Special Session Title: Speech AI for All: Challenging Deficit-based Approaches to Speech Science and Technology

Organizers

- Robin Netzorg (robert_netzorg@berkeley.edu)
    - University of California, Berkeley (USA)
    - Robin is a EECS PhD candidate advised by Gopala Anumanchipalli and Bin Yu at the University of California, Berkeley. With particular interest in building technology for gender-affirming voice training, Robin specializes in modeling speaker perception and identity alongside community members and leaders.
- Juliana Francis (jfrancis@kth.se)
    - KTH, Stockholm (Sweden)
    - Juliana is a PhD Student in the Speech, Music, and Hearing (TMH) department at KTH Royal Institute of Technology in Stockholm, Sweden, supervised by Joakim Gustafsson and Éva Székely. Her research interest is focused on expressive and controllable speech synthesis speech, in part for augmentative and alternative communication devices. Juliana is also a current editor for the ISCA LGBTQI* webpage.
- Brooklyn Sheppard (brooklyn.sheppard1@ucalgary.ca)
    - University of Calgary
    - Calgary, Alberta, Canada
    - Brooklyn (she/they) is a PhD candidate in Linguistics, supervised by Dr. Stephen Winters, where her research aims to bridge the fields of linguistics, cognitive science, and speech modelling. In industry, she has been an advocate for fair and inclusive speech and language technology, with hands-on experience leading a collaborative dataset creation process for misogynistic language in English as an NLP Scientist at Mila - Quebec AI, and an internship project with Meta dedicated to developing a conceptual framework for community-led speech dataset creation by, for, and with marginalized communities.
- Kimi Wenzel (kwenzel@andrew.cmu.edu)
    - Carnegie Mellon University (USA)
    - Kimi Wenzel is a PhD candidate at Carnegie Mellon University's Human-Computer Interaction Institute. Her research centers on understanding the downstream representational harms of speech AI, with a focus on psychological impacts. Her work examining speech diversities in AI has won awards from CHI and the Center for Machine Learning and Health. She has previously co-led workshops on speech AI and representation at CHI and FAccT.
- Norman Makoto Su (normsu@ucsc.edu)
    - University of California, Santa Cruz (USA)
    - Norman Makoto Su is an associate professor in the Department of Computational Media at UC Santa Cruz. His research interests lie in human–computer interaction (HCI) and computer–supported cooperative work (CSCW). He directs the Authentic User Experience Lab (AUX Lab), where they integrate empirical and humanistic methods in Human Computer Interaction (HCI) to study and

design with subcultures. He has published work on collective action and around new forms of and challenges with AI work and the techlash.

- Anna Leschanowsky (anna.leschanowsky@iis.fraunhofer.de)
    - Fraunhofer Institute for Integrated Circuits (IIS)
    - Anna Leschanowsky is a Senior Scientist at Fraunhofer IIS, specialising in the intersection of voice technology, HCI and privacy-preserving systems. She has been part of the FairEVA project, an open-source project gathering resources and building tools to help researchers and developers, technology activists, and voice technology users evaluate and audit bias and discrimination in voice technologies.
- Abraham Glasser (abraham.glasser@gallaudet.edu)
    - Gallaudet University
    - Washington, DC, USA
    - Dr. Abraham Glasser is an Assistant Professor at Gallaudet University in Washington, DC, USA. Dr. Glasser is a born-Deaf person who can and does use his voice, and experiences barriers with speech-based technologies. He is a faculty member in the Accessible Human-Centered Computing and Policy program, where he and his students conduct Human-Computer Interaction (HCI) research on accessible technologies for Deaf and Hard-of-Hearing users. At Gallaudet, Dr. Glasser is also co-director of the Deaf and Hard of Hearing Technology Rehabilitation Engineering Research Center (DHH-RERC), and an active member of the Coalition for Sign Language Equity in Technology (CoSET).
- Bowon Lee (bowon.lee@inha.ac.kr)
    - Inha University
    - Bowon Lee joined the faculty of the Department of Electronic Engineering at Inha University in March 2014, where he served as the department head from 2021 to 2023. He has also held a position of affiliate professor in the Department of Artificial Intelligence since 2020. His research interests include speech signal processing, speech analysis, and speech-based inclusive technologies. He is currently serving as General Chair for Workshop on Automatic Speech Recognition and Understanding (ASRU) 2025. In the past, he has served as Awards Chair for the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2018, Local Operations Chair for Interspeech 2022, and a member of the organizing committee for Young Female Researchers in Speech Workshop (YFRSW) 2022, along with the technical program committee of numerous conferences and workshops.
- Christine Murad (c2murad@uwaterloo.ca)
    - Carleton University & University of Waterloo
    - Christine Murad is a Post-Doctoral Fellow in the SPIRL lab at Carleton University, and a Research Associate at the University of Waterloo, in the Technologies for Aging Gracefully lab. Christine is currently based in Waterloo, Ontario (Canada). Christine earned her PhD (2024), MSc (2019), and her Honours B.Sc in Computer Science from the University of Toronto. Christine's research involves exploring the design of human-centered AI, with a focus on conversational voice

interfaces and the development of design heuristics and tools to assist in conversational voice interface design. Christine is a founding member of the ACM Conversational User Interfaces (CUI) Steering Committee since its inception in 2019.

- Cosmin Munteanu (cosmin.munteanu@uwaterloo.ca)
    - University of Waterloo, Canada
    - Cosmin Munteanu is an Associate Professor and Schlegel Research Chair in Technology for Healthy Aging at the Department of Systems Design Engineering, University of Waterloo, and Director of the Technologies for Ageing Gracefully lab. They are a transdisciplinary scholar, drawing from a wide range of disciplines such as engineering, computing sciences, critical theory, and technology and society studies. Cosmin takes a primarily ethnomethodological approach to study how to design intelligent applications that improve access to information, support social connections late in life, and reduce digital marginalization for underrepresented groups such as older adults. Their work is situated at the intersection of user experience design, digital inclusion, aging, natural language processing, and ethics, primarily focusing on the sociotechnical design of inclusive interfaces with and for older adults.
- Jay Cunningham (jcunni37@depaul.edu)
    - Dr. Jay L. Cunningham is a computer scientist and scholar with over a decade of interdisciplinary research spanning computing, public interest technology, and AI policy — with a focus on human-centered AI, AI ethics, data governance, and technology equity. He is dedicated to advancing responsibility and fairness in artificial intelligence and machine learning (AI/ML) systems, data science, and human-computer interaction (HCI).
- Shaomei Wu (shaomei@aimpower.org)
    - Shaomei Wu
    - Shaomei is a person who stutters and the founder and CEO of AImpower.org, a tech nonprofit dedicated to research and develop inclusive AI technologies for, with, and by the disabled community. Her recent research focuses on fair and inclusive speech research and technologies for stuttered speech.

W**hat makes the session special?**

- While commonly seen in the InterSpeech community, the pathology based speech technologies for marginalized communities are not only poorly adopted by the targeted communities but also (in some cases) reinforcing existing social discriminations and biases.
- We want to push the field to look beyond the pathological/medical model and engage with the social model of disability, e.g. seeing disfluent speech not as something to be fixed but something to accommodate.
- To close the gap between research and real world adoption and impact, we need to work more closely with the impacted communities and other stakeholders, and engage them early on. There is an opportunity to learn and adopt the user-centered, participatory approaches from the HCI field in speech science and technology.

- This special session aims to create a dedicated space to engage with the HCI community, inviting submissions that explore and demonstrate at the user-centered, community-first approach in studying and evaluating speech technology. Through discussions and exchanges with the HCI community, this session can introduce user-centered expertise and methods to the InterSpeech community and empower the InterSpeech community with new ideas and tools to bridge the gap between research and real world impact.

Summary

In recent years, InterSpeech has promoted research agendas around fair and inclusive speech science and technology, seeking to promote research on and by historically marginalized and/or pathologized groups. Despite this focus, current research directions, namely those rooted in diagnostic assessment of “pathological” speech, run the risk of branding affected communities as deficient or abnormal, necessitating both technological surveillance and “correction” of their speech.

Work at InterSpeech on fair and inclusive speech science and technology is primarily quantitative in nature, but we believe quantitative approaches alone are insufficient to truly achieve fairness and inclusivity. As other conferences that are concerned with these topics such as FAccT and CHI encourage both quantitative and qualitative work, we propose this special session in the hope of providing a home for work grounded in qualitative and HCI-based methods for system design, analysis, and critique.

This special session is a continuation of many conversations the organizers have had throughout the years both inside and outside of InterSpeech. Most recently at CHI 2025, we hosted the SpeechAI4All workshop, which brought together speech researchers, HCI researchers, and impacted communities to discuss the current and future state of speech technology. These conversations deserve a home within InterSpeech to promote truly fair, accessible, inclusive, and equitable speech science and technology. We cannot solely focus on the technological contributions of our work, but also their downstream, human impacts.

Topics Include:

- Qualitative and HCI Approaches to Speech Technology
- User Studies of Speech Technology and Interfaces
- Auditing Speech Papers and Datasets
- Community-Led Speech Research

The special session will consist of an invited panel with community members (1 hour), and a poster presentation session (1 hour).

The following researchers and their groups have expressed interest in submitting a paper to the session:

- Robin Netzorg (UC Berkeley)
- Shaomei Wu (AImpower)
- Abraham Glasser (Gallaudet University)
- Christian Vogler (Carleton University & University of Waterloo)
- Raja Kushalnagar (Gallaudet University)
- Bowon Lee (Inha University)
- Cosmin Munteanu (University of Waterloo, Canada)
- Christine Murad (independent)
- Jenny Waycott (U Melbourne)
- Barbara Barbosa Neves (University of Sydney)
- Roisin McNaney (Monash University)
- Minha Lee (Eindhoven University of Technology)
- Sneha Das (Technical University of Denmark)
- Ingo Siegert (Otto-von-Guericke-University Magdeburg)
- Juliana Francis (KTH Stockholm)
- Brooklyn Sheppard (University of Calgary)

The following are potential reviewers of the session:

- Nicole Holliday
- Rong Gong
- Alisha Pradhan
- Sergio Sayago
- Ben Cowan
- Sharifa Sultana
- Tal August
- Odette Schourenberg
- Nicholas Cummins