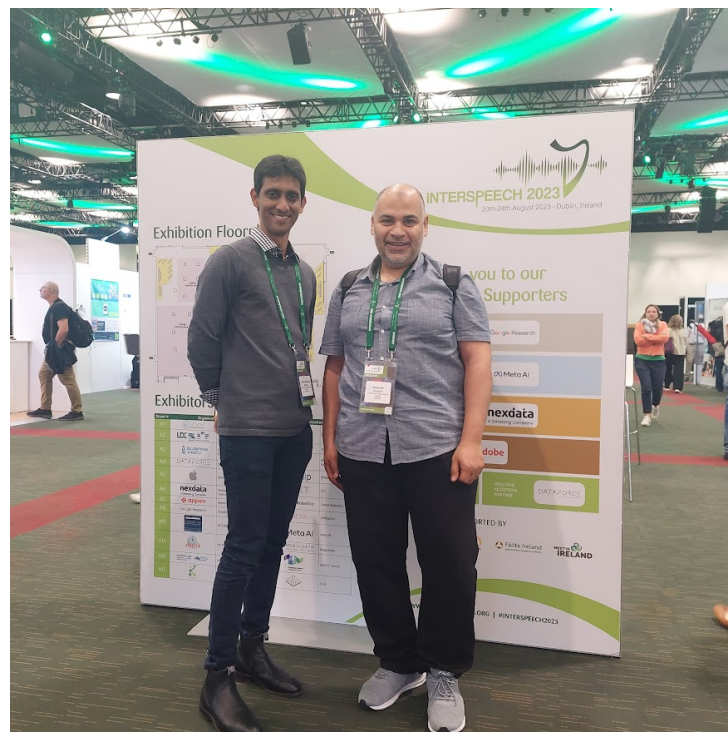


Interspeech2023 - Mostafa Shahin

Interspeech is an annual conference dedicated to the field of spoken language processing and speech communication. It serves as a global forum for researchers, academics, and professionals to exchange ideas, present the latest advancements in speech technology, and discuss challenges in the realm of human speech and its applications. This conference typically features a wide range of topics, including automatic speech recognition, speech synthesis, natural language processing, and the cognitive aspects of speech communication. Interspeech provides a valuable platform for fostering collaboration and innovation in this dynamic and evolving field.

In the heart of the vibrant and historic city of Dublin, this year's conference marked my fifth attendance at Interspeech since 2012. Over the years, the dominant focus has been on speech processing and machine learning, with limited attention to speech science. However, the noticeable shift this year was the increased interest in speech science, linguistics, and phonetics disciplines, as reflected in the growing presence of researchers from these areas. Moreover, there was a great interest in research related to health such as dementia, speech disorders, hearing disorders, medical conversations, etc. There were two special sessions focusing on speech and language in health and a show-and-tell session for health applications.



ISCA Medal for Scientific Achievement went this year to Shri Narayanan for his significant research in the field of speech and audio signal processing, particularly for his pioneering work in emotion recognition and multimodal human-computer interaction systems.

In the current year, I had the opportunity to showcase our research on an innovative automatic language identification approach, which relied on a self-supervised speech representation model. Our findings highlighted the effectiveness of training the model to capture phonological features like articulatory manners and places, substantially enhancing its proficiency in language identification through speech signals. This research was a part of our involvement in a language identification competition centred on Zoom recordings of parent-child conversations in both English and Mandarin. I am pleased to note that our team secured the second position among the five participating teams.

