

Interspeech 2022

**The 1st Workshop on  
Speech for Social Good**

**Proceedings of the Workshop**

September 24 - 25, 2022

## **Introduction**

The 1st Workshop on Speech for Social Good (S4SG) will be hosted by Interspeech 2022 and held on 24, 25 September, 2022. The workshop is being organised by Anurag Katakhar, Alan W Black, Vikram Ramanarayanan, Sakriani Sakti, Shrimai Prabhume, Sai Krishna Rallabandi, Sunayana Sitaram, and Anirudh Koul.

The purpose of this workshop is to focus on applications of speech processing to areas of social good including but not limited to low-resource languages, the medical domain, bias in speech technologies, and ethical concerns in their development and deployment. To ensure sufficient attention to each domain, the workshop is split into three focused sessions. Session 1, Low Resource Languages, casts light on challenges in collecting and curating datasets from low resource regions of India, North America and Africa, and their eventual use in building speech technologies. Session 2 will include a panel discussion on “Data Collection, Bias, and Ethical Concerns in Speech Processing”, and focus on broader applications such as assistive technology for children’s communication. Finally, session 3, Medical Applications, will focus on building assistive speech technology for the medically impaired.

The workshop provides a forum for bringing together researchers, from academia and industry, from a breadth of application areas, and finding similarities in the challenges of building speech technologies for the long tail of users. It enables discussion on future research directions and best practices for bringing these users into the mainstream.



**Organizers :**

Anurag Katakkar, NVIDIA Corporation, USA  
Alan W Black, Carnegie Mellon University, USA  
Vikram Ramanarayanan, Modality AI and UCSF, USA  
Sakriani Sakti, JAIST, Japan  
Shrimai Prabhumoye, NVIDIA Corporation, USA  
Sai Krishna Rallabandi, Fidelity Investments, USA  
Sunayana Sitaram, Microsoft Research, India  
Anirudh Koul, Pinterest, USA

**Program Committee :**

Daan van Esch, Google  
Preethi Jyothi, IIT Bombay, India  
Chi-Chun Lee, National Tsing Hua University, Taiwan  
Robert MacDonald, Google, USA  
Shrimai Prabhumoye, NVIDIA Corporation, USA  
Sai Krishna Rallabandi, Fidelity Investments, USA  
Vikram Ramanarayanan, Modality AI and UCSF, USA  
Elizabeth Salesky, Johns Hopkins University, USA  
Katrín Tomanek, Google, USA  
Shinji Watanabe, Carnegie Mellon University, USA  
Brain Yan, Carnegie Mellon University, USA  
Heiga Zen, Google, Japan

**Keynote Speakers and Panelists :**

Pratyush Kumar, IIT Madras, India  
Vikram Ramanarayanan, Modality AI and UCSF, USA  
Sakriani Sakti, JAIST, Japan  
Odette Scharenborg, TU Delft, The Netherlands  
Emily Ahn, University of Washington, USA  
Gopala Anumanchipalli, UC Berkeley, USA



# Conference Program

"Saturday September 24, 2022"

**11:00–13:30 Session 1 : Low Resource/Multilingual NLP**

Chair: Gopala Anumanchipalli

11:00–12:00 *Keynote : "What does it really take to build speech recognition systems for the next billion users?"*

Pratyush Kumar, IIT Madras

12:10–12:30 *Annotated Speech Corpus for Low Resource Indian Languages: Awadhi, Bhojpuri, Braj and Magahi*

Ritesh Kumar, Siddharth Singh, Shyam Ratan, Mohit Raj, Sonal Sinha, bornini lahiri, Vivek Seshadri, Kalika Bali and Atul Kr. Ojha

12:30–12:50 *Towards an Automatic Speech Recognizer for the Choctaw language*

Jacqueline Brixey and David Traum

12:50–13:10 *Building TTS systems for low resource languages under resource constraints*

Perez Ogayo, Graham Neubig and Alan W Black

**13:10–13:30 Breakout Room (Gather Town)**

**22:00–00:30 Session 2 : Panel Discussion and Misc. Topics**

Chair: Anurag Katakhar

22:00–23:00 *Panel : "Data Collection, Bias, and Ethical Concerns in Speech Processing"*

Odette Scharenborg (TU Delft), Emily Ahn (UW), Gopala Anumanchipalli (UC Berkeley), Sakriani Sakti (JAIST), and Alan W Black (CMU)

23:10–23:30 *Can Smartphones be a cost-effective alternative to LENA for Early Childhood Language Intervention?*

Satwik Dutta, Jacob C. Reyna, Jay F. Buzhardt, Dwight Irvin and John H.L. Hansen

23:30–23:50 *Comparing data augmentation and training techniques to reduce bias against non-native accents in hybrid speech recognition systems*

Yixuan Zhang, Yuanyuan Zhang, Tanvina Patel and Odette Scharenborg

23:50–00:10 *Text Normalization for Speech Systems for All Languages*

Athiya Deviyani and Alan W Black

**"Sunday September 25, 2022"**

**00:10–00:30** *Breakout Room (Gather Town)*

**11:00–13:30** **Session 3 : Medical Applications**

Chair: Sai Krishna Rallabandi

11:00–12:00 *Keynote : "Multimodal Dialog Technologies for Neurological and Mental Health"*  
Vikram Ramanarayanan, Modality AI and UCSF

12:10–12:30 *Assessing ASR Model Quality on Disordered Speech using BERTScore*  
Jimmy Tobin, Qisheng Li, Subhashini Venugopalan, Katie Seaver, Richard Cave  
and Katrin Tomanek

12:30–12:50 *Cross-Teager Cepstral Coefficients For Dysarthric Severity Level Classification*  
Anand Therattil, Aastha Kachhi and Hemant Patil

12:50–13:10 *Highly Intelligible Speech Synthesis for Spinal Muscular Atrophy Patients Based on  
Model Adaptation*  
Takuma Yoshimoto, Ryoichi Takashima, Chiho Sasaki and Tetsuya Takiguchi

**13:10–13:30** *Breakout Room (Gather Town)*