

IJCNLP-AACL 2025

**The 14th International Joint Conference on Natural  
Language Processing and The 4th Conference of the  
Asia-Pacific Chapter of the Association for Computational  
Linguistics**

**Proceedings of the Student Research Workshop**

December 20-24, 2025

©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 979-8-89176-304-3

## Message from the Chair of Student Research Workshop

Welcome to the IJCNLPACL 2025 Student Research Workshop (SRW)!

The IJCNLP-AACL 2025 SRW is held in conjunction with the 14th International Joint Conference on Natural Language Processing (IJCNLP) and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL).

Continuing its long-standing mission, the SRW provides a dedicated forum for student researchers in computational linguistics and natural language processing. It offers a supportive environment for students to share ideas, gain visibility, and receive constructive feedback from experienced members of the community.

As in previous years, the workshop invites submissions in two categories: research papers and thesis proposals. Authors may choose between archival submissions included in the conference proceedings and non-archival submissions, which allow presentation without limiting future publication opportunities. This flexible format accommodates both well-developed work and preliminary ideas, enabling meaningful participation from students at various stages of their research. Importantly, all submissions, whether archival or non-archival, receive equal care in review and mentorship.

This year, the SRW received 75 submissions in total: 71 through direct submission and 4 via ARR Commitment. We accepted 32 papers, resulting in an overall acceptance rate of 43%. The selection process was highly competitive, and we are pleased to note that all accepted papers demonstrate creativity and contribute to their respective fields. The accepted submissions reflect diversity not only in topics but also in the backgrounds of the student authors. Among these, 5 are archival and 27 are non-archival. During the conference, 9 papers will be presented as oral talks and 23 as poster presentations, delivered either in person or virtually.

Mentorship remains at the heart of the SRW. This year, 12 submissions participated in the pre-submission mentoring program, where students received early feedback on their writing and presentation from two experienced mentors. We are grateful to all mentors who supported authors throughout the review and revision process. Our gratitude goes to the program committee members for their thoughtful and careful reviews, and to the mentors who contributed their time or offering valuable feedback to student authors throughout the process.

We also thank our faculty advisors - Xiting Wang, Daisuke Kawahara, for their consistent guidance and support. We sincerely appreciate all of the organizers of the IJCNLP-AACL conference for their effort. And of course, we thank all authors for their enthusiasm and engagement. Your contributions make the SRW a vibrant and intellectually stimulating part of IJCNLPACL 2025.

We hope you find this year workshop inspiring and enriching.

# **Organizing Committee**

## **Faculty Advisors**

Xiting Wang, Renmin University, China  
Daisuke Kawahara, Waseda University, Japan

## **Student Chairs**

Santosh T.Y.S.S, Technical University of Munich,  
Shuichiro Shimizu, Kyoto University, Japan  
Yifan Gong, Renmin University, China

## Program Committee

### Program Chairs

Yifan Gong  
Shuichiro Shimizu  
Santosh T.y.s.s, Amazon

### Reviewers

Mohamed Abdalla, Gavin Abercrombie, Somak Aditya, Ameeta Agrawal, Georgios Alexandridis

Premjith B, Long Bai, Valerio Basile, Nathaniel Blanchard, Maria Boritchev, Davide Buscaldi, Jan Buys

Bo Chen, Alvin Cheung

Dipankar Das, Brian Davis, Lucia Donatelli, Ondrej Dusek

Minghong Fang, Alejandro Figueroa

Yanjun Gao, Venkata S Govindarajan, Camille Guinaudeau

Shohei Higashiyama

Abhik Jana, Arkadiusz Janz, Tianyu Jiang, Zhuoxuan Jiang

Sabyasachi Kamila, SeongKu Kang, Alina Karakanta, Bugeun Kim, Junyeong Kim, Taehwan Kim, Satoshi Kosugi, Nikhil Krishnaswamy, Marek Kubis, Sebastian Kula, Florian Kunneman, Yen-Ling Kuo

Yuxuan Lai, Andre Lamurias, David Langlois, Sahinur Rahman Laskar, Dongha Lee, Gaël Lejeune, Chuanyi Li, Jing Li, Sheng Li, KyungTae Lim, Dugang Liu, Hui Liu, Kunpeng Liu, Peipei Liu, Yidi Liu, Sharid Loáiciga, Jiaying Lu, Chunchuan Lyu

Lorenzo Malandri, Valentin Malykh, Edison Marrese-Taylor, Sandeep Mathias, Chandresh Kumar Maurya, Fanchao Meng, Dr. Satanik Mitra, Mainack Mondal, Raha Moraffah

Shah Nawaz, Hamada Nayel, Pengyu Nie

Jasabanta Patro, Adam Poliak

Mengyang Qiu

Leonardo Ranaldi, Christophe Rodrigues, Ramon Ruiz-Dolz

Yusuke Sakai, Debarshi Kumar Sanyal, Sunil Saumya, Alexandra Schofield, Sofia Serrano, Raksha Sharma, Zhou Sijia, Satyaki Sikdar, Mayank Singh, Nikhil Singh, Mohit Singhal, Konstantinos Skianis, Rui Sousa-Silva, Kai Sun

Santosh T.y.s.s, Yu Tian, Antonela Tommasel

Natalia Vanetik, Dan Vilenchik

Jindong Wang, Jingwen Wang, Ruibo Wang, Dittaya Wanvarie, Likang Wu, Winston Wu

Qiongkai Xu, Yongxiu Xu, Wei Xue

Jinyoung Yeo

Lei Zhang, Qing Zhang, Yang Zhang, Peide Zhu, Shaolin Zhu

## Table of Contents

<i>Interpretable Sparse Features for Probing Self-Supervised Speech Models</i>	
Iñigo Parra . . . . .	1
<i>Learning Dynamics of Meta-Learning in Small Model Pretraining</i>	
David Demitri Africa, Yuval Weiss, Paula Buttery and Richard Diehl Martinez . . . . .	10
<i>Efficient Environmental Claim Detection with Hyperbolic Graph Neural Networks</i>	
Darpan Aswal and Manjira Sinha . . . . .	24
<i>Stacked LoRA: Isolated Low-Rank Adaptation for Lifelong Knowledge Management</i>	
Heramb Vivek Patil, Vaishnavee Sanam and Minakshi Pradeep Atre . . . . .	36
<i>On Multilingual Encoder Language Model Compression for Low-Resource Languages</i>	
Daniil Gurgurov, Michal Gregor, Josef Van Genabith and Simon Ostermann . . . . .	47
<i>Do We Need Large VLMs for Spotting Soccer Actions?</i>	
Ritabrata Chakraborty, Rajat Subhra Chakraborty, Avijit Dasgupta and Sandeep Chaurasia . . . . .	59
<i>LRMGS: A Language-Robust Metric for Evaluating Question Answering in Very Low-Resource Indic Languages</i>	
Anuj Kumar, Satyadev Ahlawat, Yamuna Prasad and Virendra Singh . . . . .	66
<i>NumPert: Numerical Perturbations to Probe Language Models for Veracity Prediction</i>	
Peter Røysland Aarnes and Vinay Setty . . . . .	78
<i>Testing Simulation Theory in LLMs' Theory of Mind</i>	
Koshiro Aoki and Daisuke Kawahara . . . . .	96
<i>Turn-by-Turn Behavior Monitoring in LM-Guided Psychotherapy</i>	
Anish Sai Chedalla, Samina Ali, Jiuming Chen, starborn0128@gmail.com starborn0128@gmail.com and Eric Xia . . . . .	105
<i>BookAsSumQA: An Evaluation Framework for Aspect-Based Book Summarization via Question Answering</i>	
Ryuhei Miyazato, Ting-Ruen Wei, Xuyang Wu, Hsin-Tai Wu and Kei Harada . . . . .	123
<i>Thesis Proposal: Interpretable Reasoning Enhancement in Large Language Models through Puzzle and Ontological Task Analysis</i>	
Mihir Panchal . . . . .	134
<i>Adaptive Coopetition: Leveraging Coarse Verifier Signals for Resilient Multi-Agent LLM Reasoning</i>	
Wendy Yaqiao Liu, Rui Jerry Huang, Anastasia Miin and Lei Ding . . . . .	145
<i>AI Through the Human Lens: Investigating Cognitive Theories in Machine Psychology</i>	
Akash Kundu and Rishika Goswami . . . . .	156
<i>Thesis Proposal: A NeuroSymbolic Approach to Control Task-Oriented Dialog Systems</i>	
Anuja Tayal and Barbara Di Eugenio . . . . .	171
<i>Enriching the Low-Resource Neural Machine Translation with Large Language Model</i>	
Sachin Giri, Takashi Ninomiya and Isao Goto . . . . .	184
<i>Investigating Training and Generalization in Faithful Self-Explanations of Large Language Models</i>	
Tomoki Doi, Masaru Isonuma and Hitomi Yanaka . . . . .	193

<i>Thesis Proposal: Efficient Methods for Natural Language Generation/Understanding Systems</i>	
Nalin Kumar .....	209
<i>Two Step Automatic Post Editing of Patent Machine Translation based on Pre-trained Encoder Models and LLMs</i>	
Kosei Buma, Takehito Utsuro and Masaaki Nagata .....	218
<i>Rethinking Tokenization for Rich Morphology: The Dominance of Unigram over BPE and Morphological Alignment</i>	
Saketh Reddy Vemula, Sandipan Dandapat, Dipti Sharma and Parameswari Krishnamurthy .	232
<i>Are LLMs Good for Semantic Role Labeling via Question Answering?: A Preliminary Analysis</i>	
Ritwik Raghav and Abhik Jana .....	253
<i>Could you BE more sarcastic? A Cognitive Approach to Bidirectional Sarcasm Understanding in Language Models</i>	
Veer Chheda, Avantika Sankhe and Atharva Vinay Sankhe .....	259
<i>To What Extent Can In-Context Learning Solve Unseen Tasks?</i>	
Ryoma Shinto, Masashi Takeshita, Rafal Rzepka and Toshihiko Itoh .....	277
<i>Visualizing and Benchmarking LLM Factual Hallucination Tendencies via Internal State Analysis and Clustering</i>	
Nathan Mao, Varun Kaushik, Shreya Shivkumar, Parham Sharafolsami, Kevin Zhu and Sunishchal Dev .....	289
<i>Mitigating Forgetting in Continual Learning with Selective Gradient Projection</i>	
Anika Singh, David Martinez, Aayush Dhaulakhandi, Varun Chopade, Likhith Malipati, Vasu Sharma, Kevin Zhu, Sunishchal Dev and Ryan Lagasse .....	299
<i>VariantBench: A Framework for Evaluating LLMs on Justifications for Genetic Variant Interpretation</i>	
Humair Basharat, Simon Plotkin, Charlotte Le, Kevin Zhu, Michael Pink and Isabella Alfaro	314
<i>The ‘aftermath’ of compounds: Investigating Compounds and their Semantic Representations</i>	
Swarang Joshi .....	322