IWSLT 2024

The 21st International Conference on Spoken Language Translation

Proceedings of the Conference

August 15-16, 2024

The IWSLT 2024 organizers gratefully acknowledge the support from our Diamond sponsor Apple.

Diamond

©2024 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 979-8-89176-141-4

Introduction

The International Conference on Spoken Language Translation (IWSLT) is the premiere annual scientific conference for the study, development and evaluation of spoken language translation technology. Launched in 2004 and spun out from the C-STAR speech translation consortium before it (1992-2003), IWSLT is the main venue for scientific exchange on all topics related to speech-to-text translation, speech-to-speech translation, simultaneous and consecutive translation, speech dubbing, cross-lingual communication including all multimodal, emotional, paralinguistic, and stylistic aspects and their applications in the field. The conference organizes evaluations around challenge areas, and presents scientific papers and system descriptions. IWSLT is organized by the Special Interest Group on Spoken Language Translation (SIGSLT), which is supported by ACL, ISCA and ELRA.

This year, IWSLT featured spoken language translation shared tasks organized into seven distinct tracks: (i) speech-to-speech translation, (ii) simultaneous speech translation, (iii) subtitling, (iv) offline speech translation, (v) dubbing, (vi) low resource, and (vii) indic speech translation. Each track was coordinated by one or more chairs. The resulting evaluation campaigns attracted a total of 18 teams, from academia, research centers and industry. System submissions resulted in 26 system papers that will be presented at the conference. Following our call for papers, this year we received 10 submissions of research papers, 7 of which were accepted for oral presentation through a double-blind review process. The proceedings also include a survey paper summarizing recent research highlights, 2 test suite papers, which were peer-reviewed consistently with scientific and system papers respectively. In addition, the conference program is enriched by the presentation of 2 speech translation papers published in the Findings of the ACL over the past year.

The program committee is excited about the quality of the accepted papers and expects lively discussion and exchange at the conference. The conference chairs and organizers would like to express their gratitude to everyone who contributed and supported IWSLT. In particular, we wish to thank our Diamond sponsor Apple. We thank the shared tasks chairs, organizers, and participants, the program committee members, as well as all the authors that went the extra mile to submit system and research papers to IWSLT, and make this year's conference a big success. We also wish to express our sincere gratitude to ACL for hosting our conference and for arranging the logistics and infrastructure that allow us to hold IWSLT 2024 as a hybrid conference.

Welcome to IWSLT 2024, welcome to Bangkok!

Marine Carpuat, Program Chair
Marcello Federico and Alex Waibel, Conference Chairs

Organizing Committee

Conference Chairs

Marcello Federico, AWS AI Labs, USA
Alex Waibel, CMU, USA

Program Chair

Marine Carpuat, UMD, USA

Sponsorship Chair

Sebastian Stüker, Zoom, Germany

Evaluation Chair

Jan Niehues, KIT, Germany

Website and Publication Chair

Elizabeth Salesky, JHU, USA

Publicity Chair

Atul Kr. Ohja, University of Galway, Ireland

Program Committee

Program Committee

Milind Agarwal, GMU
Ibrahim Said Ahmad, Northeastern University
Md Mahfuz Ibn Alam, GMU
Antonios Anastasopoulos, GMU
Laurent Besacier, Naver Labs, France
Mauro Cettolo, FBK, Italy
Lizhong Chen, Oregon State University
William Chen, CMU
Josep Maria Crego, Systran, France
Anna Currey, AWS AI Labs
Qianqian Dong, ByteDance AI Lab, China
Akiko Eriguchi, Microsoft, USA
HyoJung Han, UMD
Benjamin Hsu, Amazon
Hirofumi Inaguma, Meta AI, USA
Takatomo Kano, NTT
Philipp Koehn, JHU
Surafel Melaku Lakew, Amazon AI, USA
Yves Lepage, Waseda University, Japan
Evgeny Matusov, AppTek, Germany
Chandresh Maurya, IIT Indore
Paul McNamee, JHU
Jan Niehues, KIT
Atul Kr. Ojha, U. Galway
John E. Ortega, Northeastern University
Sara Papi, FBK
Kumar Rishu, Charles University
Matthias Sperber, Apple, USA
Sebastian Stüker, Zoom, Germany
Katsuhito Sudoh, NAIST, Japan
Yun Tang, Meta AI, USA
Brian Thompson, AWS AI Labs, USA
Ioannis Tsiamas, Polytechnic University of Catalonia (UPC)
Marco Turchi, Zoom, Germany
David Vilar, Google, Germany
Yogesh Virkar, Amazon
Matthew Wiesner, JHU
Krzystof Wolk, Polish-Japanese Academy of Information Technology, Poland
Tong Xiao, Northeastern University
Rodolfo Zevallos, Universitat Pompeu Fabra

Table of Contents

FINDINGS OF THE IWSLT 2024 EVALUATION CAMPAIGN
Ibrahim Said Ahmad, Antonios Anastasopoulos, Ondřej Bojar, Claudia Borg, Marine Carpuat, Roldano Cattoni, Mauro Cettolo, William Chen, Qianqian Dong, Marcello Federico, Barry Haddow, Dávid Javorský, Mateusz Krubiński, Tsz Kim Lam, Xutai Ma, Prashant Mathur, Evgeny Matusov, Chandresh Kumar Maurya, John P. McCrae, Kenton Murray, Satoshi Nakamura, Matteo Negri, Jan Niehues, Xing Niu, Atul Kr. Ojha, John E. Ortega, Sara Papi, Peter Polák, Adam Pospíšil, Pavel Pecina, Elizabeth Salesky, Nivedita Sethiya, Balaram Sarkar, Jiatong Shi, Claytone Sikasote, Matthias Sperber, Sebastian Stüker, Katsuhito Sudoh, Brian Thompson, Alex Waibel, Shinji Watanabe, Patrick Wilken, Petr Zemánek and Rodolfo Zevallos 1

Pause-Aware Automatic Dubbing using LLM and Voice Cloning
Yuang Li, Jiaxin GUO, Min Zhang, Ma Miaomiao, Zhiqiang Rao, Weidong Zhang, Xianghui He, Daimeng Wei and Hao Yang 60

NICT's Cascaded and End-To-End Speech Translation Systems using Whisper and IndicTrans2 for the Indic Task
Raj Dabre and Haiyue Song 65

Transforming LLMs into Cross-modal and Cross-lingual Retrieval Systems
Frank Palma Gomez, Ramon Sanabria, Yun-hsuan Sung, Daniel Cer, Siddharth Dalmia and Gustavo Hernandez Abrego 71

Conditioning LLMs with Emotion in Neural Machine Translation
Charles Brazier and Jean-Luc Rouas 81

The NYA's Offline Speech Translation System for IWSLT 2024
Yingxin Zhang, Guodong Ma and Binbin Du 87

Improving the Quality of IWLST 2024 Cascade Offline Speech Translation and Speech-to-Speech Translation via Translation Hypothesis Ensembling with NMT models and Large Language Models
Zhanglin Wu, Jiaxin GUO, Daimeng Wei, Zhiqiang Rao, Zongyao Li, Hengchao Shang, Yuanchang Luo, Shaojun Li and Hao Yang 94

HW-TSC's Speech to Text Translation System for IWSLT 2024 in Indic track
bin wei, Zongyao Li, Jiaxin GUO, Daimeng Wei, Zhanglin Wu, Xiaoyu Chen, Zhiqiang Rao, Shaojun Li, Yuanchang Luo, Hengchao Shang, hao yang and yanfei jiang 101

Multi-Model System for Effective Subtitling Compression
Carol-Luca Gasan and Vasile Păiș 105

FBK@IWSLT Test Suites Task: Gender Bias evaluation with MuST-SHE
Beatrice Savoldi, Marco Gaido, Matteo Negri and Luisa Bentivogli 113

SimulSeamless: FBK at IWSLT 2024 Simultaneous Speech Translation
Sara Papi, Marco Gaido, Matteo Negri and Luisa Bentivogli 120

The SETU-DCU Submissions to IWSLT 2024 Low-Resource Speech-to-Text Translation Tasks
Maria Zafar, Antonio Castaldo, Prashanth Nayak, Rejwanul Haque, Neha Gajakos and Andy Way 128

Automatic Subtitling and Subtitle Compression: FBK at the IWSLT 2024 Subtitling track
Marco Gaido, Sara Papi, Mauro Cettolo, Roldano Cattoni, Andrea Piergentili, Matteo Negri and Luisa Bentivogli 134

UM IWSLT 2024 Low-Resource Speech Translation: Combining Maltese and North Levantine Arabic
Sara Nabhani, Aiden Williams, Miftahul Jannat, Kate Rebecca Belcher, Melanie Galea, Anna Taylor, Kurt Micallef and Claudia Borg 145

UOM-Constrained IWSLT 2024 Shared Task Submission - Maltese Speech Translation
Kurt Abela, Md Abdur Razzaq Riyadh, Melanie Galea, Alana Busuttil, Roman Kovalev, Aiden Williams and Claudia Borg 156

Compact Speech Translation Models via Discrete Speech Units Pretraining
Tsz Kin Lam, Alexandra Birch and Barry Haddow 162

QUESPA Submission for the IWSLT 2024 Dialectal and Low-resource Speech Translation Task
John E. Ortega, Rodolfo Joel Zevallos, Ibrahim Said Ahmad and William Chen 173

Speech Data from Radio Broadcasts for Low Resource Languages
Bismarck Bamfo Odoom, Leibny Paola Garcia Perera, Prangthip Hansanti, Loic Barrault, Christophe Ropers, Matthew Wiesner, Kenton Murray, Alexandre Mourachko and Philipp Koehn 182

JHU IWSLT 2024 Dialectal and Low-resource System Description
Nathaniel Romney Robinson, Kaiser Sun, Cihan Xiao, Niyati Bafna, Weiting Tan, Haoran Xu, Henry Li Xinyuan, Ankur Kejriwal, Sanjeev Khudanpur, Kenton Murray and Paul McNamee 188

CMU's IWSLT 2024 Simultaneous Speech Translation System
Xi Xu, Siqi Ouyang and Lei Li 202

HW-TSC's Submissions To the IWSLT2024 Low-resource Speech Translation Tasks
zheng jiawei, Hengchao Shang, Zongyao Li, Zhanglin Wu, Daimeng Wei, Zhiqiang Rao, Shaojun Li, Jiaxin GUO, bin wei, Yuanchang Luo and Hao Yang 208

CMU's IWSLT 2024 Offline Speech Translation System: A Cascaded Approach For Long-Form Robustness
Brian Yan, Patrick Fernandes, Jinchuan Tian, Siqi Ouyang, William Chen, Karen Livescu, Lei Li, Graham Neubig and Shinji Watanabe 212

NAIST Simultaneous Speech Translation System for IWSLT 2024
Yuka Ko, Ryo Fukuda, Yuta Nishikawa, Yasumasa Kano, Tomoya Yanagita, Kosuke Doi, Mana Makinae, Haotian Tan, Makoto Sakai, Sakriani Sakti, Katsuhito Sudoh and Satoshi Nakamura 218

Blending LLMs into Cascaded Speech Translation: KIT's Offline Speech Translation System for IWSLT 2024
Sai Koneru, Thai Binh Nguyen, Ngoc-Quan Pham, Danni Liu, Zhaolin Li, Alexander Waibel and Jan Niehues 231

ALADAN at IWSLT24 Low-resource Arabic Dialectal Speech Translation Task
waad ben kheder, Josef Jon, André Beyer, Abdel Messaoudi, Rabea Affan, Claude Barras, Maxim Tychonov and Jean-Luc Gauvain 240

Enhancing Translation Accuracy of Large Language Models through Continual Pre-Training on Parallel Data
Minato Kondo, Takehito Utsuro and Masaaki Nagata 251

The KIT Speech Translation Systems for IWSLT 2024 Dialectal and Low-resource Track
Zhaolin Li, Enes Yavuz Ugan, Danni Liu, Carlos Mullov, Tu Anh Dinh, Sai Koneru, Alexander Waibel and Jan Niehues 269

Empowering Low-Resource Language Translation: Methodologies for Bhojpuri-Hindi and Marathi-Hindi ASR and MT
Harpreet Singh Anand, Amulya Ratna Dash and Yashvardhan Sharma 277

Recent Highlights in Multilingual and Multimodal Speech Translation
Danni Liu and Jan Niehues 283

Word Order in English-Japanese Simultaneous Interpretation: Analyses and Evaluation using Chunk-wise Monotonic Translation
Kosuke Doi, Yuka Ko, Mana Makinae, Katsuhito Sudoh and Satoshi Nakamura 302

Leveraging Synthetic Audio Data for End-to-End Low-Resource Speech Translation
Yasmin Moslem 313

HW-TSC's Simultaneous Speech Translation System for IWSLT 2024
Shaojun Li, Zhiqiang Rao, bin wei, Yuanchang Luo, Zhanglin Wu, Zongyao Li, Hengchao Shang, Jiaxin GUO, Daimeng Wei and Hao Yang 322

UoM-DFKI submission to the low resource shared task
Kumar Rishu, Aiden Williams, Claudia Borg and Simon Ostermann 328

HW-TSC's submission to the IWSLT 2024 Subtitling track
Yuhao Xie, Yuanchang Luo, Zongyao Li, Zhanglin Wu, Xiaoyu Chen, Zhiqiang Rao, Shaojun Li, Hengchao Shang, Jiaxin GUO, Daimeng Wei and Hao Yang 334

Charles Locock, Lowcock or Lockhart? Offline Speech Translation: Test Suite for Named Entities
Maximilian Awiszus, Jan Niehues, Marco Turchi, Sebastian Stüker and Alex Waibel 339

Fixed and Adaptive Simultaneous Machine Translation Strategies Using Adapters
Abderrahmane Issam, Yusuf Can Semerci, Jan Scholtes and Gerasimos Spanakis 346

IWSLT 2024 Indic Track system description paper: Speech-to-Text Translation from English to multiple Low-Resource Indian Languages
Deepanjali Singh, Ayush Anand, Abhyuday Chaturvedi and Niyati Baliyan 359

Program

Thursday, August 15, 2024

09:00 - 09:15 *Welcome Remarks*

09:15 - 10:30 *Findings of the IWSLT 2024 Evaluation Campaign*

10:30 - 11:00 *Coffee Break*

11:00 - 12:00 *Findings of the IWSLT 2024 Evaluation Campaign*

12:00 - 12:30 *Oral Session A: Test Suite Papers (2)*

12:30 - 14:00 *Lunch Break*

14:00 - 15:30 *Poster Session I: System Papers for Low-Resource and Indic Tracks (13)*

15:30 - 16:00 *Coffee Break*

16:00 - 16:30 *Oral Session B: Scientific Papers (1) and Highlight Papers (1)*

16:30 - 17:30 *Panel Discussion*

Friday, August 16, 2024

09:00 - 10:30	*Oral Session C: Scientific Papers (6)*
10:30 - 11:00	*Coffee Break*
11:00 - 12:30	*Poster Session II: Findings Papers (1) and System Papers for Speech-to-Speech, Simultaneous, Subtitling, Offline & Dubbing tracks (12)*
12:30 - 14:00	*Lunch Break*
14:00 - 14:15	*Oral Session D: Findings Papers (1)*
14:15 - 15:30	*Planning Session for 2025*
15:30 - 16:00	*Coffee Break*
16:00 - 16:15	*Closing Remarks*