Zheng Zhang

247 Fitzpatrick Hall, Notre Dame, IN, 46556
zzhang37@nd.edu (651) 276-1867 http://zhengzhang.me

Education

Ph.D. in Computer Science and Engineering 2021–Present
University of Notre Dame *Notre Dame, IN*
Advisor: Toby Jia-Jun Li

Ph.D. in Computer Science (Transfer out) 2019–2021
University of Rochester *Rochester, NY*
Advisor: Zhen Bai

M.S. in Computer Science and Engineering 2017–2019
University of Minnesota, Twin Cities *Minneapolis, MN*
Advisor: Haiyi Zhu

B.Eng. in Software Engineering 2013–2017
Shaanxi Normal University *Xi'an, China*

Industrial Experience

Apple Inc. Seattle, WA
AI/ML research intern May 2022–Aug 2022
Mentor: Alistair Conkie, Ladan Golipour
Project: Zero-shot learning for foreign accent conversion

- Work with Siri text-to-speech team to build zero-shot learning pipeline converting speaker's foreign English accent to native accent without further fine-tuning upon specific foreign speakers
- Develop and train accent and speaker encoder with Speech Accent Archive and Voxceleb1 dataset respectively
- Build and train foreign accent conversion models based on Tacotron and FastSpeech that converts the concatenation of linguistic feature vector, accent and speaker embedding into mel-spectrogram

Research Projects

PaTAT: Human-AI collaborative qualitative coding with explainable interactive rule synthesis
Leading student researcher Nov 2021–Sep 2022

- Developed pattern synthesizer based on users' open coding
- Developed interactive qualitative coding system that is able to learn coding pattern and recommend coding for unannotated data
- Conducted lab study to evaluate the usability and effectiveness of PaTAT system
- Wrote a paper submitted to CHI'23

Label Sleuth: Human-AI collaborative textual data annotation platform
Leading student researcher/developer Apr 2022–Aug 2022

- Led the development of Label Sleuth: a platform enabling data owner and model developer to dispatch textual data annotation tasks to non-expert annotators while training and evaluating classification model via active learning
- Closely worked with design and product team at IBM Research to deliver the Label Sleuth platform

Error taxonomy, detection, and repair of NL2SQL model
Leading student researcher Jun 2021–Sep 2022

- Developed an SQL error labeling system and coding the error types of the state-of-the-art NL2SQL models
- Developed user interface that supports a range of NL2SQL error detection and repair methods
- Wrote a paper to be submitted to CHI 2023 as the co-first author

PEANUT: Human-in-the-loop system for audio-visual data annotation
Leading student researcher Jul 2021–Apr 2022

- Developed a human-AI collaborative audio-visual annotation system in which AI can automatically complete the annotation of sounding objects in video frames based on user's guidance
- Conducted conditional user study to test the usability, efficiency and effectiveness of the annotation system
- Wrote a paper to be submitted to CHI 2023 as the first author

StoryBuddy: interactive storytelling system with automatic question generation
Leading student researcher Jun 2021–Sep 2021
- Developed a storytelling system that supports automatic story question generation and flexible parent involvement
- Conducted user study to test the usability and effectiveness of the storytelling system
- Published a full paper as first author at CHI 2022

HAICOR: Human-in-the-loop tool for enhancing the rationality of AI commonsense reasoning
Leading student researcher Sep 2019–May 2021
- Designed and developed a platform facilitating lay users' debugging rationality of AI commonsense reasoning with intelligent argumentative scaffolding
- Developed GPT-2 based knowledge-aware commonsense reasoning model for naïve psychological need prediction on ROCStories dataset

Publications

(* means equal contribution)

Zheng Ning*, **Zheng Zhang***, Tianyi Sun, Yuan Tian, Tianyi Zhang, Toby Jia-Jun Li. An Empirical Study of Model Errors and User Error Discovery and Repair Strategies in Natural Language Database Queries
IUI 2023

Eyal Shnarch, Alon Halfon, Ariel Gera, Marina Danilevsky, Yannis Katsis, Leshem Choshen, Martin Santillan Cooper, Dina Epelboim, **Zheng Zhang**, Dakuo Wang, Lucy Yip, Liat Ein-Dor, Lena Dankin, Ilya Shnayderman, Ranit Aharonov, Yunyao Li, Naftali Liberman, Philip Levin Slesarev, Gwilym Newton, Shila Ofek-Koifman, Noam Slonim, Yoav Katz. Label Sleuth: From Unlabeled Text to a Classifier in a Few Hours
EMNLP 2022 (System Demo Track)

Ying Xu, Dakuo Wang, Mo Yu, Daniel Ritchie, Bingsheng Yao, Tongshuang Wu, **Zheng Zhang**, Toby Jia-Jun Li, Nora Bradford, Branda Sun, Tran Hoang, Yisi Sang, Yufang Hou, Xiaojuan Ma, Diyi Yang, Nanyun Peng, Zhou Yu, Mark Warschauer. Fantastic Questions and Where to Find Them: FairytaleQA–An Authentic Dataset for Narrative Comprehension. 2022 Annual Meeting of the Association for Computational Linguistics
ACL 2022

Bingsheng Yao, Dakuo Wang, Tongshuang Wu, **Zheng Zhang**, Toby Jia-Jun Li, Mo Yu, Ying Xu. It is AI's Turn to Ask Human a Question: Question and Answer Pair Generation for Children Storybooks in FairytaleQA Dataset. 2022 Annual Meeting of the Association for Computational Linguistics
ACL 2022

Zheng Zhang, Ying Xu, Yanhao Wang, Bingsheng Yao, Daniel Ritchie, Tongshuang Wu, Mo Yu, Dakuo Wang, Toby Jia-Jun Li. StoryBuddy: A Human-AI Collaborative Agent for Parent-Child Interactive Storytelling with Flexible Parent Involvement. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems
CHI 2022

Weiwen Leung*, **Zheng Zhang***, Daviti Jibuti, Jinhao Zhao, Maximilian A Klein, Casey Pierce, Lionel Robert, Haiyi Zhu. Race, Gender and Beauty: The Effect of Information Provision Affects Online Hiring Biases. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems
CHI 2020

Hao Fei Cheng, Ruotong Wang, **Zheng Zhang**, Fiona O'Connell, Terrance Gray, Max Harper, Haiyi Zhu. Explaining Decision-Making Algorithms through UI: Strategies to Help Non-Expert Stakeholders. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems
CHI 2019

Skills

Research expertise: Human-AI collaboration, Natural language processing, Speech processing
Machine learning: Pytorch, Tensorflow, SciPy, SpaCy
Web development: React, Vue.js, Flask, MongoDB, SQL
Programming language: Python, Java, C, C++, Javascript
HCI: Qualitative and quantitative research methods (e.g. survey, interview, statistical testing, field study, etc)