Introduction

Welcome to the EACL 2023 Student Research Workshop!

The EACL 2023 Student Research Workshop (SRW) is a forum for students in the field of Computational Linguistics and Natural Language Processing to come together to discuss and advance their research with help from more experienced researchers from both academia and industry.

Following the tradition of the previous student research workshops, we have two tracks: research papers and thesis proposals. The research paper track is a venue for PhD students, Master’s students, and advanced undergraduate students to describe completed work or work-in-progress along with preliminary results. The thesis proposal track is offered for students who have decided on a thesis topic and are interested in receiving feedback for their proposal with suggestions for both making the ideas achievable, as well as discussions related to future directions for their work.

The student research workshop has received considerable attention, and papers have addressed research questions in various areas and in several different languages. After excluding the withdrawn submissions, we received 40 submissions in total: 7 thesis proposals and 33 research papers. We accepted 6 thesis proposals and 13 research papers, resulting in an overall acceptance rate of ≈ 48%. Excluding non-archival papers, 16 papers appear in these proceedings. All the accepted papers will be presented as part of the EACL conference.

Mentoring is a core part of the SRW. In keeping with previous years, we organized pre-submission mentoring to improve the writing style and presentation of submissions. A total of 8 papers participated in this program. It offered students the opportunity to receive guidance from an experienced researcher before their submission was peer-reviewed for acceptance. We also offered post-submission mentorship for improving the quality of the poster presentation of accepted papers. A total of 13 submissions participated in this program.

We thank our program committee members for providing careful and comprehensive reviews for the papers, and all of our mentors for donating their time to provide feedback to our student authors. Thanks to our faculty advisors, Valerio Basile and Natalie Schluter, for the essential advice and suggestions, and to the EACL 2023 organizing committee for their support in the entire process. Finally, we would like to thank all the authors whose participation has made the workshop a success!
Program Committee

Organizers

Elisa Bassignana, IT University of Copenhagen
Matthias Lindemann, University of Edinburgh
Alban Petit, Université Paris-Saclay

Faculty advisers

Valerio Basile, University of Turin
Natalie Schluter, IT University of Copenhagen, Apple Inc.

Program Committee

Gavin Abercrombie, Heriot Watt University
Manex Agirrezaabal, University of Copenhagen
Ilseyar Alimova, Kazan Federal University
Miguel A. Alonso, Universidade da Coruña
Mithun Balakrishna, Amazon.com Inc
Maria Barrett, IT University of Copenhagen
Pierpaolo Basile, Department of Computer Science, University of Bari Aldo Moro
Chris Biemann, Universität Hamburg
Russsa Biswas, FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, AIFB Institute, KIT
Gerlof Bouma, University of Gothenburg
Johan Boye, KTH
Davide Buscaldi, LIPN, Université Sorbonne Paris Nord
Ruken Cakici, METU, Ankara
Burcu Can, University of Stirling
Giovanni Cassani, Tilburg University
Li-hsin Chang, University of Turku
Lin Chen, Engineering Manager, Meta Platform Inc.
Alexandra Chronopoulou, LMU Munich
Juan Manuel Coria, Université Paris-Saclay CNRS, LISM
Ruixiang Cui, University of Copenhagen
Verna Dankers, University of Edinburgh
Dana Dannells, Språkbanken Text, Dept. of Swedish, University of Gothenburg
Orphee De Clercq, LT3, Ghent University
Iria De-dios-flores, Universidade de Santiago de Compostela. CiTIUS center on Intelligent Technologies.
Roberto Dessi, Facebook AI Research / Universitat Pompeu Fabra
Mattia Di Gangi, AppTek GmbH
Elisa Di Nuovo, European Parliament
Antoine Doucet, University of La Rochelle
Mahmoud El-haj, Lancaster University
Desmond Elliott, University of Copenhagen
Dayne Freitag, SRI International
Saadia Gabriel, University of Washington
Björn Gambäck, Norwegian University of Science and Technology
Marcos Garcia, Universidade de Santiago de Compostela
Aina Gari Soler, LTCI, Télécom-Paris, Institut Polytechnique de Paris
Filip Ginter, University of Turku
Jonas Groschwitz, Saarland University
Katharina Haemmerl, Center for Information and Language Processing, LMU
Jiatong Han, National University of Singapore
Helmut Horacek, German Research Center for AI (DFKI)
Md.murad Hossain, PhD student
Dirk Hovy, Bocconi University
Filip Ilievski, USC Information Sciences Institute
Ashwin Ittoo, ULiege
Tommi Jauhiainen, University of Helsinki
Richard Johansson, University of Gothenburg
Katharina Kann, University of Colorado Boulder
Alina Karakanta, Fondazione Bruno Kessler (FBK), University of Trento
Amr Keleg, The University of Edinburgh
Najoung Kim, Boston University
Florian Kunneman, Vrije Universiteit Amsterdam
Andrey Kutuzov, University of Oslo
Els Lefever, LT3, Ghent University
Jens Lemmens, University of Antwerp
Hongmin Li, California State University, East Bay
Sheng Liang, Center for Information and Language Processing, LMU Munich
Lucy Lin, Spotify
Kevin Lin, Microsoft
Xinyuan Lu, National University of Singapore
Enrique Manjavacas Arevalo, University of Leiden
Alessandro Mazzei, Università degli Studi di Torino
Louise McNally, Universitat Pompeu Fabra
Syed Mehtab Alam, CIRAD, TETIS
Paola Merlo, University of Geneva
Nada Mimouni, CNAM Paris
Gosse Minnema, University of Groningen
Aditya Mogadala, Amazon
Olof Mogren, RISE Research institutes of Sweden
Adib Mosharrof, University of Kentucky
Max Müller-eberstein, IT University of Copenhagen
Costanza Navarretta, University of Copenhagen
Massimo Nicosia, Google
Anna Nikulásdóttir, Grammatek ehf
Joakim Nivre, Research Institutes of Sweden
Stefan Olafsson, Reykjavík University
Patrizia Paggio, University of Copenhagen and University of Malta
Alexander Panchenko, Skolkovo Institute of Science and Technology
Ludovica Pannitto, CIfMeC - University of Trento
Lucia Passaro, University of Pisa
Stellan Petersson, University of Gothenburg
Marco Polignano, University of Pisa
Yanxia Qin, School of Computing, National University of Singapore
Taraka Rama, Walmart Global Tech
Abhinav Ramesh Kashyap, ASUS AICS
Paul Rayson, Lancaster University
Lina M. Rojas Barahona, Orange Innovation Research
Sepideh Sadeghi, Google Inc.
Manuela Sanguinetti, University of Cagliari, Department of Mathematics and Computer Science
David Sasu, IT University of Copenhagen
Yves Scherrer, University of Helsinki
Sina Sheikholeslami, KTH Royal Institute of Technology
Manish Shrivastava, International Institute of Information Technology Hyderabad
Vered Shwartz, University of British Columbia
Kairit Sirts, University of Tartu
Pia Sommerauer, Vrije Universiteit Amsterdam
Marco Antonio Stranisci, University of Turin
Sara Stymne, Uppsala University
Alane Suhr, AI2
Alberto Testoni, University of Trento
Arda Tezcan, LT3, Language and Translation Technology Team, Ghent University
Isabel Trancoso, INESC-ID / IST Univ. Lisbon
Enrica Troiano, Vrije Universiteit
Dennis Ulmer, IT University of Copenhagen
Sowmya Vajjala, National Research Council
Antal Van Den Bosch, Utrecht University
Rob Van Der Goot, IT University of Copenhagen
Bonnie Webber, University of Edinburgh
Xinnuo Xu, University of Edinburgh
Mike Zhang, IT University of Copenhagen
Shiyue Zhang, The University of North Carolina at Chapel Hill
Heike Zinsmeister, Universität Hamburg

Mentors

Gavin Abercrombie, Heriot Watt University
Maria Barrett, IT University of Copenhagen
Burcu Can, University of Stirling
Orphee De Clercq, LT3, Ghent University
Jonas Groschwitz, Saarland University
Dirk Hovy, Bocconi University
Florian Kunneman, Vrije Universiteit Amsterdam
Costanza Navarretta, University of Copenhagen
Yves Scherrer, University of Helsinki
Vered Shwartz, University of British Columbia
Pia Sommerauer, Vrije Universiteit Amsterdam
Antal Van Den Bosch, Utrecht University
Rob Van Der Goot, IT University of Copenhagen
Bonnie Webber, University of Edinburgh
Table of Contents

*Revealing Weaknesses of Vietnamese Language Models Through Unanswerable Questions in Machine Reading Comprehension*
  Son Quoc Tran, Phong Nguyen-Thuan Do, Kiet Van Nguyen and Ngan Luu-Thuy Nguyen . . . . . . . . 1

*Incorporating Dropped Pronouns into Coreference Resolution: The case for Turkish*
  Tuğba Pamay Arslan and Gűlşen Eryiğit ................................................................. 14

*Towards Generation and Recognition of Humorous Texts in Portuguese*
  Marcio Lima Inácio and Hugo Gonçalo Oliveira .................................................... 26

*GAP-Gen: Guided Automatic Python Code Generation*
  Junchen Zhao, Yurun Song, Junlin Wang and Ian Harris ........................................ 37

*Development of pre-trained language models for clinical NLP in Spanish*
  Claudio Aracena and Jocelyn Dunstan ........................................................................ 52

*Which One Are You Referring To? Multimodal Object Identification in Situated Dialogue*
  Holy Lovenia, Samuel Cahyawijaya and Pascale Fung ............................................. 61

*A Unified Framework for Emotion Identification and Generation in Dialogues*
  Avinash Madasu, Mauajama Firdaus and Asif Ekbal ................................................ 73

*Improving and Simplifying Template-Based Named Entity Recognition*
  Muñali Kondragunta, Olatz Perez-de-Viñaspre and Maite Oronoz ................................ 79

*Polite Chatbot: A Text Style Transfer Application*
  Sourabrata Mukherjee, Vojtěch Hudeček and Ondřej Dušek .................................... 87

*Template-guided Grammatical Error Feedback Comment Generation*
  Steven Coyne ............................................................................................................... 94

*Clinical Text Anonymization, its Influence on Downstream NLP Tasks and the Risk of Re-Identification*
  Iyadh Ben Cheikh Larbi, Aljoscha Burchardt and Roland Roller ................................ 105

*Automatic Dialog Flow Extraction and Guidance*
  Patrícia Ferreira ........................................................................................................ 112

*Diverse Content Selection for Educational Question Generation*
  Amir Hadifar, Semere Kiros Bitew, Johannes Deleu, Veronique Hoste, Chris Develder and Thomas Demeester .................................................................................. 123

*Towards Automatic Grammatical Error Type Classification for Turkish*
  Harun Uz and Gűlşen Eryiğit .................................................................................... 134

*Theoretical Conditions and Empirical Failure of Bracket Counting on Long Sequences with Linear Recurrent Networks*
  Nadine El-Naggar, Pranava Madhyastha and Tillman Weyde .................................... 143

*Addressing Domain Changes in Task-oriented Conversational Agents through Dialogue Adaptation*
  Tiziano Labruna and Bernardo Magnini ..................................................................... 149