

CALL FOR PAPERS
IEEE Journal of Selected Topics in Signal Processing
Special Issue on *End-to-End Speech and Language Processing*

End-to-end (E2E) systems have achieved competitive results compared to conventional hybrid Hidden Markov-deep neural network model-based automatic speech recognition (ASR) systems. Such E2E systems are attractive because they do not require initial alignments between input acoustic features and output graphemes or words. Very deep convolutional networks and recurrent neural networks have also been very successful in ASR systems due to their added expressive power and better generalization. ASR is often not the end goal of real-world speech information processing systems. Instead, an important end goal is information retrieval, in particular keyword search (KWS), that involves retrieving speech documents containing a user-specified query from a large database. Conventional keyword search uses an ASR system as a front-end that converts the speech database into a finite-state transducer (FST) index containing a large number of likely word or sub-word sequences for each speech segment, along with associated confidence scores and time stamps. A user-specified text query is then composed with this FST index to find the putative locations of the keyword along with confidence scores. More recently, inspired by E2E approaches, ASR-free keyword search systems have been proposed with limited success. Machine learning methods have also been very successful in Question-Answering, parsing, language translation, analytics and deriving representations of morphological units, words or sentences. Challenges such as the Zero Resource Speech Challenge aim to construct systems that learn an end-to-end Spoken Dialog (SD) system, in an unknown language, from scratch, using only information available to a language learning infant (zero linguistic resources). The principal objective of the recently concluded IARPA Babel program was to develop a keyword search system that delivers high accuracy for any new language given very limited transcribed speech, noisy acoustic and channel conditions, and limited system build time of one to four weeks. This special issue will showcase the power of novel machine learning methods not only for ASR, but for keyword search and for the general processing of speech and language.

Topics of interest in the special issue include (but are not limited to):

- Novel end-to-end speech and language processing
- Query-by-example search
- Deep learning based acoustic and word representations
- Query-by-example search
- Question answering systems
- Multilingual dialogue systems
- Multilingual representation learning
- Low and zero resource speech processing
- Deep learning based ASR-free keyword search
- Deep learning based media retrieval
- Kernel methods applied to speech and language processing
- Acoustic unit discovery
- Computational challenges for deep end-to-end systems
- Adaptation strategies for end to end systems
- Noise robustness for low resource speech recognition systems
- Spoken language processing: speech to speech translation, speech retrieval, extraction, and summarization
- Machine learning methods applied to morphological, syntactic, and pragmatic analysis
- Computational semantics: document analysis, topic segmentation, categorization, and modeling
- Named entity recognition, tagging, chunking, and parsing
- Sentiment analysis, opinion mining, and social media analytics
- Deep learning in human computer interaction

Dates:

Manuscript submission: April 1, 2017

First review completed: June 1, 2017

Revised Manuscript Due: July 15, 2017

Second Review Completed: August 15, 2017

Final Manuscript Due: September 15, 2017

Publication: December, 2017

Guest Editors:

Nancy F. Chen, Institute for Infocomm Research (I2R), A*STAR, Singapore

Mary Harper, Army Research Laboratory, USA

Brian Kingsbury, IBM Watson, IBM T.J. Watson Research Center, USA

Kate Knill, Cambridge University, U.K.

Bhuvana Ramabhadran, IBM Watson, IBM T.J. Watson Research Center, USA