IberSPEECH
2024
Aveiro, 11-13 November 2024

IBERSPEECH
Aveiro, 2024

IberSPEECH 2024, held in Aveiro, November 11-13th 2024, was organized by IEETA Research Unit (www.ieeta.pt) of the Universidade de Aveiro (www.ua.pt) and received inestimable support from the Spanish Thematic Network on Speech Technology (RTTH), the ISCA Special Interest Group on Iberian Languages (SIG-IL), the International Speech Communication Association (ISCA), Ministerio de Ciencia e Innovación (Spain) and the Cátedra RTVE, as well as the sponsors Veridas and Vicomtech.

The IberSPEECH 2024 event – the seventh of its kind using this name – brings together the XIII Jornadas en Tecnologías del Habla and the IX Iberian SLTech Workshop events. Both events pursue the aims of being a meeting point to present and discuss the results of the research on speech and language technologies on Iberian languages and promoting industry/university collaboration.

Aveiro is a picturesque city in north of Portugal, often referred to as the "Venice of Portugal" due to its charming network of canals and colorful moliceiro boats. Located along the country's central coast, Aveiro is known for its unique Art Nouveau architecture, vibrant culture, and rich history. The city is home to several prominent educational and research institutions, including the University of Aveiro, renowned for its innovative research, especially in technology and environmental sciences. Visitors to the Aveiro region can enjoy its beautiful beaches, traditional cuisine, and a variety of cultural events, making it a dynamic blend of tradition and modernity.

The venue, the Rectory of Universidade de Aveiro, in the Campus Universitário de Santiago, is very close to the city center.

Following the tradition of previous editions, IberSPEECH 2024 was a three-day event, bringing together the best research and practitioners in speech and language technology in Iberian languages to promote interaction and discussion. The organizing committee planned a wide variety of scientific, including technical paper presentations, keynote lectures, special sessions for the presentation of projects, recent PhD theses, entrepreneurship experiences and challenge evaluations, and a new entrepreneurship round table.

The core of the IberSPEECH 2024 scientific technical program includes a total of 46 full paper contributions that will be distributed among 5 oral and 2 posters sessions. To ensure the quality of all contributions, each submitted paper was reviewed by three members of the scientific review committee. All the papers in the conference will be accessible through the International Speech Communication Association (ISCA) Online Archive. Paper selection was based on the scores and comments provided by the committee, which includes over 70 researchers from different institutions (mainly from Spain and Portugal, but also from other European countries).

In addition to regular paper sessions, the IberSPEECH 2024 scientific program featured the following activities: the ALBAYZIN evaluation challenges; a special session including presentation of demos, research projects, recent PhD theses, entrepreneurship experiences, two keynote lectures; and a round table about entrepreneurship in the field of speech and language technologies.

Following the success of previous editions since 2006, the 2024 Albayzín Evaluations consist of three challenges: (1) Speech to Text Challenge (S2TC), organized by Radio Televisión

Española (RTVE) and University of Zaragoza, consisting in automatically transcribe different types of TV shows as in previous challenges; (2) Speaker Diarization and Identity Assignment Challenge (SDIAC), also organized by RTVE and University of Zaragoza, targeting segmentation of broadcast audio documents according to different speakers, linking those segments which originate from the same speaker and identify a closed set of speakers; (3) Bilingual Basque-Spanish ASR System Challenge (BBS-S2TC), organized by the University of the Basque Country (UPV/EHU), the task consisted on transcribing short segments of speech (ranging from 3 to 10 seconds) in Spanish, Basque, or both languages (with code switchings). This year, 10 teams have participated, providing different solutions to the proposed challenges.

Additionally, 16 special session papers were also included in the conference program. These were intended to describe either progress in current or recent research and ongoing projects, demonstration systems, or PhD Theses extended summaries to compete for the PhD Award.

The technical program also featured an Industrial Round Table, coordinated by Dr. Dayana Ribas (Special Sessions Chair), bringing together industry representatives to discuss their experiences and insights on advancing research to development in speech and language technologies.

Moreover, IberSPEECH 2024 featured two remarkable keynote speakers: Andreas Stolcke (Distinguished Scientist at Uniphore, USA), Prof. Juan Godino-Llorente (Professor at Universidad Politécnica de Madrid, Spain), to whom we would like to acknowledge for their extremely valuable participation.

The social program of IberSPEECH 2024 included a welcome reception and the gala dinner, held at *Quinta do Encontro*, a charming winery and restaurant nestled in the heart of Portugal's Bairrada wine region, offering a blend of contemporary architecture and traditional flavors, where guests can enjoy locally inspired dishes paired with exquisite wines crafted from the estate's vineyards.

Organizers would like to thank all those whose effort made possible this conference, including the members of the organizing committee, the local organizing committee, the Albayzin committee, the scientific review committee, the authors, the conference attendees, the supporting institutions and companies, and so many people who gave their best to achieve a successful conference.

Aveiro, November 2024
The General and Technical Program Chairs,
António Teixeira
Carlos Martinez-Hinarejos
Eduardo Lleida

Committees

General Chairs:

Antonio Teixeira, Universidade de Aveiro, Portugal
Carlos David Martínez Hinarejos, Universitat Politècnica de València, Spain

Technical Program Chair:

Eduardo Lleida, Universidad de Zaragoza, Spain

Technical Program Co-Chairs:

Carmen Garcia Mateo, U Vigo, Spain
Eva Navas, University of the Basque Country, UPV- EHU, Spain
José Andrés González López, Universidad de Granada, Spain
Francesc Alías Pujol, La Salle — Universitat Ramon LLull, Spain
Samuel Silva, Universidade de Aveiro, Portugal
Luis Fernando D'Haro, UPM, Spain

Special Sessions Chair:

Dayana Ribas, SONOC, BTS Group, Zaragoza, Spain

Evaluation Chairs:

Eduardo Lleida, Universidad de Zaragoza, Spain
Alfonso Ortega, Universidad de Zaragoza, Spain
Javier Tejedor Noguerales, Universidad San Pablo CEU, Spain
Luis Javier Rodríguez Fuentes, Universidad del País Vasco, Spain
Doroteo Torre Toledano, AUDIAS, Universidad Autónoma de Madrid

Local Committee:

Ana Rocha, Universidade de Aveiro, Portugal
Bernardo Marques, Universidade de Aveiro, Portugal
Catarina Oliveira, Universidade de Aveiro, Portugal
Fábio Barros, Universidade de Aveiro, Portugal
Gabriel Silva, Universidade de Aveiro, Portugal
Mário Rodrigues, Universidade de Aveiro, Portugal
Nuno Almeida, Universidade de Aveiro, Portugal
Samuel Silva, Universidade de Aveiro, Portugal

Scientific Review Committee

Alberto Abad	IST/INESC-ID
Luciana Albuquerque	University of Aveiro
Aitor Alvarez	Vicomtech
Francesc Alías Pujol	La Salle - Universitat Ramon Llull
Xavier Anguera Miro	ELSA Corp.
Jorge Baptista	U. Algarve & L2F-Spoken Language Lab,INESC ID Lisboa (Portugal)
Plinio Barbosa	University of Campinas
Jon Barker	The University of Sheffield
Fernando Batista	INESC-ID & ISCTE-IUL
Antonio Bonafonte	sanas.ai
Joao Cabral	TrinityCollege Dublin
Valentín Cardeñoso-Payo	Universidad de Valladolid
Maria Jose Castro-Bleda	Universitat Politècnica de València
Ricardo Cordoba	Grupo de Tecnologia del Habla
Conceicao Cunha	IPS Munich
Luis Fernando D'Haro	Speech Technology Group - UPM
Arantza Del Pozo	Vicomtech
Laura Docio-Fernandez	University of Vigo
David Escudero	University of Valladolid
Nicholas Evans	EURECOM
Mireia Farrús	Universitat de Barcelona
Liliana Ferreira	Fraunhofer Portugal AICOS
Javier Ferreiros	Universidad Politécnica de Madrid
Roberto Font	SDG Group
Ascension Gallardo-Antolin	Universidad Carlos III de Madrid
Fernando Garcia Granada	Universitat Politècnica de València
Carmen Garcia Mateo	University of Vigo
Kafentzis George	University of Crete
Omid Ghahabi	EML Speech Technology GmbH
Juan Ignacio Godino Llorente	Universidad Politécnica de Madrid
Angel Gomez	University of Granada
Jose Andres Gonzalez Lopez	University of Granada
César González-Ferreras	Universidad de Valladolid
David Griol	University of Granada
Inma Hernaez	Hitz Basque Center for Language Technology - University of the Basque Country (UPV/EHU)
Javier Hernando	Universitat Politècnica de Catalunya
Lluís-F. Hurtado	Universitat Politècnica de València
Eduardo Lleida	University of Zaragoza
Jordi Luque	Telefónica Research
Iván López-Espejo	University of Granada
Carlos David Martinez Hinarejos	Pattern Recognition and Human Language Technologies Research Center - Universitat Politècnica de València
Juan Manuel Martín-Doñas	Vicomtech

Antonio Miguel	University of Zaragoza
Climent Nadeu	Universitat Politècnica de Catalunya
Juan L. Navarro-Mesa	Universidad de Las Palmas de Gran Canaria
Eva Navas	University of the Basque Country
Hermann Ney	RWTH Aachen University
Catarina Oliveira	University of Aveiro
Alfonso Ortega	University of Zaragoza
Antonio M. Peinado	Dpto. Teoria de la señal, telematica y comunicaciones - University of Granada
Carmen Peláez-Moreno	University Carlos III Madrid
Jiri Pribil	Slovak Academy of Sciences
Jorge Proença	ELSA Corp.
José L. Pérez-Córdoba	University of Granada
Paulo Quaresma	Universidade de Evora
Ganna Raboshchuk	Google
Dayana Ribas	BTS - Business Telecommunications Services
Mário Rodrigues	ESTGA / IEETA - University of Aveiro
Luis Javier Rodriguez-Fuentes	University of the Basque Country UPV/EHU
Rubén San-Segundo	Grupo de Tecnologias del Habla. Dpto. Ingeniería Electrócnia. Universidad Politécnica de Madrid
Jon Sanchez	Aholab - EHU/UPV
Emilio Sanchis	Universitat Politècnica de València
Diana Santos	University of Oslo
Encarna Segarra	Departamento Sistemas Informáticos y Computación (Universitat Politècnica de València)
Carlos Segura Perales	Telefónica I+D
Samuel Silva	DETI / IEETA - Universidade de Aveiro
Alberto Simões	2Ai Lab - IPCA
Rubén Solera-Ureña	INESC-ID Lisboa
António Teixeira	DETI / IEETA, University of Aveiro
Javier Tejedor	Universidad CEU San Pablo
Doroteo Toledano	UAM
M. Inés Torres	Universidad del Pas Vasco
Isabel Trancoso	INESC ID Lisboa / IST
Andrej Zgank	University of Maribor
Catalin Zorila	Toshiba Cambridge Research Laboratory

Organizing Institutions

Organized by:

With the support of:

Ayuda RED2022-134270-T financiada por:

Sponsored by:

Venue

The IberSPEECH 2024 conference took place at the "**Sala de Atos**" at the Rectory Building of University of Aveiro, located on the Campus Universitario de Santiago (Aveiro).

The campus is located just a short distance from the city center, providing easy access to the vibrant local community, historic sites, and natural attractions like the Ria de Aveiro lagoon. Spread over a modern, spacious area, the campus combines innovative architectural design with green spaces, offering an inspiring environment for students and staff. This central location also means the university benefits from nearby transportation links, including the Aveiro train station, which connects the city to major hubs like Porto and Lisbon, making it an accessible destination for both domestic and international students.

Invited Speakers

Andreas Stolcke is a distinguished scientist at Uniphore. He obtained his PhD from UC Berkeley and then worked as a researcher/scientist at SRI International, Microsoft, and Amazon. His research interests include computational linguistics, language modeling, speech recognition, speaker recognition and diarization (keeping track of multiple speakers), and paralinguistics (e.g., sentiment and emotion recognition), with over 300 papers and patents in these areas. His open-source SRI Language Modeling Toolkit is widely used in academia. For over 30 years, Andreas has a strong track record inventing novel algorithms for speech and language processing. He is a Fellow of the IEEE, the International Speech Communication Association, and the Asia-Pacific Artificial Intelligence Association.

Juan Godino-Llorente was born in Madrid, Spain, in 1969. He received the B.Sc. and M.Sc. degrees in Telecommunications Eng., and the PhD. degree in Computer Science in 1992, 1996 and 2002, respectively, all from Universidad Politécnica de Madrid (UPM), Spain. From 1996 to 2003 he was with the UPM as Ass. Professor at the Circuits and Systems Eng. Dept. From 2003 to 2005 he joined the Signal Theory and Communications Dept. at the University of Alcala. From 2005, he joined again UPM, being the Head of the Circuits and Systems Eng. Dept. from 2006 till 2010. Since 2011 he is Full Professor in the field of Signal Theory and Communications. In 2006, he won the associate professor position after a national qualifying competitive call with 130 candidates, in which was ranked 1st.

During the academic term 2003-2004, he was a Visiting Professor at Salford University, Manchester, UK; and in 2016, he has been a Visiting Researcher at the Massachusetts Institute of Technology, USA funded by a Fulbright grant. He has served as editor for the IEEE Journal of Selected Topics in Signal Processing, for the IEEE Trans. on Audio, Speech and Language Processing, for the Speech Communication Journal, and for the EURASIP Journal of Advances in Signal Processing; and has also been a member of the scientific committee of INTERSPEECH, IEEE ICASSP, EUSIPCO, BIOSIGNALS, and other top ranked events for more than 10 years. He has participated as invited speaker in several international advanced schools and has delivered invited speeches at different universities and events, including Harvard University, Johns Hopkins University, Tampere University and National University of Colombia.

He has chaired the 3rd Advanced Voice Function Assessment Workshop, and of the 1st and 2nd Automatic Assessment of Parkinsonian Speech Workshop. Likewise, since 2004, he is part of different panels of experts of the European Commission, and has been national coordinator of COST Action 2103, funded by the European Science Foundation. He is also expert evaluator of research proposals for the Spanish, Portuguese, Latvian, Polish, Israeli, Czech, Icelandic, Romanian, Belgian, and Norwegian research agencies.

He has published more than 70 papers in international journals included in the Science Citation Index and more than 50 in top ranked conferences. The international impact of his research activities is supported by the large number of publications in international journals, which have attracted more than 5500 citations (h-index=40). He has also led 10 competitive projects and 12 research projects financed by companies and public institutions. In total, he has received more than 2,5 M€ of funding from competitive calls and from the industry for research purposes.

Program

Start time	Monday 11th	Tuesday 12th	Wednesday 13th
08:00/8:30	8:00 - Registration desk opens	8:30 - Registration desk opens	8:30 - Registration desk opens
9:00	9:00 - 9:20 Opening Ceremony	9:00 - 10:40 Oral Session 3 (5p) Language Technologies and Applications	9:00 - 10:40 Oral Session 5 (5p) Speech Enhancement, Processing, and Acoustic Event Detection
9:20	9:20 - 10:40 Oral Session 1 (4p) Speech Technology and Applications		
10:40	10:40 - 11:00 Coffee Break	10:40 - 11:00 Coffee Break	10:40 - 11:00 Coffee Break
11:00	11:00 - 12:00 Keynote 1	11:00 - 12:00 Keynote 2	11:00 - 11:30 Special Talk
12:00	12:00 - 13:00 Posters Session 1	12:00 - 13:00 Posters Session 2	11:30 - 12:30 Special Session: Projects, Demos & Theses
12:30			12:30 - 13:00 Closing Ceremony
13:00	13:00 -14:30 Lunch Break	13:30 -15:00 Lunch Break	
15:00	15:00 - 17:00 Oral Session 2 (6p) Voice and Speech Analysis for Diagnosis and Monitoring	15:00 - 17:00 Oral Session 4 (6p) Human Speech Production and Synthesis	
17:00	17:00 - 17:20 Coffee Break	17:00 - 17:20 Coffee Break	
17:20	17:20 - 18:20 Industrial Round Table	17:20 - 19:00 Albayzin Evaluations	
18:20			
19:00	18:20 - 19:30 RTTH Assembly	19:00 - BUS TO CONFERENCE DINNER	
19:30	19:30 - WELCOME RECEPTION	19:30 - CONFERENCE DINNER	