

# Foreword



The International Workshop on Spoken Language Translation (IWSLT) is an annual scientific workshop, associated with an open evaluation campaign on Spoken Language Translation, where both scientific papers and system descriptions are presented. The 11th International Workshop on Spoken Language Translation takes place in Lake Tahoe, USA on Dec. 04 and 05, 2014. Since 2004, the annual workshop has been held in Kyoto, Pittsburgh, Kyoto, Trento, Honolulu, Tokyo, Paris, San Francisco, Hong Kong, and Heidelberg, and this year in Lake Tahoe.

One of the prominent research activities in Spoken Language Translation is the work conducted by the Consortium for Speech Translation Advanced Research (C-STAR), which was an international partnership of research laboratories engaged in automatic translation of spoken language started in early 90s. The C-STAR members had initiated the first shared task-type Spoken Language Translation workshop in 2004 and the IWSLT has been growing up with more participants and steering committee members.

The IWSLT includes scientific papers in dedicated technical sessions, either in oral or poster form. The contributions cover theoretical and practical issues in the field of Machine Translation (MT) in general and Spoken Language Translation (SLT), including Automatic Speech Recognition (ASR), Text-to-Speech Synthesis (TTS), and MT, in particular:

- Speech and text MT
- Integration of ASR and MT
- MT and SLT approaches
- MT and SLT evaluation
- Language resources for MT and SLT
- Open source software for MT and SLT
- Adaptation in MT
- Simultaneous speech translation
- Speech translation of lectures
- Efficiency in MT
- Stream-based algorithms for MT
- Multilingual ASR and TTS
- Rich transcription of speech for MT
- Translation of on-verbal events

Submitted manuscripts were carefully peer-reviewed by three members of the program committee and papers were selected based on their technical merit and relevance to the conference. In addition to core statistical machine translation papers, the technical program covers a wide spectrum of topics related to Spoken Language Translation,

ranging from issues related to real-time interpretation or to the translation of dialogs to more practical issues related to the integration of speech and translation technologies. Several important new annotated corpora will also be presented during the workshop. In summary, the large number of submissions as well as the high quality of the submitted papers indicates the interest on Spoken Language Translation as a research field and the growing interest in these technologies and their practical applications.

The results of the Spoken Language Translation evaluation campaigns organized in the framework of the workshop are also an important part of IWSLT. Those evaluations are organized in the manner of competition. While participants compete for achieving the best result in the evaluation, they come together afterwards, and discuss and share their techniques that they used in their systems. In this respect, IWSLT proposes challenging research tasks and an open experimental infrastructure for the scientific community working on spoken and written language translation. This year, the IWSLT evaluation offered a very challenging and appealing task on the Spoken Language Translation of public speeches (TALK) in a variety of topics, including a dedicated task to automatic speech recognition in order to cover the full pipeline of speech translation.

For each task, monolingual and bilingual language resources, as needed, are provided to participants in order to train their systems, as well as sets of manual and automatic speech transcripts (with n-best and lattices) and reference translations, allowing researchers working only on written language translation to also participate. Moreover, blind test sets are released and all translation outputs produced by the participants are evaluated using several automatic translation quality metrics. For the primary submissions of all MT and SLT tasks, a human evaluation was carried out as well.

Each participant in the evaluation campaign has been requested to submit a paper describing the system and the utilized resources. A survey of the evaluation campaigns is presented by the organizers.

This time IWSLT 2014 is co-located with the 2014 IEEE Spoken Language Technology Workshop (SLT 2014). SLT will be held in South Lake Tahoe, Nevada, on Dec. 7-10, 2014. The main theme of the SLT workshop will be "machine learning in spoken language technologies". We expect that the co-location of IWSLT 2014 and SLT 2014 will attract more participants for further discussion on multi-lingual spoken language technologies.

Apart from the technical content of the conference, spectacular scenery of Lake Tahoe will welcome all participants to around-the-clock awesomeness on the shore of the largest alpine lake in North America.

Welcome to Lake Tahoe!

*Satoshi Nakamura, General Chair IWSLT 2014*