

<b>Day 1, 14 November 2022, Monday (Hybrid mode)</b> <b>Time Zone: (IST)</b>		<b>PID</b>
09:30-10:00	<b>Zoom connection</b>	
10:00-10:30	<b>Opening ceremony</b>	
10:30-12:00	Session 1 (6 presentations, 15 min) - Asia/India (hybrid mode) <b>Audio Signal Processing</b>	<b>6719, 3745, 1192, 7422, 8071, 465</b>
12:00-12:20	<b>Coffee break</b>	
12:20-13:50	Session 2 (6 presentations, 15 min) - Europe/Russia (hybrid) <b>Multimodal Interaction</b>	<b>7244, 1330, 6491, 3011, 7957, 2438</b>
13:50-14:50	<b>Lunch break</b>	
14:50-16:50	Session 3 (8 presentations, 15 min) - Europe/Russia (online) <b>Speech Recognition and Synthesis</b>	<b>3033, 1422, 6086, 8471, 9074, 3687, 2006, 1694</b>
16:50-17:00	<b>Coffee break</b>	
17:00-18:00	Keynote lecture by Prof. SHRIKANTH NARAYANAN (USA, online) <b>Enabling Trustworthy Speech-centric Behavioral Machine Intelligence</b>	
<b>Day 2, 15 November 2022, Tuesday (Hybrid mode)</b> <b>Time Zone: (IST)</b>		<b>PID</b>
09:00-09:30	<b>Zoom connection</b>	
09:30-11:30	Session 4 (8 presentations, 15 min) - Asia /India (hybrid mode) <b>Speech Analysis in Medicine</b>	<b>2278, 2975, 4491, 4987, 7425, 2434, 807, 7931</b>
11:30-11:50	<b>Coffee break</b>	
11:50-13:20	Session 5 (6 presentations, 15 min)- Europe/Russia (hybrid) <b>Computational Paralinguistics</b>	<b>7819, 8441, 5260, 6245, 3497, 3996</b>
13:20-14:30	<b>Lunch break</b>	
14:30-15:30	Keynote lecture by Prof. GERARD BAILLY (France, online): <b>Exploring latent spaces of end to end text to speech synthesis systems</b>	
15:30-15:50	<b>Coffee break</b>	
15:50-17:20	Session 6 (6 presentations, 15 min) - America/Russia (online) <b>Speaker Recognition</b>	<b>2267, 3513, 282, 4678, 8010, 3235</b>
17:30 -	<b>Banquet</b>	
<b>Day 3, 16 November 2022, Wednesday (Hybrid mode)</b> <b>Time Zone: (IST)</b>		<b>PID</b>
09:00-09:30	<b>Zoom connection</b>	
09:30-11:30	Session 7 (8 presentations, 15 min) - Asia /India (hybrid mode) <b>Speech and Language Resources</b>	<b>5391, 9853, 475, 6960, 7847, 741, 127, 6370</b>
11:30-11:50	<b>Coffee break</b>	
11:50-13:20	Session 8 (6 presentations, 15 min) - Europe /Russia (hybrid) <b>Spoken Language Processing</b>	<b>7206, 1903, 8900, 2051, 1580, 1383</b>
13:20-14:30	<b>Lunch break</b>	
14:30-15:30	Keynote lecture by Prof. HEMA A MURTHY (India, in -person) <b>Development of text to speech synthesis systems for Indian languages</b>	
15:30-15:50	<b>Coffee break</b>	
15:50-17:20	Session 9 (6 presentations, 15 min)- America /Russia (online) <b>Speech Prosody</b>	<b>5624, 2418, 9282, 288, 4021, 4183</b>
17:20-17:30	<b>Closing ceremony</b>	

## Day 1

### 10:30-12:00 Session 1 - Audio Signal Processing (6 presentations, India hybrid mode)

- 6719 Speech Music Overlap Detection using Spectral Peak Evolutions - Mrinmoy Bhattacharjee, S. R. Mahadeva Prasanna and Prithwijit Guha (India)
- 3745 Overlapped Speech Detection Using AM-FM based Time-Frequency Representations - Shikha Baghel, S. R. Mahadeva Prasanna and Prithwijit Guha (India)
- 1192 Device Robust Acoustic Scene Classification using Adaptive Noise Reduction and Convolutional Recurrent Attention Neural Network - Spoorthy V and Shashidhar G. Koolagudi (India)
- 7422 Significance of Distance on Pop Noise for Voice Liveness Detection - Priyanka Gupta and Hemant A. Patil (India)
- 8071 Analysis of Time-Averaged Feature Extraction Techniques on Infant Cry Classification - Aditya Pusuluri, Aastha Kachhi and Hemant A. Patil (India)
- 465 Classifying Mahout and Social Interactions of Asian Elephants based on Trumpet Calls - Seema Lokhandwala, Priyankoo Sarmah and Rohit Sinha (India)

### 12:20-13:50 Session 2 - Multimodal Interaction (6 presentations, Europe & Russia hybrid)

- 7244 DyCoDa: A Multi-Modal Data Collection of Multi-User Remote Survival Game Recordings - Denis Dresvyanskiy, Yamini Sinha, Matthias Busch, Ingo Siegert, Alexey Karpov and Wolfgang Minker (Germany, Russia)
- 1330 Emotional Speech Recognition Based on Lip-Reading - Elena Ryumina and Denis Ivanko (Russia)
- 6491 Effects of Depth of Field on Focus using a Virtual Reality Escape Room - Nikolaos Tsiftsis, Konstantinos Moustakas and Nikolaos Fakotakis (Greece)
- 3011 Should We Believe Our Eyes or Our Ears? Processing Incongruent Audiovisual Stimuli by Russian Listeners - Elena Riekhakaynen and Elena Zatevalova (Russia)
- 7957 Nonverbal Constituents of Argumentative Discourse: Gesture and Prosody Interaction - Anna Leonteva and Tatiana Sokoreva (Russia)
- 2438 Dynamics of Frequency Characteristics of Visually Evoked Potentials of Electroencephalography During the Work with Brain-Computer Interfaces - Yaroslav Turovsky, Daniyar Wolf, Roman Meshcheryakov and Anastasia Iskhakova (Russia)

### 14:50-16:50 Session 3 - Speech Recognition and Synthesis (8 presentations, World hybrid)

- 3033 Influence of Accented Speech in Automatic Speech Recognition: A Case Study on Assamese L1 Speakers Speaking Code Switched Hindi-English - Joyshree Chakraborty, Rohit Sinha and Priyankoo Sarmah (India)
- 1422 Study of Speech Recognition System Based on Transformer and Connectionist Temporal Classification Models for Low Resource Language - Shweta Bansal, Shambhu Sharan and Shyam S Agrawal (India)
- 6086 Low-Cost Training of Speech Recognition System for Hindi ASR Challenge 2022 - Alexander Zatvornitskiy (Russia)

- 8471 Investigation of Transfer Learning for End-to-End Russian Speech Recognition - Irina Kipyatkova (Russia)
- 9074 Comparing NLP Solutions for the Disambiguation of French Heterophonic Homographs for End-to-End TTS Systems - Maria-Loulou Hajj, Gerard Bailly, Martin Lenglet and Olivier Perrotin (France)
- 3687 Low-resource Emotional Speech Synthesis: Transfer Learning, Data requirements and Adversarial Training - Anton Nesterenko, Ruslan Akhmerov, Yulia Matveeva, Anna Goremykina, Dmitry Astankov, Evgeniy Shuranov and Alexandra Shirshova (Russia)
- 2006 An Initial Study on Birdsong Re-synthesis using Neural Vocoders Rhythm - Rajiv Bhatia and Tomi H. Kinnunen (Finland)
- 1694 CRIM's Speech Recognition System for OpenASR21 Evaluation with Conformer and Voice Activity Detector Embeddings - Vishwa Gupta and Gilles Boulianne (Canada)

## Day 2

### 09:30-11:30 Session 4 - Speech Analysis in Medicine (8 presentations, India & Russia hybrid)

- 2278 Continuous Wavelet Transform for Severity-Level Classification of Dysarthria - Aastha Kachhi, Anand Therattil, Priyanka Gupta and Hemant A. Patil (India)
- 2975 Significance of Energy Features for Severity Classification of Dysarthria - Aastha Kachhi, Anand Therattil, Ankur T. Patil, Hardik B. Sailor and Hemant A. Patil (India)
- 4491 Neural Network Based Curve Fitting to Enhance the Intelligibility of Dysarthric Speech - K S Pavithra, H M Chandrashekar and Veena Karjigi (India)
- 4987 Assessment of Speech Quality during Speech Rehabilitation Based on the Solution of the Classification Problem - Evgeny Kostyuchenko, Ivan Rakhmanenko and Lidiya Balatskaya (Russia)
- 7425 Fuzzy Classifier For Speech Assessment in Speech Rehabilitation - Dariya Novokhrestova, Ilya Hodashinsky, Evgeny Kostyuchenko, Konstantin Sarin and Marina Bardamova (Russia)
- 2434 Significance of Dimensionality Reduction in CNN-based Vowel Classification from Imagined Speech using Electroencephalogram Signals - Oindrila Banerjee, Govind D, Akhilesh Kumar Dubey and Suryakanth V Gangashetty (India)
- 807 Recognition of the Emotional State of Children with Down Syndrome by Video, Audio and Text Modalities: Human and Automatic - Elena Lyakso, Olga Frolova, Anton Matveev, Yuri Matveev, Aleksey Grigorev, Olesia Makhnytkina and Nersisson Ruban (Russia, India)
- 7931 Detection of Speech Related Disorders by Pre-Trained Embedding Models Extracted Biomarkers - Attila Zoltán Jenei, Gábor Kiss and Dávid Sztahó (Hungary)

### 11:50-13:20 Session 5 - Computational Paralinguistics (6 presentations, World hybrid)

- 7819 Self-Configuring Genetic Programming Feature Generation in Affect Recognition Tasks - Danila Mamontov, Wolfgang Minker and Alexey Karpov (Germany, Russia)
- 8441 On the Use of Ensemble X-Vector Embeddings for Improved Sleepiness Detection - José Vicente Egas López, Róbert Busa-Fekete and Gábor Gosztolya (Hungary, USA)

- 5260 Fake Speech Detection using OpenSMILE Features - Devesh Kumar, Pavan Kumar V Patil, Ayush Agarwal and S. R. Mahadeva Prasanna (India)
- 6245 Fake Speech Detection using Modulation Spectrogram - Raghav Magazine, Ayush Agarwal, Anand Hegde and S. R. Mahadeva Prasanna (India)
- 3497 Multi-level Fusion of Fisher Vector Encoded BERT and wav2vec 2.0 Embeddings for Native Language Identification - Dani Krebbers, Heysem Kaya and Alexey Karpov (The Netherlands, Russia)
- 3996 An Electroglottographic Method for Assessing the Emotional State of the Speaker - Aleksey Grigorev, Anna Kurazhova, Egor Kleshnev, Aleksandr Nikolaev, Olga Frolova and Elena Lyakso (Russia)

**15:50-17:20 Session 6 - Speaker Recognition (6 presentations, World hybrid)**

- 2267 Comparing Unsupervised Detection Algorithms for Audio Adversarial Examples - Shanatip Choosaksakunwiboon, Karla Pizzi and Ching-Yu Kao (Germany)
- 3513 Importance of Supra-Segmental Information and Self-Supervised Framework for Spoken Language Diarization Task - Jagabandhu Mishra and S. R. Mahadeva Prasanna (India)
- 282 Deep Speaker Embeddings based Online Diarization - Anastasia Avdeeva and Sergey Novoselov (Russia)
- 4678 Neural Embedding Extractors for Text-Independent Speaker Verification - Jahangir Alam, Woohyun Kang and Abderrahim Fathan (Canada)
- 8010 An Analytic Study on Clustering-based Pseudo-Labels for Self-Supervised Deep Speaker Verification - Woohyun Kang, Jahangir Alam and Abderrahim Fathan (Canada)
- 3235 Multiresolution Decomposition Analysis via Wavelet Transforms for Audio Deepfake Detection - Abderrahim Fathan, Jahangir Alam and Woo Hyun Kang (Canada)

**Day 3**

**09:30-11:30 Session 7 - Speech and Language Resources (8 presentations, India & Russia hybrid)**

- 5391 A Multi-Modal Approach to Mining Intent from Code-Mixed Hindi-English Calls in the Hyperlocal-Delivery Domain - Jose Mathew, Pranjal Sahu, Bhavuk Singhal, Aniket Joshi, Krishna Medikonda and Jairaj Sathyanarayana (India)
- 9853 Automatic Rhythm and Speech Rate Analysis of Missing Spontaneous Speech - Parismita Gogoi, Priyankoo Sarmah and S. R. Mahadeva Prasanna (India)
- 475 Analysis-by-Synthesis Modeling of Bengali Intonation - Moumita Pakrashi and Shakuntala Mahanta (India)
- 6960 Coherence Based Automatic Essay Scoring Using Sentence Embedding and Recurrent Neural Networks - Ramesh Dadi and Suresh Kumar Sanampudi (India)
- 7847 Analysis of Automatic Evaluation Metric on Low-Resourced Language: BERTScore Vs BLEU Score - Goutam Datta, Nisheeth Joshi and Kusum Gupta (India)
- 741 Comparison of Word Embeddings of Unaligned Audio and Text Data Using Persistent Homology - Zhandos Yessenbayev and Zhanibek Kozhirbayev (Kazakhstan)

127 ClusterVote: Automatic Summarization Dataset Construction with Document Clusters - Daniil Chernyshev and Boris Dobrov (Russia)

6370 Thematic Diversity of Everyday Russian Discourse: a Case Study Based on the ORD corpus - Eleonora Akinshina and Tatiana Sherstinova (Russia)

**11:50-13:20 Session 8 - Spoken Language Processing (6 presentations, World hybrid)**

7206 Exploring the Use of Machine Learning for Resume Recommendations - Anna Shestakova and Andrea Corradini (UK, Denmark)

1903 Personalizing Retrieval-based Dialogue Agents - Pavel Posokhov, Anastasia Matveeva, Olesia Makhnytkina, Anton Matveev and Yuri Matveev (Russia)

8900 Categorization of Threatening Speech Acts - Liliya Komalova and Lyubov Kalyuzhnaya (Russia)

2051 Dictionary with the Evaluation of Positivity/Negativity Degree of the Russian Words - Valery Solovyev, Musa Islamov and Venera Bayrasheva (Russia)

1580 Logistics Translator. Concept Vision on Future Interlanguage Computer Assisted Translation - Rodmonga Potapova, Vsevolod Potapov and Oleg Kuzmin (Russia)

1383 Forensic Identification of Foreign-Language Speakers by the Method of Structural-Melodic Analysis of Phonograms - Rodmonga Potapova, Vsevolod Potapov and Irina Kuryanova (Russia)

**15:50-17:20 Session 9 - Speech Prosody (6 presentations, World online)**

5624 The Role of Pause in Interaction: A Case of Polylogue - Tatiana Sokoreva and Tatiana Shevchenko (Russia)

2418 Celtic English Continuum in Pitch Patterns of Spontaneous Talk: Evidence of Long-Term Contacts - Maria Chubarova and Tatiana Shevchenko (Russia)

9282 Prosodic Features of Verbal Irony in Russian and French: Universal vs. Language-Specific - Uliana Kochetkova, Pavel Skrelin, Rada German and Daria Novoselova (Russia)

288 Joint Changes in First and Second Formants of /a/, /i/, /u/ Vowels in Babble Noise - a New Statistical Approach - Alisa Gvozdeva, Alexander Lunichkin, Larisa Zaytseva, Elena Ogorodnikova and Irina Andreeva (Russia)

4021 Multi-Label Dysfluency Classification - Melanie Jouaiti and Kerstin Dautenhahn (Canada)

4183 Harnessing Uncertainty - Multi-Label Dysfluency Classification with Uncertain Labels - Melanie Jouaiti and Kerstin Dautenhahn (Canada)