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

In cooperation with

[Logos of the sponsors]
Introduction

Equality, Diversity and Inclusion (EDI) is an important agenda across every field throughout the world. Language as a major part of communication should be inclusive and treat everyone with equality. Today’s large internet community uses language technology (LT) and has a direct impact on people across the globe. EDI is crucial to ensure everyone is valued and included, so it is necessary to build LT that serves this purpose. Recent results have shown that big data and deep learning are entrenching existing biases and that some algorithms are even naturally biased due to problems such as ‘regression to the mode’. Our focus is on creating LT that will be more inclusive of gender, racial, sexual orientation, persons with disability. The workshop will focus on creating speech and language technology to address EDI not only in English, but also in less resourced languages.
Organizing Committee

General Chair

Bharathi Raja Chakravarthi, National University of Ireland Galway
B Bharathi, SSN College of Engineering, Tamil Nadu, India
John P McCrae, National University of Ireland Galway
Manel Zarrouk, Institut Galilee Universite, Paris
Kalika Bali, Microsoft Research, India
Paul Buitelaar, National University of Ireland Galway
Program Committee

Adeep Hande, Indian Institute of Information Technology Tiruchirappalli
Akshat Gupta, Carnegie Mellon University
Alberto Barrón-Cedeño, Università di Bologna
Anna Pillar, Radboud University
Anne Lauscher, Bocconi University
António Câmara, Columbia University
Arianna Muti, University of Bologna
Blaž Škrlj, Jožef Stefan Institute
Boshko Koloski, International Postgraduate School Jožef Stefan
Courtney Mansfield, LivePerson Inc.
Debora Nozza, Bocconi University
Ewoenam Kwaku, University of Antwerp
Fazlourrahman Balouchzahi, Instituto Politécnico Nacional
Filip Nilsson, Department of Computer Science, Electrical and Space Engineering
Frank Rudzicz, University of Toronto
Harrison Santiago, University of Florida
Ishan Sanjeev, International Institute of Information Technology Hyderabad
Iñaki San, Elhuyar Fundazioa
Jamell Dacon, Michigan State University
Jetske Adams, Radboud University
José Antonio, Universidad de Murcia
KV Aditya, International Institute of Information Technology, Hyderabad
Kangwook Lee, University of Wisconsin, Madison
Katerina Korre, University of Bologna
Kyle Swanson, Stanford University
Kyrill Poelmans, Tilburg University
Manex Agirrezabal, University of Copenhagen
Marion Bartl, University College Dublin
Martha Larson, Radboud University
Michael Gira, University of Wisconsin - Madison
Nankai Lin, Guangdong University of Foreign Studies
Nawshad Farruque, University of Alberta
Nina Markl, Institute for Language, Cognition and Computation, University of Edinburgh
Olga Zamaraeva, Universidad de La Coruña
Oren Mishali, Computer Science Department, Technion-Israel Institute of Technology
Pieter Delobelle, KU Leuven
Poonam Goyal, BITS Pilani, Birla Institute of Technology and Science
Pradeep Kumar, NIT Patna
Rafael Valencia-García, Universidad de Murcia
Rajalakshmi Sivanaiah, Sri Sivasubramaniya Nadar College of Engineering
Sunil Saumya, IIIT Dharwad
Senthil Kumar, Anna University
Sharal Coelho, Mangalore University
Suhasini S, SSN College of Engineering
Suman Dowlagar, International Institute of Information Technology Hyderabad
Keynote Talk: Towards Equitable Language Technologies
Su Lin
Microsoft Research, Montreal

Abstract: Language technologies are now ubiquitous. Yet the benefits of these technologies do not accrue evenly to all people, and they can be harmful; they can reproduce stereotypes, prevent speakers of “non-standard” language varieties from participating fully in public discourse, and reinscribe historical patterns of linguistic discrimination. In this talk, I will take a tour through the rapidly emerging body of research examining bias and harm in language technologies. I will offer some perspective on the many challenges of this work, ranging from how we conceptualize and measure language-related harms to how we grapple with the complexities of where and how language technologies are encountered. I will conclude by discussing some future directions towards more equitable technologies.

Bio: She is a postdoctoral researcher in the Fairness, Accountability, Transparency, and Ethics (FATE) group at Microsoft Research Montréal. She is interested in examining the social and ethical implications of natural language processing technologies; She develop approaches for anticipating, measuring, and mitigating harms arising from language technologies, focusing on the complexities of language and language technologies in their social contexts, and on supporting NLP practitioners in their ethical work. She has also worked on using NLP approaches to examine language variation and change (computational sociolinguistics), for example developing models to identify language variation on social media.
# Table of Contents

*Mind the data gap(s): Investigating power in speech and language datasets*
  Nina Markl ................................................................. 1

*Regex in a Time of Deep Learning: The Role of an Old Technology in Age Discrimination Detection in Job Advertisements*
  Anna Pillar, Kyrill Poelmans and Martha Larson ............................. 13

*Doing not Being: Concrete Language as a Bridge from Language Technology to Ethnically Inclusive Job Ads*
  Jetske Adams, Kyrill Poelmans, Iris Hendrickx and Martha Larson ............. 19

*Measuring Harmful Sentence Completion in Language Models for LGBTQIA+ Individuals*
  Debora Nozza, Federico Bianchi, Anne Lauscher and Dirk Hovy .................. 26

*Using BERT Embeddings to Model Word Importance in Conversational Transcripts for Deaf and Hard of Hearing Users*
  Akhter Al Amin, Saad Hassan, Cecilia Alm and Matt Huenerfauth ............... 35

*Detoxifying Language Models with a Toxic Corpus*
  Yoona Park and Frank Rudzicz ............................................. 41

*Inferring Gender: A Scalable Methodology for Gender Detection with Online Lexical Databases*
  Marion Bartl and Susan Leavy .............................................. 47

*Debiasing Pre-Trained Language Models via Efficient Fine-Tuning*
  Michael Gira, Ruisu Zhang and Kangwook Lee ................................ 59

*Disambiguation of morpho-syntactic features of African American English – the case of habitual be*
  Harrison Santiago, Joshua Martin, Sarah Moeller and Kevin Tang ............ 70

*Behind the Mask: Demographic bias in name detection for PII masking*
  Courtney Mansfield, Amandalynne Paullada and Kristen Howell ............... 76

*Mapping the Multilingual Margins: Intersectional Biases of Sentiment Analysis Systems in English, Spanish, and Arabic*
  António Câmara, Nina Taneja, Tamjeed Azad, Emily Allaway and Richard Zemel 90

*Monte Carlo Tree Search for Interpreting Stress in Natural Language*
  Kyle Swanson, Joy Hsu and Mirac Suzgun ................................... 107

*IIITSurat@LT-EDI-ACL2022: Hope Speech Detection using Machine Learning*
  Pradeep Kumar Roy, Snehaan Bhawal, Abhinav Kumar and Bharathi Raja Chakravarthi 120

*The Best of both Worlds: Dual Channel Language modeling for Hope Speech Detection in low-resourced Kannada*
  Adeep Hande, Siddhanth U Hegde, Sangeetha S, Ruba Priyadharshini and Bharathi Raja Chakravarthi ........................................... 127

*NYCU_TWD@LT-EDI-ACL2022: Ensemble Models with VADER and Contrastive Learning for Detecting Signs of Depression from Social Media*
  Wei-Yao Wang, Yu-Chien Tang, Wei-Wei Du and Wen-Chih Peng ................. 136

*UMUTeam@LT-EDI-ACL2022: Detecting homophobic and transphobic comments in Tamil*
  José Antonio García-Díaz, Camilo Caparros-Laíz and Rafael Valencia-García ... 140
<table>
<thead>
<tr>
<th>Paper ID</th>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMUTeam@LT-EDI-ACL2022</td>
<td>Detecting Signs of Depression from text</td>
<td>José Antonio García-Díaz and Rafael Valencia-García</td>
<td>145</td>
</tr>
<tr>
<td>bitssa_nlp@LT-EDI-ACL2022</td>
<td>Leveraging Pretrained Language Models for Detecting Homophobia and Transphobia in Social Media Comments</td>
<td>Vitthal Bhandari and Poonam Goyal</td>
<td>149</td>
</tr>
<tr>
<td>ABLIMET @ LT-EDI-ACL2022</td>
<td>A Roberta based Approach for Homophobia/Transphobia Detection in Social Media</td>
<td>Abulimiti Maimaitituoheti</td>
<td>155</td>
</tr>
<tr>
<td>MUCIC@LT-EDI-ACL2022</td>
<td>Hope Speech Detection using Data Re-Sampling and 1D Conv-LSTM</td>
<td>Anusha M D Gowda, Fazlourrahman Balouchzahi, Hosahalli Lakshmaiah Shashirekha and Grigori Sidorov</td>
<td>161</td>
</tr>
<tr>
<td>DeepBlues@LT-EDI-ACL2022</td>
<td>Depression level detection modelling through domain specific BERT and short text Depression classifiers</td>
<td>Nawshad Farruque, Osmar Zaiane, Randy Goebel and Sudhakar Sivapalan</td>
<td>167</td>
</tr>
<tr>
<td>SUH_ASR@LT-EDI-ACL2022</td>
<td>Transformer based Approach for Speech Recognition for Vulnerable Individuals in Tamil</td>
<td>Suhasini S and Bharathi B</td>
<td>177</td>
</tr>
<tr>
<td>LPS@LT-EDI-ACL2022</td>
<td>An Ensemble Approach about Hope Speech Detection</td>
<td>Yue Ying Zhu</td>
<td>183</td>
</tr>
<tr>
<td>CURAI_IITDWD@LT-EDI-ACL2022</td>
<td>Hope Speech Detection in English YouTube Comments using Deep Learning Techniques</td>
<td>Vanshita Jha, Ankit Kumar Mishra and Sunil Saumya</td>
<td>190</td>
</tr>
<tr>
<td>SSN_MLRG3@LT-EDI-ACL2022</td>
<td>Depression Detection System from Social Media Text using Transformer Models</td>
<td>Sarika Esackimuthu, Shruthi Hariprasad, Rajalakshmi Sivaniaiah, Angel Deborah S, Sakaya Milton Rajendram and Mirnalinee T T</td>
<td>196</td>
</tr>
<tr>
<td>BERT 4EVER@LT-EDI-ACL2022</td>
<td>Detecting signs of Depression from Social Media Detecting Depression in Social Media using Prompt-Learning and Word-Emotion Cluster</td>
<td>Xiaotian Lin, Yingwen Fu, Ziyu Yang, Nankai Lin and Shengyi Jiang</td>
<td>200</td>
</tr>
<tr>
<td>CIC@LT-EDI-ACL2022</td>
<td>Are transformers the only hope? Hope speech detection for Spanish and English comments</td>
<td>Fazlourrahman Balouchzahi, Sabur Butt, Grigori Sidorov and Alexander Gelbukh</td>
<td>206</td>
</tr>
<tr>
<td>scubeMSEC@LT-EDI-ACL2022</td>
<td>Detection of Depression using Transformer Models</td>
<td>Sivamanikandan S, Santhosh V, Sanjaykumar N, Jerin Mahibha C and Thenmozhi Durairaj</td>
<td>212</td>
</tr>
<tr>
<td>SSNCSE_NLP@LT-EDI-ACL2022</td>
<td>Hope Speech Detection for Equality, Diversity and Inclusion using sentence transformers</td>
<td>Bharathi B, Dhanya Srinivasan, Josephine Varsha, Thenmozhi Durairaj and Senthil Kumar B</td>
<td>218</td>
</tr>
</tbody>
</table>