich Schuetze . . . . .	129
<i>From Annotation to Adaptation: Metrics, Synthetic Data, and Aspect Extraction for Aspect-Based Sentiment Analysis with Large Language Models</i>	
Nikita Neveditsin, Pawan Lingras and Vijay Kumar Mago . . . . .	142
<i>Developing Japanese CLIP Models Leveraging an Open-weight LLM for Large-scale Dataset Translation</i>	
Issa Sugiura, Shuhei Kurita, Yusuke Oda, Daisuke Kawahara and Naoaki Okazaki . . . . .	162
<i>Self-Vocabularizing Training for Neural Machine Translation</i>	
Pin-Jie Lin, Ernie Chang, Yangyang Shi and Vikas Chandra . . . . .	171

<i>CCT-Code: Cross-Consistency Training for Multilingual Clone Detection and Code Search</i>	
Nikita Sorokin, Tikhonov Anton, Dmitry Abulkhanov, Ivan Sedykh, Irina Piontkovskaya and Valentin Malykh .....	178
<i>Text Compression for Efficient Language Generation</i>	
David Gu, Peter Belcak and Roger Wattenhofer .....	186
<i>Multilingual Native Language Identification with Large Language Models</i>	
Dhiman Goswami, Marcos Zampieri, Kai North, Shervin Malmasi and Antonios Anastasopoulos	193
<i>Generating Synthetic Free-text Medical Records with Low Re-identification Risk using Masked Language Modeling</i>	
Samuel Belkadi, Libo Ren, Nicolo Micheletti, Lifeng Han and Goran Nenadic .....	200
<i>How many words does it take to understand a low-resource language?</i>	
Emily Chang and Nada Basit .....	207
<i>Linear Relational Decoding of Morphology in Language Models</i>	
Eric Xia and Jugul Kalita .....	225
<i>SPY: Enhancing Privacy with Synthetic PII Detection Dataset</i>	
Maksim Savkin, Timur Ionov and Vasily Kononov .....	236
<i>Tighter Clusters, Safer Code? Improving Vulnerability Detection with Enhanced Contrastive Loss</i>	
Pranav Kapparad and Biju R Mohan .....	247
<i>Text Extraction and Script Completion in Images of Arabic Script-Based Calligraphy: A Thesis Proposal</i>	
Dilara Zeynep Güler, Ümit Atlamaz and Şaziye Betül Özateş .....	253
<i>Subasa - Adapting Language Models for Low-resourced Offensive Language Detection in Sinhala</i>	
Shanilka Haturusinghe, Tharindu Cyril Weerasooriya, Christopher M Homan, Marcos Zampieri and Sidath Ravindra Liyanage .....	260
<i>Integrating Symbolic Execution into the Fine-Tuning of Code-Generating LLMs</i>	
Marina Sakharova, Abhinav Anand and Mira Mezini .....	271
<i>Through the Looking Glass: Common Sense Consistency Evaluation of Weird Images</i>	
Elisei Rykov, Kseniia Petrushina, Kseniia Titova, Anton Razzhigaev, Alexander Panchenko and Vasily Kononov .....	279
<i>ColorFoil: Investigating Color Blindness in Large Vision and Language Models</i>	
Ahnaf Mozib Samin, M Firoz Ahmed and Md. Mushtaq Shahriyar Rafee .....	294
<i>Towards Practical and Knowledgeable LLMs for a Multilingual World: A Thesis Proposal</i>	
Bryan Li .....	301
<i>MDC3: A Novel Multimodal Dataset for Commercial Content Classification in Bengali</i>	
Anik Mahmud Shanto, Mst. Sanjida Jamal Priya, Fahim Shakil Tamim and Mohammed Moshikul Hoque .....	311
<i>DateLogicQA: Benchmarking Temporal Biases in Large Language Models</i>	
Gagan Bhatia, Ming Ze Tang, Cristina Mahanta and Madiha Kazi .....	321
<i>AMR-RE: Abstract Meaning Representations for Retrieval-Based In-Context Learning in Relation Extraction</i>	
Peitao Han, Lis Pereira, Fei Cheng, Wan Jou She and Eiji Aramaki .....	333

<i>Linguistic Analysis of Veteran Job Interviews to Assess Effectiveness in Translating Military Expertise to the Civilian Workforce</i>	
Caroline J. Wendt, Ehsanul Haque Nirjhar and Theodora Chaspari . . . . .	343
<i>MetaMeme: A Dataset for Meme Template and Meta-Category Classification</i>	
Benjamin Lambright, Jordan Youner and Constantine Lignos . . . . .	356
<i>Representing and Clustering Errors in Offensive Language Detection</i>	
Jood Otey, Laura Biester and Steven R Wilson . . . . .	368
<i>ELIOT: Zero-Shot Video-Text Retrieval through Relevance-Boosted Captioning and Structural Information Extraction</i>	
Xuye Liu, Yimu Wang and Jian Zhao . . . . .	381
<i>Can Large Language Models Advance Crosswalks? The Case of Danish Occupation Codes</i>	
Bolei Ma, Cynthia A. Huang and Anna-Carolina Haensch . . . . .	392
<i>Paraphrase-based Contrastive Learning for Sentence Pair Modeling</i>	
Seiji Sugiyama, Risa Kondo, Tomoyuki Kajiwara and Takashi Ninomiya . . . . .	400
<i>Do Video Language Models really understand the video contexts?</i>	
Jeongwan Shin, Jinhyeong Lim and Hyeyoung Park . . . . .	408
<i>Evaluating Text Style Transfer Evaluation: Are There Any Reliable Metrics?</i>	
Sourabrata Mukherjee, Atul Kr. Ojha, John Philip McCrae and Ondrej Dusek . . . . .	418
<i>(CPER) From Guessing to Asking: An Approach to Resolving Persona Knowledge Gap in LLMs during Multi-Turn Conversations</i>	
Sarvesh Baskar, Manas Gaur, Srinivasan Parthasarathy and Tanmay Tulsidas Verlekar . . . . .	435
<i>Streamlining LLMs: Adaptive Knowledge Distillation for Tailored Language Models</i>	
Prajvi Saxena, Sabine Janzen and Wolfgang Maass . . . . .	448
<i>LLM DEBATE OPPONENT: Counter-argument Generation focusing on Implicit and Critical Premises</i>	
Taisei Ozaki, Chihiro Nakagawa, Naoya Inoue, Shoichi Naito and Kenshi Yamaguchi . . . . .	456
<i>AutoML Meets Hugging Face: Domain-Aware Pretrained Model Selection for Text Classification</i>	
Parisa Safikhani and David Broneske . . . . .	466
<i>Paraphrasing Attack Resilience of Various Machine-Generated Text Detection Methods</i>	
Andrii Shportko and Inessa Verbitsky . . . . .	474
<i>Detecting, Generating, and Evaluating in the Writing Style of Different Authors</i>	
Mosab Rezaei . . . . .	485
<i>Collaborative Data Exploration through Visualization: A Thesis Proposal Analyzing Impact of Conversational Assistants</i>	
Abari Bhattacharya and Barbara Di Eugenio . . . . .	492
<i>MENDER: Multi-hop Commonsense and Domain-specific CoT Reasoning for Knowledge-grounded Empathetic Counseling of Crime Victims</i>	
Abid Hossain, Priyanshu Priya, Armita Mani Tripathi, Pradeepika Verma and Asif Ekbal . . . . .	501
<i>SkipCLM: Enhancing Crosslingual Alignment of Decoder Transformer Models via Contrastive Learning and Skip Connection</i>	
Nikita Sushko, Alexander Panchenko and Elena Tutubalina . . . . .	517

<i>Towards LLMs Robustness to Changes in Prompt Format Styles</i>	
Lilian Ngweta, Kiran Kate, Jason Tsay and Yara Rizk . . . . .	529
<i>Reliability of Distribution Predictions by LLMs: Insights from Counterintuitive Pseudo-Distributions</i>	
Toma Suzuki, Ayuki Katayama, Seiji Gobara, Ryo Tsujimoto, Hibiki Nakatani, Kazuki Hayashi, Yusuke Sakai, Hidetaka Kamigaito and Taro Watanabe . . . . .	538
<i>Rosetta-PL: Propositional Logic as a Benchmark for Large Language Model Reasoning</i>	
Shaun Lee Baek, Shaun Esua-Mensah, Cyrus Tsui, Sejan Vigneswaralingam, Abdullah Alali, Michael Lu, Vasu Sharma and Kevin Zhu . . . . .	551