LTEDI 2022

The Second Workshop on Language Technology for Equality, Diversity and Inclusion

Proceedings of the Workshop

May 27, 2022

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

In cooperation with

©2022 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-955917-43-8

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

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
Luke Melas-Kyriazi, University of Oxford, University of Oxford
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 Departmen, 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

Susan Leavy, University College Dublin
Thenmozhi Durairaj, Sri Sivasubramaniya Nadar College Of Engineering
Toon Calders, Universiteit Antwerpen
Usman Naseem, University of Sydney
Vishesh Gupta, IIT(ISM) Dhanbad
Vitthal Bhandari, BITS Pilani, Birla Institute of Technology and Science
Wei-Wei Du, National Yang Ming Chiao Tung University
Wei-Yao Wang, National Yang Ming Chiao Tung University
Yingwen Fu, Guangdong University of Foreign Studies
Zining Zhu, University of Toronto
Shaina Raza, University of Toronto

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-Laiz and Rafael Valencia-García 140

UMUTeam@LT-EDI-ACL2022: Detecting Signs of Depression from text
José Antonio García-Díaz and Rafael Valencia-García 145

bitsa_nlp@LT-EDI-ACL2022: Leveraging Pretrained Language Models for Detecting Homophobia and Transphobia in Social Media Comments
Vitthal Bhandari and Poonam Goyal 149

ABLIMET @LT-EDI-ACL2022: A Roberta based Approach for Homophobia/Transphobia Detection in Social Media
Abulimiti Maimaitituoheti 155

MUCIC@LT-EDI-ACL2022: Hope Speech Detection using Data Re-Sampling and 1D Conv-LSTM
Anusha M D Gowda, Fazlourrahman Balouchzahi, Hosahalli Lakshmaiah Shashirekha and Grigori Sidorov 161

DeepBlues@LT-EDI-ACL2022: Depression level detection modelling through domain specific BERT and short text Depression classifiers
Nawshad Farruque, Osmar Zaiane, Randy Goebel and Sudhakar Sivapalan 167

SSN_ARMM@ LT-EDI -ACL2022: Hope Speech Detection for Equality, Diversity, and Inclusion Using ALBERT model
Praveenkumar Vijayakumar, Prathyush S, Aravind P, Angel Deborah S, Rajalakshmi Sivanaiah, Sakaya Milton Rajendram and Mirnalinee T T 172

SUH_ASR@LT-EDI-ACL2022: Transformer based Approach for Speech Recognition for Vulnerable Individuals in Tamil
Suhasini S and Bharathi B 177

LPS@LT-EDI-ACL2022:An Ensemble Approach about Hope Speech Detection
Yue Ying Zhu 183

CURAJ_IIITDWD@LT-EDI-ACL 2022: Hope Speech Detection in English YouTube Comments using Deep Learning Techniques
Vanshita Jha, Ankit Kumar Mishra and Sunil Saumya 190

SSN_MLRG3 @LT-EDI-ACL2022-Depression Detection System from Social Media Text using Transformer Models
Sarika Esackimuthu, Shruthi Hariprasad, Rajalakshmi Sivanaiah, Angel Deborah S, Sakaya Milton Rajendram and Mirnalinee T T 196

BERT 4EVER@LT-EDI-ACL2022-Detecting signs of Depression from Social MediaDetecting Depression in Social Media using Prompt-Learning and Word-Emotion Cluster
Xiaotian Lin, Yingwen Fu, Ziyu Yang, Nankai Lin and Shengyi Jiang 200

CIC@LT-EDI-ACL2022: Are transformers the only hope? Hope speech detection for Spanish and English comments
Fazlourrahman Balouchzahi, Sabur Butt, Grigori Sidorov and Alexander Gelbukh 206

scubeMSEC@LT-EDI-ACL2022: Detection of Depression using Transformer Models
Sivamanikandan S, Santhosh V, Sanjaykumar N, Jerin Mahibha C and Thenmozhi Durairaj . . 212

SSNCSE_NLP@LT-EDI-ACL2022:Hope Speech Detection for Equality, Diversity and Inclusion using sentence transformers
Bharathi B, Dhanya Srinivasan, Josephine Varsha, Thenmozhi Durairaj and Senthil Kumar B 218

SOA_NLP@LT-EDI-ACL2022: An Ensemble Model for Hope Speech Detection from YouTube Comments
Abhinav Kumar, Sunil Saumya and Pradeep Kumar Roy 223

IIT Dhanbad @LT-EDI-ACL2022- Hope Speech Detection for Equality, Diversity, and Inclusion
Vishesh Gupta, Ritesh Kumar and Rajendra Pamula 229

IISERB@LT-EDI-ACL2022: A Bag of Words and Document Embeddings Based Framework to Identify Severity of Depression Over Social Media
Tanmay Basu 234

SSNCSE_NLP@LT-EDI-ACL2022: Homophobia/Transphobia Detection in Multiple Languages using SVM Classifiers and BERT-based Transformers
Krithika Swaminathan, Bharathi B, Gayathri G L and Hrishik Sampath 239

KUCST@LT-EDI-ACL2022: Detecting Signs of Depression from Social Media Text
Manex Agirrezabal and Janek Amann 245

E8-IJS@LT-EDI-ACL2022 - BERT, AutoML and Knowledge-graph backed Detection of Depression
Ilija Tavchioski, Boshko Koloski, Blaž Škrlj and Senja Pollak 251

Nozza@LT-EDI-ACL2022: Ensemble Modeling for Homophobia and Transphobia Detection
Debora Nozza 258

KADO@LT-EDI-ACL2022: BERT-based Ensembles for Detecting Signs of Depression from Social Media Text
Morteza Janatdoust, Fatemeh Ehsani-Besheli and Hossein Zeinali 265

Sammaan@LT-EDI-ACL2022: Ensembled Transformers Against Homophobia and Transphobia
Ishan Sanjeev Upadhyay, Kv Aditya Srivatsa and Radhika Mamidi 270

OPI@LT-EDI-ACL2022: Detecting Signs of Depression from Social Media Text using RoBERTa Pre-trained Language Models
Rafał Poświata and Michał Wiktor Perełkiewicz 276

FilipN@LT-EDI-ACL2022-Detecting signs of Depression from Social Media: Examining the use of summarization methods as data augmentation for text classification
Filip Nilsson and György Kovács 283

NAYEL @LT-EDI-ACL2022: Homophobia/Transphobia Detection for Equality, Diversity, and Inclusion using SVM
Nsrin Ashraf, Mohamed Taha, Ahmed Taha Abd Elfattah and Hamada Nayel 287

giniUs @LT-EDI-ACL2022: Aasha: Transformers based Hope-EDI
Harshul Raj Surana and Basavraj Chinagundi 291

SSN_MLRG1@LT-EDI-ACL2022: Multi-Class Classification using BERT models for Detecting Depression Signs from Social Media Text
Karun Anantharaman, Angel Deborah S, Rajalakshmi Sivanaiah, Saritha Madhavan and Sakaya Milton Rajendram 296

DepressionOne@LT-EDI-ACL2022: Using Machine Learning with SMOTE and Random UnderSampling to Detect Signs of Depression on Social Media Text.
Suman Dowlagar and Radhika Mamidi 301

LeaningTower@LT-EDI-ACL2022: When Hope and Hate Collide
Arianna Muti, Marta Marchiori Manerba, Katerina Korre and Alberto Barrón-Cedeño 306

MUCS@Text-LT-EDI@ACL 2022: Detecting Sign of Depression from Social Media Text using Supervised Learning Approach
Asha Hegde, Sharal Coelho, Ahmad Elyas Dashti and Hosahalli Lakshmaiah Shashirekha . . . 312

SSNCSE_NLP@LT-EDI-ACL2022: Speech Recognition for Vulnerable Individuals in Tamil using pretrained XLSR models
Dhanya Srinivasan, Bharathi B, Thenmozhi Durairaj and Senthil Kumar B 317

IDIAP_TIET@LT-EDI-ACL2022 : Hope Speech Detection in Social Media using Contextualized BERT with Attention Mechanism
Deepanshu Khanna, Muskaan Singh and Petr Motlicek 321

SSN@LT-EDI-ACL2022: Transfer Learning using BERT for Detecting Signs of Depression from Social Media Texts
Adarsh S and Betina Antony 326

Findings of the Shared Task on Detecting Signs of Depression from Social Media
Kayalvizhi S, Thenmozhi Durairaj, Bharathi Raja Chakravarthi and Jerin Mahibha C 331

Findings of the Shared Task on Speech Recognition for Vulnerable Individuals in Tamil
Bharathi B, Bharathi Raja Chakravarthi, Subalalitha CN, Sripriya N, Arunaggiri Pandian and Swetha Valli 339

DLRG@LT-EDI-ACL2022:Detecting signs of Depression from Social Media using XGBoost Method
Herbert Goldwin Sharen and Ratnavel Rajalakshmi 346

IDIAP Submission@LT-EDI-ACL2022 : Hope Speech Detection for Equality, Diversity and Inclusion
Muskaan Singh and Petr Motlicek 350

IDIAP Submission@LT-EDI-ACL2022: Homophobia/Transphobia Detection in social media comments
Muskaan Singh and Petr Motlicek 356

IDIAP Submission@LT-EDI-ACL2022: Detecting Signs of Depression from Social Media Text
Muskaan Singh and Petr Motlicek 362

Overview of The Shared Task on Homophobia and Transphobia Detection in Social Media Comments
Bharathi Raja Chakravarthi, Ruba Priyadharshini, Thenmozhi Durairaj, John Philip McCrae, Paul Buitelaar, Prasanna Kumar Kumaresan and Rahul Ponnusamy 369

Overview of the Shared Task on Hope Speech Detection for Equality, Diversity, and Inclusion
Bharathi Raja Chakravarthi, Vigneshwaran Muralidaran, Ruba Priyadharshini, Subalalitha CN, John Philip McCrae, Miguel Ángel García, Salud María Jiménez-Zafra, Rafael Valencia-García, Prasanna Kumar Kumaresan, Rahul Ponnusamy, Daniel García-Baena and José Antonio García-Díaz . 378