BEA 2024

The 19th Workshop on Innovative Use of NLP for Building Educational Applications

Proceedings of the Workshop

June 20, 2024

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

Gold Level

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

ISBN 979-8-89176-100-1

Introduction

This year marks the 19th edition of the *Workshop on Innovative Use of NLP for Building Educational Applications*. As in previous years, we are happy to welcome a plethora of work on various aspects and types of educational applications – from some of the traditionally popular tasks around language learning including automated essay scoring, grammatical error detection and correction, readability assessment, and vocabulary acquisition, among others, to topics related to math and programming education, questions around empathy in teachers' responses and evaluation of teacher encouragement, fairness and explainability, bias alleviation, and ethics in AI models applied to the educational domain and many other exciting developments.

In total, we received 88 submissions, and from these, we have accepted 4 papers as talks and 34 as poster and demo presentations, for an overall acceptance rate of 43 percent. Like the rest of the NLP community, we are observing a paradigm shift, with more and more researchers applying Large Language Models (LLMs) in the context of educational applications for a variety of purposes including implementation and evaluation. A large number of papers that we have received and accepted this year investigate the topics around the integration of LLMs into educational applications. With research excellence being one of the main factors considered when making paper acceptance decisions, we hope we have also brought together a diverse program. As before, we also put a particular emphasis on multilingualism of the work included in our program, and this year BEA features work done not only on English, but also on other languages including Catalan, Danish, Dutch, Filipino, French, German, Japanese, Italian, Portuguese, Romanian, Russian, Sinhala, Spanish, and Swedish.

In addition to the diverse oral, poster, and demo presentations, this year, Alla Rozovskaya, an Assistant Professor in the Department of Computer Science at Queens College, City University of New York, will give a keynote on *Multilingual Low-Resource Natural Language Processing for Language Learning*. Furthermore, BEA 2024 has hosted two shared tasks – on *Automated Prediction of Item Difficulty and Item Response Time (APIDIRT)* and on *Multilingual Lexical Simplification Pipeline (MLSP)*. Both tasks have attracted a large number of participants, and the program includes oral presentations on the shared task descriptions from the organizers as well as extended poster sessions for shared task participants presenting their systems.

Last but not least, we would like to thank everyone who has been involved in organizing the BEA workshop this year. We are particularly grateful to our sponsors who keep supporting BEA: this year, our sponsors include British Council, Cambridge University Press & Assessment, CATALPA, Duolingo English Test, Educational Testing Service, and the National Board of Medical Examiners. We would like to also thank all the authors who showed interest and submitted a paper this year.

Due to the record number of submissions received, we had to extend our invitation to become part of the Program Committee to all the authors of submitted papers, and many have helped us and provided their valuable feedback and thoughtful reviews. Without this help from the community, it would not be possible to spread the reviewing load reasonably, and we are very grateful to our regular reviewers as well as to emergency reviewers and all the authors who joined our PC this year and who, we hope, may become our regular PC members. In particular, we would like to extend our gratitude to the following emergency and outstanding reviewers: Michael Gringo Angelo Bayona, Jeanette Bewersdorff, Jie Cao, Scott Crossley, Sam Davidson, Kordula De Kuthy, Jasper Degraeuwe, Rujun Gao, Handoko Handoko, Michael Holcomb, Helen Jin, John Sie Yuen Lee, Hunter McNichols, Arun Balajiee Lekshmi Narayanan, Huy Viet Nguyen, Adam Nohejl, Eda Okur, Udita Patel, Martí Quixal, Manav Rathod, Alla Rozovskaya, Abhijit Suresh, Chee Wei Tan, Gladys Tyen, Justin Vasselli, Elena Volodina, ManFai Wong, Kevin Yancey, Roman Yangarber, Torsten Zesch.

Ekaterina Kochmar, MBZUAI
Marie Bexte, FernUniversität in Hagen
Jill Burstein, Duolingo
Andrea Horbach, Hildesheim University and CATALPA, FernUniversität in Hagen
Ronja Laarmann-Quante, Ruhr University Bochum
Anaïs Tack, KU Leuven, imec
Victoria Yaneva, National Board of Medical Examiners
Zheng Yuan, King's College London

Organizing Committee

General Chair

Ekaterina Kochmar, MBZUAI

Program Chairs

Andrea Horbach, Universität Hildesheim and CATALPA, FernUniversität in Hagen
Ronja Laarmann-Quante, Ruhr University Bochum
Marie Bexte, FernUniversität in Hagen

Publication Chair

Anaïs Tack, KU Leuven, imec

Shared Tasks Chairs

Victoria Yaneva, National Board of Medical Examiners
Jill Burstein, Duolingo

Sponsorship Chair

Zheng Yuan, King's College London

Program Committee

Chairs

Ekaterina Kochmar, MBZUAI
Marie Bexte, FernUniversität in Hagen
Jill Burstein, Duolingo
Andrea Horbach, Universität Hildesheim
Ronja Laarmann-Quante, Ruhr University Bochum
Anaïs Tack, KU Leuven; imec; UCLouvain
Victoria Yaneva, National Board of Medical Examiners
Zheng Yuan, King's College London

Program Committee

Tazin Afrin, Educational Testing Service
Prabhat Agarwal, Pinterest
Erfan Al-Hossami, University of North Carolina at Charlotte
Desislava Aleksandrova, CBC/Radio-Canada
Giora Alexandron, Weizmann Institute of Science
David Alfter, UCLouvain
Fernando Alva-Manchego, Cardiff University
Jatin Ambasana, Unitedworld Institute of Technology, Karnavati University
Nico Andersen, DIPF | Leibniz Institute for Research and Information in Education
Alejandro Andrade, Pearson
Tesfa Tegegne Asfaw, Bahir Dar University
Nischal Ashok Kumar, University of Massachusetts Amherst
Berk Atil, Pennsylvania State University
Shiva Baghel, Extramarks
Rabin Banjade, University of Memphis
Stefano Banno, University of Cambridge
Michael Gringo Angelo Bayona, Trinity College Dublin
Lee Becker, Pearson
Beata Beigman Klebanov, Educational Testing Service
Lisa Beinborn, Vrije Universiteit Amsterdam
Enrico Benedetti, University of Bologna
Luca Benedetto, University of Cambridge
Jeanette Bewersdorff, FernUniversität in Hagen
Ummugul Bezirhan, Boston College, TIMSS and PIRLS International Study Center
Smita Bhattacharya, Saarland University
Abhidip Bhattacharyya, University of Massachusetts, Amherst
Serge Bibauw, UCLouvain
Robert-Mihai Botarleanu, National University of Science and Technology POLITEHNICA Bucharest
Allison Bradford, University of California, Berkeley
Ted Briscoe, MBZUAI
Jie Cao, University of Colorado
Dan Carpenter, North Carolina State University
Dumitru-Clementin Cercel, University Politehnica of Bucharest
Imran Chamieh, Hochschule Ruhr West

Jeevan Chapagain, UniversityofMemphis
Mei-Hua Chen, Department of Foreign Languages and Literature, Tunghai University
Luis Chiruzzo, Universidad de la Republica
Yan Cong, Purdue University
Mark Core, University of Southern California
Steven Coyne, Tohoku University / RIKEN
Scott Crossley, Georgia State University
Sam Davidson, University of California, Davis
Orphee De Clercq, LT3, Ghent University
Kordula De Kuthy, Universität Tübingen
Michiel De Vrindt, KU Leuven
Jasper Degraeuwe, Ghent University
Dorottya Demszky, Stanford University
Yang Deng, The Chinese University of Hong Kong
Aniket Deroy, IIT Kharagpur
Chris Develder, Ghent University
Yuning Ding, FernUniversität in Hagen
Rahul Divekar, Educational Testing Service
George Duenas, Universidad Pedagogica Nacional
Matthew Durward, University of Canterbury
Yo Ehara, Tokyo Gakugei University
Yao-Chung Fan, National Chung Hsing University
Effat Farhana, VanderbiltUniversity
Mariano Felice, University of Cambridge
Nigel Fernandez, University of Massachusetts Amherst
Michael Flor, Educational Testing Service
Jennifer-Carmen Frey, EURAC Research
Kotaro Funakoshi, Tokyo Institute of Technology
Thomas Gaillat, Rennes 2 university
Diana Galvan-Sosa, University of Cambridge
Ashwinkumar Ganesan, Amazon Alexa AI
Achyutarama Ganti, Oakland University
Rujun Gao, Texas A&M University
Ritik Garg, Extramarks Education Pvt. Ltd.
Dominik Glandorf, University of Tübingen, Yale University
Christian Gold, Fernuniversitaet Hagen
Sebastian Gombert, DIPF | Leibniz Institute for Research and Information in Education
Kiel Gonzales, University of the Philippines Diliman
Cyril Goutte, National Research Council Canada
Prasoon Goyal, The University of Texas at Austin
Pranav Gupta, Cornell University
Abigail Gurin Schleifer, Weizmann Institute of Science
Handoko Handoko, Universitas Andalas
Ching Nam Hang, Department of Computer Science, City University of Hong Kong
Jiangang Hao, Educational Testing Service
Ahatsham Hayat, University of Nebraska-Lincoln
Nicolas Hernandez, Nantes University
Nils Hjortnaes, Indiana University Bloomington
Michael Holcomb, University of Texas Southwestern Medical Center
Heiko Holz, Ludwigsburg University of Education
Sukhyun Hong, Hyperconnect, Matchgroup

Chung-Chi Huang, Frostburg State University
Chieh-Yang Huang, MetaMetrics Inc
Anna Huelsing, University of Hildesheim
Syed-Amad Hussain, Ohio State University
Catherine Ikae, Applied Machine Intelligence, Bern University of Applied Sciences, Switzerland
Joseph Marvin Imperial, University of Bath
Radu Tudor Ionescu, University of Bucharest
Suriya Prakash Jambunathan, New York University
Qinjin Jia, North Carolina State University
Helen Jin, University of Pennsylvania
Ioana Jivet, FernUniversität in Hagen
Léane Jourdan, Nantes University
Anisia Katinskaia, University of Helsinki
Elma Kerz, RWTH Aachen University
Fazel Keshtkar, St. John's University
Mamoru Komachi, Hitotsubashi University
Charles Koutcheme, Aalto University
Roland Kuhn, National Research Council of Canada
Alexander Kwako, University of California, Los Angeles
Kristopher Kyle, University of Oregon
Antonio Laverghetta Jr., Pennsylvania State University
Celine Lee, Cornell University
John Lee, City University of Hong Kong
Seolhwa Lee, Technical University of Darmstadt
Jaewook Lee, UMass Amherst
Arun Balajiee Lekshmi Narayanan, University of Pittsburgh
Yayun Li, City University of Hong Kong
Yudong Liu, Western Washington University
Zhexiong Liu, University of Pittsburgh
Naiming Liu, Rice University
Julian Lohmann, Christian Albrechts Universität Kiel
Anastassia Loukina, Grammarly Inc
Jiaying Lu, Emory University
Crisron Rudolf Lucas, UniversityCollegeDublin
Collin Lynch, NCSU
Sarah Löber, University of Tübingen
Jakub Macina, ETH Zurich
Nitin Madnani, Educational Testing Service
Jazzmin Maranan, University of the Philippines Diliman
Arianna Masciolini, University of Gothenburg
Sandeep Mathias, Presidency University
Hunter McNichols, University of Massachusetts Amherst
Jose Marie Mendoza, University of the Philippines Diliman
Amit Mishra, AmityUniversityMadhyaPradesh
Masato Mita, CyberAgent Inc.
Daniel Mora Melanchthon, Pontificia Universidad Católica de Valparaíso
Phoebe Mulcaire, Duolingo
Laura Musto, Universidad de la Republica
Ricardo Muñoz Sánchez, Språkbanken Text, Göteborgs Universitet
Farah Nadeem, LUMS
Sungjin Nam, ACT, Inc

Diane Napolitano, The Washington Post
Tanya Nazaretsky, EPFL
Kamel Nebhi, Education First
Seyed Parsa Neshaei, EPFL
Huy Nguyen, Amazon
Gebregziabihier Nigusie, Mizan-Tepi University
Christina Niklaus, University of St. Gallen
S Jaya Nirmala, National Institute of Technology Tiruchirappalli
Adam Nohejl, Nara Institute of Science and Technology
Kai North, George Mason University
Eda Okur, Intel Labs
Kostiantyn Omelianchuk, Grammarly
Amin Omidvar, PhD student at the Department of Electrical Engineering and Computer Science, York University
Benjamin Paddags, Department of Computer Science, University of Copenhagen
Ulrike Pado, HFT Stuttgart
Jeiyoon Park, Korea University
Chanjun Park, Upstage
Udita Patel, Amazon.com
Long Qin, Alibaba
Mengyang Qiu, University at Buffalo
Martí Quixal, University of Tübingen
Vatsal Raina, University of Cambridge
Manav Rathod, University of California, Berkeley
Hanumant Redkar, Goa University, Goa
Edsel Jedd Renovalles, University of the Philippines Diliman
Robert Reynolds, Brigham Young University
Saed Rezayi, National Board of Medical Examiners
Luisa Ribeiro-Flucht, University of Tuebingen
Frankie Robertson, University of Jyväskylä
Donya Rooein, Bocconi University
Aiala Rosá, Instituto de Computación, Facultad de Ingeniería, Universidad de la República
Allen Roush, University of Oregon
Alla Rozovskaya, Queens College, City University of New York
Josef Ruppenhofer, Fernuniviersität in Hagen
Horacio Saggion, Universitat Pompeu Fabra
Omer Salem, Cairo University
Nicy Scaria, Indian Institute of Science
Nils-Jonathan Schaller, Leibniz Institute for Science and Mathematics Education
Martha Shaka, University College Cork
Ashwath Shankarnarayan, New York University
Matthew Shardlow, Manchester Metropolitan University
Gyu-Ho Shin, University of Illinois Chicago
Li Siyan, Columbia University
Yixiao Song, University of Massachusetts Amherst
Mayank Soni, ADAPT Centre, Trinity College Dublin
Maja Stahl, Leibniz University Hannover
Felix Stahlberg, Google Research
Katherine Stasaski, Salesforce Research
Kevin Stowe, Educational Testing Services (ETS)
Helmer Strik, Centre for Language and Speech Technology (CLST), Centre for Language Studies

(CLS), Radboud University Nijmegen
David Strohmaier, University of Cambridge
Katsuhito Sudoh, Nara Women's University
Hakyung Sung, University of Oregon
Abhijit Suresh, Graduate Student
CheeWei Tan, NanyangTechnologicalUniversity
Zhongwei Teng, Vanderbilt University
Xiaoyi Tian, University of Florida
Gladys Tyen, University of Cambridge
Shriyash Upadhyay, University of Pennsylvania
Felipe Urrutia, Center for Advanced Research in Education
Masaki Uto, The University of Electro-Communications
Sowmya Vajjala, National Research Council
Justin Vasselli, Nara Institute of Science and Technology
Giulia Venturi, Institute of Computational Linguistics Antonio Zampolli"(ILC-CNR)
Anthony Verardi, Duolingo
Elena Volodina, University of Gothenburg
Jiani Wang, East China Normal University
Taro Watanabe, Nara Institute of Science and Technology
Michael White, The Ohio State University
Alistair Willis, The Open University
Anna Winklerova, Faculty of Informatics Masaryk University
Man Fai Wong, City University of Hong Kong
Simon Woodhead, Eedi
Changrong Xiao, Tsinghua University
Kevin P. Yancey, Duolingo
Roman Yangarber, University of Helsinki
Su-Youn Yoon, EduLab
Marcos Zampieri, George Mason University
Fabian Zehner, DIPF | Leibniz Institute for Research and Information in Education
Kamyar Zeinalipour, University of Siena
Torsten Zesch, Computational Linguistics, FernUniversität in Hagen
Jing Zhang, Emory University
Yang Zhong, University of Pittsburgh
Yiyun Zhou, NBME
Jessica Zipf, University of Konstanz
Michael Zock, CNRS-LIS
Bowei Zou, Institute for Infocomm Research

Keynote Talk

Multilingual Low-Resource Natural Language Processing for Language Learning

Alla Rozovskaya
Queens College, City University of New York

Abstract: Recent studies on a wide range of NLP tasks have demonstrated the effectiveness of training paradigms that integrate large language models. However, such methods require large amounts of labeled and unlabeled data, limiting their success to a small set of well-resourced languages. This talk will discuss low-resource approaches for two language learning applications. We will begin with work on generating vocabulary exercises. We will describe an approach that does not require labeled training data and can be used to adapt the exercises to the linguistic profile of the learner. Next, we will discuss our recent work on multilingual grammatical error correction (GEC), addressing the issue of training GEC models for languages with little labeled training data, and the issue of evaluating system performance when high-quality benchmarks are lacking.

Bio: Alla Rozovskaya is an Assistant Professor in the Department of Computer Science at Queens College, City University of New York (CUNY), and a member of the Doctoral Faculty of the Computer Science and Linguistics programs at the CUNY Graduate Center. She earned her Ph.D. in Computational Linguistics at the University of Illinois at Urbana-Champaign, under the supervision of Prof. Dan Roth. Her research interests lie broadly in the area of low-resource and multilingual NLP and educational applications.

Table of Contents