



About the Summer School

Summer School on Speech Signal Processing (S4P) is being organized as a part of educational outreach activities at Speech Research Lab, DAU Gandhinagar. It will provide opportunities to students, faculties, researchers, and professionals to enhance their fundamentals and get exposed to research areas in the field of speech signal processing. The school constitutes of a theme topic and tutorials surrounding it. S4P 2025 is the 7th summer school at DAU following the successful earlier six summer schools, focusing on different topics in the broad area of speech signal processing.

FOCUS: Automatic Speech Recognition (ASR)

ASR system is now a key component of commercially successful Intelligent Personal Assistants (IPAs), such as Apple's Siri, Microsoft's Cortana, Google Assistant, etc. Design of ASR system depends upon various factors, such as near-field vs. far-field speech, recording and transmission channel conditions, acoustic model, language model, signal degradation conditions (acoustic noise), etc. Understanding these technological challenges is the major goal of S4P 2025.

Invited Speakers

Akihiko K. Sugiyama has 40 years of experience developing telecommunications, speech, and audio signal processing signal systems for consumer and network system products. The sole inventor or a co-inventor of 273 registered patents in Japan and overseas as well as 6 registered trademarks. He is a Fellow of IEEE and IEICE.

Thomas Hain is Professor of Computer Science at the University of Sheffield, UK. He holds the degree 'Dipl.-Ing' in Electrical and Communication Eng. from the University of Technology, Vienna, and a PhD in Information Eng. from the Cambridge University. He is a Fellow of ISCA.

Petr Motlicek is a senior researcher at IDIAP Research Institute in Martigny, Switzerland. He is an associate professor at Brno University of Technology (BUT), and an external teacher at EPFL. His research interests include audio and speech processing, and its application for various tasks, i.e., embedded platforms, robotics, forensic voice comparison, etc.

Mathew Magimai Doss is a senior researcher at IDIAP Research Institute in Martigny, Switzerland. He is also a lecturer at EPFL. He is an Associate Editor of the IEEE/ACM Transactions on Audio, Speech, and Language Processing. He is the coordinator of SNSF Sinergia project SMILE-II. His research interest lies in signal processing, statistical pattern recognition, artificial neural networks and computational linguistics.

Yu Tsao is Deputy Director at Research Center for Information Technology Innovation, Academia Sinica, Taiwan. He is currently an Associate Editor for the IEEE/ACM Trans. on Audio, Speech and Language Process., and IEEE Signal Process. Letters.

Nancy F. Chen received her PhD from MIT and Harvard in 2011. She is with Institute for Infocomm Research (I2R), Singapore. She works in generative AI, conversational AI, natural language generation and machine learning with applications related to education, healthcare, journalism, and defence. She is board member of ISCA, Senior Editor for the IEEE/ACM Trans. on Audio, Speech and Lang. Process and she was IEEE SPS Distinguished Lecturer in 2023. She is a fellow of ISCA.

Tatsuya Kawahara is currently a Professor with the School of Informatics, Kyoto University, Japan. He has authored and coauthored more than 450 academic papers on ASR, spoken language processing, and spoken dialogue systems. He is President of APSIPA and Secretary General of ISCA, Fellow of IEEE, and General Chair of IEEE ASRU 2007.

Invited Speakers

Bayya Yegnanarayana is INSA Hon Scientist at IIIT Hyderabad. His interest interests are in signal processing, speech, image processing and neural networks. He has published over 400 papers, and supervised over 80 theses at Master and PhD levels. He is Fellow of IEEE, ISCA, INAE, INSA, IASc, APAS and life Fellow of IIT Kharagpur. He led MeitY sponsored consortium project on phonetic engine (Phase-I&II)

Hema A. Murthy is Emeritus Professor (Hon) at IIT Madras. Having published over 300 research papers, she is a Fellow of ISCA, INAE, and APAIA. She was a member of ISCA Advisory Council. She competently led MeitY Sponsored consortium projects in TTS (Phase-I&II), and NLTM.

Srinivasan Umesh is a Professor in Electrical Engineering at IIT Madras. He received his PhD degree from University of Rhode Island, USA. His research interests include ASR, Speaker Normalization & Adaptation, Self Supervised Learning, Deep Learning and ML, and Speaker Recognition and Diarisation. He competently led MeitY Sponsored consortium projects in ASR (Phase-I&II), and NLTM.

Sriram Ganapathy is an Associate Professor with the EE, IISc Bangalore, where he leads LEAP Lab. He is also an associated with the Google Research India. He was a Research Staff Member with the IBM Watson Research Center, USA. He is the subject Editor for Elsevier Speech Communications Journal.

K. Sri Rama Murty obtained PhD from IIT Madras, now with IIT Hyderabad. His research interests include signal processing, speech analysis, recognition, synthesis, phase processing and modeling, and machine learning.

Vipul Arora obtained his PhD from IIT Kanpur, and he is with IIT Kanpur. He was a Research Scientist at Amazon Alexa, Boston. His research interests include machine learning, audio processing, machine learning for physics, and time series analysis.

Anil Kumar Vuppala obtained his PhD from IIT Kharagpur, and he is with IIIT Hyderabad. His research interests primarily lies in speech processing. He has given more than 100 invited talks in various workshops and conferences.

Vinayak Abrol obtained his PhD from IIT Mandi and now with IIIT Delhi. His research focuses on the design and analysis of numerical algorithms for information-inspired applications. His current work focuses on XAI methods for acoustic models and generative speech/audio language modeling.

Hemant A. Patil obtained his PhD degree from IIT Kharagpur and now with DAU Gandhinagar. He served as Distinguished Lecturer of ISCA and APSIPA, Associate Editor of IEEE Signal Processing Magazine. He published 350+ research publications. He is a member of IEEE, ISCA, APSIPA, EURASIP, and IUPRAI.

Industry Perspective Talks

Sri Garimella obtained Ph.D from Johns Hopkins University, Baltimore, USA, and now the Director, Applied Science, heading the Amazon Artificial General Intelligence (AGI, Speech) teams in India and Europe.

Industry Perspective Talks

Sunil Kumar Kopparapu obtained PhD from IIT Bombay and now is a Principal Scientist at TCS Research & Innovations Labs, Mumbai. He focuses on speech, script, and NLP, aiming to create practical solutions for Indian conditions. He is a senior member of IEEE and ACM.

Bidisha Sharma obtained PhD from IIT Guwahati and now she is a Senior AI Scientist at Uniphore, India. Her research interests lie in speech processing, encompassing areas like ASR, TTS synthesis, speech enhancement and end-to-end spoken language understanding and music processing. She is an active member of IEEE ARSU, SIGDIAL, IWSDS and COCOSDA.

Nirmesh J. Shah obtained PhD and MTech from DAU Gandhinagar and now he is working as a senior research scientist at Sony Research India since 2020. He is primarily leading speech related activities at Sony Research India. He has been serving as a reviewer for many IEEE Journals and top conferences, namely, ICASSP and INTERSPEECH.

Premjit Singh obtained PhD from IIT Kharagpur and now he is Lead Engineer at Samsung R&D Institute, Bengaluru, India. His notable publications are in several esteemed conferences such as INTERSPEECH, SPCM, and EUSIPCO, which involve using Modulation spectral features, constant-Q filterbank based representations, deep scattering network for SER.

Debmalya Chakraborty obtained PhD from John Hopkins University, USA and now he is a Senior Applied Scientist at Amazon AGI Alexa Machine Learning team with expertise in speech processing, auditory scene analysis, and machine learning for acoustic modeling. He has contributed significantly to improving Alexa's speech recognition frameworks.

Nagaraj Adiga obtained PhD from IIT Guwahati and now he is a Principal Research Scientist at Krutrim AI, building Multimodal and Multi-lingual LLM architectures. His work also includes Speech LLMs for multiple tasks such as speech understanding and generation tasks.

Dipesh K. Singh is a Speech Recognition Engineer at Aognito, India, specializing in Healthcare. His team achieved top performance in language diarization task as part of the DISPLACE challenge held during INTERSPEECH 2023, Dublin, Ireland. His research interests include speech processing for the healthcare, hidden ASR, and voice privacy.

Thosith S leads Speech AI at gnani.ai, Bengaluru, India, where he drives the development and deployment of cutting-edge voice technologies at scale. With over seven years of experience in the field, he has contributed across, the evolution of speech systems – from classic HMM-GMM and TDNN pipelines, through multi-lingual streaming RNNT and Conformer, to today's end-to-end voice-to-voice large language models and much more. His research interests include speech processing for the healthcare, hidden ASR, and voice privacy. The ASR system he has architected handle over 10 million voice interactions per day.

5 Minute PhD Thesis (5MPT) Contest

About 5MPT: The summer school at DAU announces the 6th edition of 5 minute PhD thesis (5MPT) contest after five such successful events in India during S4P 2024, S4P 2019, S4P 2018, S4P 2017, and S4P 2016.

Based on the reviews of the presentations by a panel of international experts, four scholars will be awarded cash prizes.

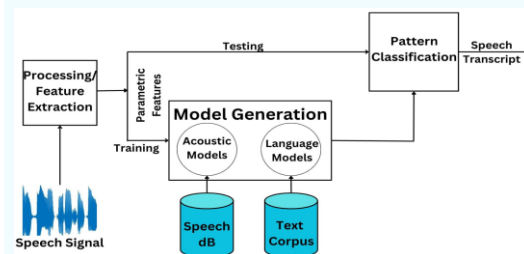
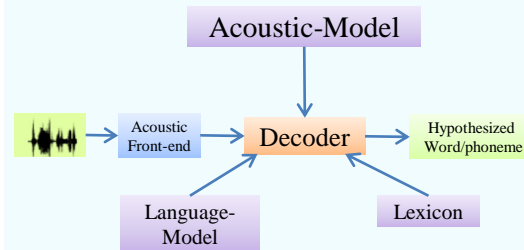
First prize: Rs. 15,000 /- **Second prize:** Rs. 10,000 /-

Third prize: Rs. 5,000 /- **Fourth prize:** Rs. 5,000 /-

Poster Presentation Contest

About PPC: The summer school at DAU announces the 2nd edition of Poster Presentation Contest (PPC) in S4P 2025. Based on the reviews of the presentations by a panel of international experts, four scholars will be awarded cash prizes of Rs. 5,000/-

Focused Areas



SCAN/Click to REGISTER:



SCAN/Click for PAYMENT:



Committees

Patron:
Tathagata Bandyopadhyay, Director General, DAU Gandhinagar

Organizing Committee:

- Hemant A. Patil, DAU Gandhinagar
- Eng-Siong Chng, Nanyang Technological University (NTU), Singapore
- Mathew Magimai Doss, IDIAP Switzerland
- Hardik B. Sailor, the Institute of Infocomm Research (I2R), Singapore
- Rodrigo Capobianco Guido, São Paulo State University (UNESP), Brazil

International Program Committee:

- Hermann Ney, RWTH Aachen, Germany
- Herve BOURLARD, EPFL, Honorary Professor and former Director of Idiap Research Institute, Switzerland
- Shrikanth Narayanan, University of Southern California, USA
- Tatsuya Kawahara, School of Informatics, Kyoto University, Japan
- Dilek Hakkani-Tur, University of Illinois Urbana-Champaign, USA
- Dong Yu, Tencent AI Lab, USA
- Akihiko Sugiyama, Founder, Damascus Corporation, Japan
- Lori Lamel, CNRS, France
- Elmar Nöth, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany
- Khalid Choukri, Secretary General, European Language Resources Association (ELRA), France
- Nicholas Evans, EURECOM, France
- Torbjorn Karl Svendsen, Norwegian University of Science and Technology, Norway
- Tiago H. Falk, MuSAE Lab, Montreal/Gatineau, Canada
- Kong Aik Lee, The Hong Kong Polytechnic University, Hong-Kong
- Nancy Zlatintsi, NTUA, Greece
- Hexin Liu, Nanyang Technological University, Singapore

DAU Hostel Accommodation:
Charges are *waived off* for participants.

DAU Student Grants:
50 DAU students (who registered for Summer School) will get a *grant of Rs. 1500/-* after attending this event.

Important Links

- S4P 2025 URL: <https://bit.ly/42NIMgv>
- Download the poster of S4P 2025: <https://bit.ly/3Exfj02>
- On-spot poster session: <http://bit.ly/4nsMNPo>
- For PhD Scholars: <https://bit.ly/4jB0tUL>
- Mobile Number: +91-75750-64449

Refund Policy Request for refund of registration fee must be made before 10th June, 2025 by writing an e-mail to s4p.daiict@gmail.com

Confirmed Sponsors



Technical Co-Sponsor

