

# Fangzheng (Frank) Xu

5568 Forbes Ave, Pittsburgh, PA, USA 15217

☎ (+1) 412-482-8915 / (+86) 188-1827-5852 | ✉ frankxu2004@gmail.com | 🌐 frankxfz.me

## Research Interests

Natural language processing, machine learning, computational social science, data mining, human-computer interaction.

## Education

### Language Technology Institute, Carnegie Mellon University

PH.D. IN COMPUTER SCIENCE

- Advisor: Prof. Graham Neubig

Pittsburgh, Pennsylvania

Sep. 2019 - Present

### Shanghai Jiao Tong University

M.Sc. IN COMPUTER SCIENCE

- Thesis: *Indirect Supervision for Information Extraction*
- Research Field: Natural language processing, text mining, knowledge graph, spatial-temporal data mining
- Advisor: Prof. Kenny Q. Zhu

Shanghai, China

Sep. 2016 - Mar. 2019

### Shanghai Jiao Tong University

B.ENG. IN COMPUTER SCIENCE, HONOR DEGREE (IEEE HONORED CLASS)

- Thesis: *Traffic Prediction for Urban Planning*
- **Best Bachelor's Thesis Award (Top 1% in SJTU & Top 1 in the CSE Department)**
- Advisor: Prof. Kenny Q. Zhu

Shanghai, China

Sep. 2012 - Jun. 2016

## Honors & Awards

- |         |                                                                             |              |
|---------|-----------------------------------------------------------------------------|--------------|
| 2018    | Best Poster Award Honorable Mention, WWW                                    | Lyon, France |
| 2017    | 2nd in emerging and rare entity recognition shared task, WNUT@EMNLP         | Copenhagen   |
| 2016-18 | SJTU Graduate Student Scholarship                                           | SJTU         |
| 2016    | Best Bachelor's Thesis Award (Top 1% in SJTU & Top 1 in the CSE Department) | SJTU         |
| 2015    | Excellence Award in <i>DataBang</i> Data Innovation Contest                 | Shanghai     |
| 2015    | 2nd in EMC Smart Campus Open Data Contest                                   | Shanghai     |

## Academic & Professional Experience

### Microsoft Research Asia

RESEARCH INTERN

- Mentor: Dr. Lei Ji, Dr. Nan Duan
- Working on structured multimodal procedural knowledge acquisition and instruction summarization from narrative videos and texts.

Beijing, China

May. 2019 - Aug. 2019

### HCII, Carnegie Mellon University

VISITOR

- Advisor: Prof. Jeffrey P. Bigham
- Worked on AI and human collaboration, enabling human by incorporating multimodal ML techniques, including NLP and CV into daily life. Help blind people perceive the world, construct knowledge and understand natural language queries with state-action graphs of user interface, help memorize and organize events from users' daily conversations, etc.

Pittsburgh, PA

Nov. 2018 - May. 2019

### SAILING Group, Carnegie Mellon University

VISITOR

- Advisor: Zhiting Hu, Prof. Eric Xing
- Worked on deep generative models for natural language. Extend data-to-text generation to text content manipulation. Instead of generating from scratch, we generate by "modifying" other sentences retrieved through domain knowledge with the critical information in data records. The task differs from traditional text style transfer and we propose a dual-copy mechanism.

Pittsburgh, PA

Mar. 2018 - May. 2018

## VISITOR

Jun. 2017 - Dec. 2017

- Advisor: Prof. Jiawei Han (UIUC), Prof. Xiang Ren (USC)
- Worked on various text problems such as relation extraction, sequence labeling, event extraction, information extraction, social network, etc.
- Research papers accepted by WSDM 2018, AAAI 2018, KBCOM 2018, WWW 2018.

**ADAPT Lab, Shanghai Jiao Tong University**

Shanghai, China

## RESEARCH ASSISTANT

Oct. 2015 - Mar. 2019

- Advisor: Prof. Kenny Q. Zhu
- Working on natural language processing, computational social science, information extraction, text mining, with papers published in various venues. I used to work on data mining, traffic prediction and urban computing as well.

**Baixing.com**

Shanghai, China

## SOFTWARE DEVELOPER INTERN

Aug. 2015 - May. 2016

- Developed open APIs, crawlers, backend code with mainly Python and PHP.
- Designed ML algorithm modules to prevent spam and inappropriate posts, e.g. posts filtering and classification, image identification.

**IIOT Lab, Shanghai Jiao Tong University**

Shanghai, China

## UNDERGRADUATE RESEARCH ASSISTANT

Jan. 2015 - Jul. 2015

- Advisor: Prof. Xinbing Wang
- Contributed to academic crawler, paper topic modeling, search algorithms and infrastructure for *AceMap*, an academic search system.

**OMNILAB, Shanghai Jiao Tong University**

Shanghai, China

## UNDERGRADUATE RESEARCH ASSISTANT

Jun. 2014 - Dec. 2014

- Advisor: Prof. Yaohui Jin
- Worked on open data platform for urban air quality, human mobility and social media, ranging from web crawlers to data analysis and visualizations.

## Teaching Experience

2016 **Teaching Assistant**, Database System Technology (SE305), in English

SJTU

## Publication

\* indicates equal contribution

**PREPRINTS****Toward Unsupervised Text Content Manipulation**Wentao Wang, Zhiting Hu, Zichao Yang, Haoran Shi, [Frank F. Xu](#), Eric Xing.*arXiv: 1901.09501***CONFERENCE PAPERS****AlpacaTag: An Active Learning-based Crowd Annotation Framework for Sequence Tagging**Bill Y. Lin\*, Dong-Ho Lee\*, [Frank F. Xu](#), Ouyu Lan, Xiang Ren.*In Proceedings of 57th Annual Meeting of the Association for Computational Linguistics (ACL Demo)*, 2019**Mining Cross-Cultural Differences and Similarities in Social Media**[Frank F. Xu](#)\*, Bill Y. Lin\*, Kenny Q. Zhu, Seung-won Hwang.*In Proceedings of 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2018**Automatic Extraction of Commonsense Located Near Knowledge**[Frank F. Xu](#)\*, Bill Y. Lin\*, Kenny Q. Zhu.*In Proceedings of 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2018**ExTRA: Extracting Prominent Review Aspects from Customer Feedback**Zhiyi Luo, Shanshan Huang, [Frank F. Xu](#), Bill Yuchen Lin, Hanyuan Shi, Kenny Zhu.*In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018**Empower Sequence Labeling with Task-Aware Neural Language Model**Liyuan Liu, Jingbo Shang, [Frank F. Xu](#), Xiang Ren, Huan Gui, Jian Peng, Jiawei Han.*In Proceedings of Thirty-second AAAI Conference on Artificial Intelligence (AAAI)*, 2018**ReQuest: Indirect Supervision for Relation Extraction using Question-Answer Pairs**Zeqiu Wu, Xiang Ren, [Frank F. Xu](#), Ji Li, Jiawei Han.*In Proceedings of The 11th ACM International Conference on Web Search and Data Mining (WSDM)*, 2018

### **Mining Cross-Cultural Differences of Named Entities: A Preliminary Study**

Frank F. Xu\*, Bill Y. Lin\*, Kenny Q. Zhu.

*Presented on First Workshop on Knowledge Base Construction, Reasoning and Mining (KBCOM) at WSDM, 2018*

### **Entropy-Based Subword Mining with an Application to Word Embeddings**

Ahmed El-Kishky, Frank F. Xu, Aston Zhang, Stephen Macke, Jiawei Han.

*In Proceedings of the Second Workshop on Subword/Character Level Models (SCLem) at NAACL, 2018*

### **Open Information Extraction with Global Structure Constraints**

Qi Zhu, Xiang Ren, Jingbo Shang, Yu Zhang, Frank F. Xu, Jiawei Han.

*In Proceedings of The Web Conference (WWW Poster), 2018 (Best Poster Award Honorable Mention)*

### **Commonsense LocatedNear Relation Extraction**

Frank F. Xu\*, Bill Y. Lin\*, Kenny Q. Zhu.

*Presented on 6th Workshop on Automated Knowledge Base Construction (AKBC) at NIPS, 2017*

### **Multi-channel BiLSTM-CRF Model for Emerging Named Entity Recognition in Social Media**

Frank F. Xu\*, Bill Y. Lin\*, Zhiyi Luo, Kenny Q. Zhu.

*In Proceedings of the 3rd Workshop on Noisy User-generated Text (W-NUT) at EMNLP, 2017 (2nd place in the shared task)*

### **Transfer Learning for Traffic Speed Prediction: A Preliminary Study**

Frank F. Xu\*, Bill Y. Lin\*, Eve Q. Liao, Kenny Q. Zhu.

*In Proceedings of SmartIoT 2018 Workshop: AI enhanced IoT data processing for Intelligent Applications at AAAI, 2018*

### **Cross-region Traffic Prediction for China on OpenStreetMap.**

Frank F. Xu, Bill Y. Lin, Qi Lu, Yifei Huang and Kenny Q. Zhu.

*In Proceedings of the 9th ACM SIGSPATIAL International Workshop on Computational Transportation Science, 2016*

## **THESIS**

### **Indirect Supervision for Information Extraction**

Frank F. Xu

*Master's Thesis, Shanghai Jiao Tong University, 2019*

### **Traffic Prediction for Urban Planning**

Frank F. Xu

*Bachelor's Thesis (Best Thesis Award, Top 1 in CSE), Shanghai Jiao Tong University, 2016*

## **Professional Service**

---

### **PROGRAM COMMITTEE**

First Workshop on Knowledge Base Construction, Reasoning and Mining at WSDM 2018 (KBCOM)

### **CONFERENCE REVIEWER**

AAAI Conference on Artificial Intelligence (AAAI 2016, 17)

ACM International Conference on Information and Knowledge Management (CIKM 2017)

ACM International Conference on Web Search and Data Mining (WSDM 2018)

The Web Conference (WWW 2018)

ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2018)

The Annual Meeting of the Association for Computational Linguistics (ACL 2018, NAACL 2019, ACL 2019)

## **Reference**

---

### **Prof. Kenny Q. Zhu**

Professor of Computer Science

Shanghai Jiao Tong University, Shanghai

Homepage: <http://cs.sjtu.edu.cn/~kzhu/>

Email: [kzhu@cs.sjtu.edu.cn](mailto:kzhu@cs.sjtu.edu.cn)

### **Prof. Jiawei Han**

Abel Bliss Professor of Computer Science

University of Illinois at Urbana-Champaign

Homepage: <http://hanj.cs.illinois.edu/>

Email: [hanj@cs.uiuc.edu](mailto:hanj@cs.uiuc.edu)

**Prof. Xiang Ren**

Assistant Professor of Computer Science  
University of Southern California and Information Science Institute  
Homepage: <http://www-bcf.usc.edu/~xiangren/>  
Email: [xiangren@usc.edu](mailto:xiangren@usc.edu)

**Prof. Jeffrey P. Bigham**

Associate Professor of Computer Science  
Carnegie Mellon University  
Homepage: <https://www.cs.cmu.edu/~jbigham/>  
Email: [jbigham@cs.cmu.edu](mailto:jbigham@cs.cmu.edu)