

AAACL-IJCNLP 2020

**The Second Workshop on Life-long Learning for Spoken  
Language Systems**

**Proceedings of the Workshop**

December 7, 2020  
Online

©2020 The 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](mailto:acl@aclweb.org)

ISBN 978-1-952148-97-2

## Introduction

The 2nd workshop on Life-long Learning for Spoken Language Systems covers challenges in a lifelong process of a natural language system, where new users or functionalities are added, and existing functionalities are modified. Improving such systems often require updating training data, and statistical models to account for data distribution shifts. This workshop will be co-located with ACL 2020 and will be held online.

This year we received six submissions for the main workshop track and accepted three of them. We would like to thank the Program Committee members who reviewed the papers and provided valuable feedback. We would also like to thank the workshop participants.

William M. Campbell, Alex Waibel, Dilek Hakkani-Tur, Timothy J. Hazen, Kevin Kilgour, Eunah Cho, Varun Kumar, Hadrien Glaude  
Workshop Organizers



**Organizers:**

William M. Campbell, Alexa AI, Amazon  
Alex Waibel, Carnegie Mellon University/Karlsruhe Institute of Technology  
Dilek Hakkani-Tur, Alexa AI, Amazon  
Timothy J. Hazen, Twitter  
Kevin Kilgour, Google Research  
Eunah Cho, Alexa AI, Amazon  
Varun Kumar, Alexa AI, Amazon  
Hadrien Glaude, Alexa AI, Amazon

**Program Committee:**

Jan Niehues, Maastricht University  
Elizabeth Salesky, CMU  
Christian Federmann, Microsoft  
Abhyuday Jagannatha, UMass Amherst  
Beat Gfeller, Google Research  
Ankur Gandhe, Alexa Speech  
Hassan Rom, Google Research  
Matthias Sperber, Apple  
John Lalor, University of Notre Dame  
Teresa Herrmann, Echobot Media Technologies GmbH/Karlsruhe Institute of Technology  
Varun Nagaraja, Alexa Speech  
Nick Ruiz, Interactions LLC  
Krishna C. Puvvada, Alexa Speech  
Jervis Pinto, Electronic Arts  
Harsha Sundar, Alexa Speech  
Stan Peshterliev, Facebook  
Abhishek Sethi, Google  
Aditya Siddhant, Google



## Table of Contents

|                                                                                                                                                                                                                    |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <i>Deep Active Learning for Sequence Labeling Based on Diversity and Uncertainty in Gradient</i><br>Yekyung Kim .....                                                                                              | 1  |
| <i>Supervised Adaptation of Sequence-to-Sequence Speech Recognition Systems using Batch-Weighting</i><br>Christian Huber, Juan Hussain, Tuan-Nam Nguyen, Kaihang Song, Sebastian Stüker and Alexander Waibel ..... | 9  |
| <i>Data Augmentation using Pre-trained Transformer Models</i><br>Varun Kumar, Ashutosh Choudhary and Eunah Cho .....                                                                                               | 18 |





## Conference Program

*Deep Active Learning for Sequence Labeling Based on Diversity and Uncertainty in Gradient*

Yekyung Kim

*Supervised Adaptation of Sequence-to-Sequence Speech Recognition Systems using Batch-Weighting*

Christian Huber, Juan Hussain, Tuan-Nam Nguyen, Kaihang Song, Sebastian Stüker and Alexander Waibel

*Data Augmentation using Pre-trained Transformer Models*

Varun Kumar, Ashutosh Choudhary and Eunah Cho

