

Raphael Schumann

rschuman@cl.uni-heidelberg.de

Education

- **Heidelberg University** **Heidelberg, Germany**
PhD Student in Computational Linguistics *12/2019 - current*
 - Supervised by Stefan Riezler
 - **University of Waterloo** **Waterloo, Canada**
Visiting Graduate Student at Computer Science Department *03/2019 - 08/2019*
 - Supervised by Olga Vechtomova and Lili Mou
 - **Heidelberg University** **Heidelberg, Germany**
Master's in Computational Linguistics *graduated 2019*
 - Thesis about unsupervised sentence summarization [2]
 - **Heidelberg University** **Heidelberg, Germany**
Bachelor's in Computational Linguistics *graduated 2015*
 - Minor in Computer Science
-

Experience

- **Nuance** **Montréal, Canada**
Research Intern *04/2018 – 10/2018*
 - Text summarization of doctor patient conversations in the health-care domain
 - Integrated autoencoder pretraining and language model fusion methods into existing code base
 - Ran extensive experiments to analyze impact of fused external language model
 - **Bosch** **Palo Alto, US**
Research Intern *07/2017 – 10/2017*
 - Natural Language Understanding for personal assistant dialog system
 - Built a joint sequence-to-sequence model for ASR error correction, intent detection and slot filling
 - Raised classification performance by ~3% for intent detection and ~1% for slot filling [4]
 - **Nuance** **Aachen, Germany**
Research Intern *04/2017 – 06/2017*
 - Clinical Language Understanding for extracting ICD-10 Codes from medical records
 - Implemented new layers and architectures for internal deep learning framework using C++ and CUDA
 - Improved classification performance of existing model by ~5% [5]
 - **Heidelberg Institute for Theoretical Studies** **Heidelberg, Germany**
Research Assistant *05/2016 – 03/2017*
 - Author Disambiguation project at the Natural Language Processing research group
 - Worked on a machine learning model to disambiguate author names in academic databases like DBLP
 - Combined traditional string similarity metrics like Levenshtein with character embedding methods
 - **IBM Research and Development** **Böblingen, Germany**
Machine Learning Intern *10/2015 – 04/2016*
 - Improved a web-based bug tracker tool for hardware verification by adding machine learning features
 - Bug owner classification with features like error signature and recent code interactions
 - Reoccurring bugs are identified and clustered to quickly access past solutions
-

Publications

1. **Raphael Schumann** and Stefan Riezler
Generating Landmark Navigation Instructions from Maps as a Graph-to-Text Problem
In Proceedings of ACL 2021, Bangkok, Thailand (virtual).
 2. **Raphael Schumann**, Lili Mou, Yao Lu, Olga Vechtomova and Katja Markert
Discrete Optimization for Unsupervised Sentence Summarization with Word-Level Extraction
In Proceedings of ACL 2020, Seattle, US (virtual).
 3. **Raphael Schumann**, Ines Rehbein
Active Learning via Membership Query Synthesis for Semi-supervised Sentence Classification
In Proceedings of CoNLL 2019, Hong Kong.
 4. **Raphael Schumann**, Pongtep Angkititrakul
Incorporating ASR Errors with Attention-based, Jointly Trained RNN for Intent Detection and Slot Filling
In Proceedings of ICASSP 2018, Calgary, Canada.
 5. Marilisa Amoia, Frank Diehl, Jesús Giménez, Joel Pinto, **Raphael Schumann**, Fabian Stemmer, Paul Vozila, Yi Zhang
Scalable Wide and Deep Learning for Computer Assisted Coding
In Proceedings of NAACL-HLT 2018: Industry Track, New Orleans, US.
-

Teaching

- Grounded Natural Language Processing for Spatial Navigation
 - Instructor, undergraduate/graduate course, summer 2020, Heidelberg University
-

Talks

- Generating Landmark Navigation Instructions from Maps as a Graph-to-Text Problem
 - **Google AI**, NLP group Zurich virtual, November 2020
 - Active Learning via Membership Query Synthesis for Semi-supervised Sentence Classification
 - **CoNLL 2019**, oral presentation Hong Kong, November 2019
 - Transcription Errors in Context of Intent Detection and Slot Filling
 - **Carnegie Mellon University**, LTI Sphinx Lunch Pittsburgh, November 2018
 - **M*Modal (now 3M)** Pittsburgh, November 2018
 - Incorporating ASR Errors with Attention-based, Jointly Trained RNN for Intent Detection and Slot Filling
 - **ICASSP 2018**, oral presentation Calgary, April 2018
-

Natural Languages

- German (native), English