

Fangzheng (Frank) Xu

5000 Forbes Ave, Pittsburgh, PA, USA 15213

☎ (+1) 412-482-8915 | ✉ fangzheng@cs.cmu.edu | 🌐 frankxfz.me

Research Interests

Large language models, agents, code generation, retrieval-augmented models.

Education

Language Technology Institute, Carnegie Mellon University

Pittsburgh, Pennsylvania

PH.D. IN COMPUTER SCIENCE

Sep. 2019 - Present

- Advisor: Prof. Graham Neubig
- Research Interest: Large language models for both code generation and agents. The eventual goal is to enable users to use natural language to instruct and interact with the computers in the future. I put special emphasis on augmenting models with external knowledge, e.g. retrieval-augmented generation, non-parametric models.
- **Supported by IBM Ph.D. Fellowship since Fall 2022.**

Shanghai Jiao Tong University

Shanghai, China

M.SC. IN COMPUTER SCIENCE

Sep. 2016 - Mar. 2019

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

Shanghai Jiao Tong University

Shanghai, China

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

Sep. 2012 - Jun. 2016

- 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

Honors & Awards

2022 IBM Ph.D. Fellowship

2018 Best Poster Award Honorable Mention, WWW

2017 2nd in emerging and rare entity recognition shared task, WNUT@EMNLP

2016 Best Bachelor's Thesis Award (Top 1% in SJTU & Top 1 in the CSE Department)

Lyons, France

Copenhagen

SJTU

Academic & Professional Experience

NeuLab, Carnegie Mellon University

Pittsburgh, PA

RESEARCH ASSISTANT

Aug. 2019 - Present

- Advisor: Prof. Graham Neubig
- Working on natural language processing, LLMs, code generation, retrieval-augmented models and web agents.

Google Research

Boston, MA

STUDENT RESEARCHER

May. 2022 - Aug. 2022

- Mentor: Nicholas Monath, Andrew McCallum, Manzil Zaheer
- Working on instance-based machine learning models for sequence modeling.

Microsoft Research Redmond

Virtual

RESEARCH INTERN

Jun. 2021 - Aug. 2021

- Mentor: Christian Bird, Shuvendu Lahiri, Todd Mytkowicz, Alexey Svyatkovskiy
- Working on automatic merge of code version control conflicts with reinforcement learning.

Microsoft Research Asia

Beijing, China

RESEARCH INTERN

May. 2019 - Aug. 2019

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

HCII, Carnegie Mellon University

Pittsburgh, PA

VISITOR

Nov. 2018 - May. 2019

- 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.

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.

Teaching Experience

Sp. 2024 **Guest Lecture**, Advanced NLP (11-711)

CMU

Fall 2021 **Teaching Assistant**, Advanced NLP (11-711)

CMU

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

SJTU

Publication

* indicates equal contribution

PAPERS

OpenDevin: An Open Platform for AI Software Developers as Generalist Agents

Web Agent Lead of OpenDevin Team.

arXiv, 2024

CodeRAG-Bench: Can Retrieval Augment Code Generation?

Zora Zhiruo Wang, Akari Asai, Xinyan Velocity Yu, [Frank F. Xu](#), Yiqing Xie, Graham Neubig, Daniel Fried.

arXiv, 2024

WebArena: A Realistic Web Environment for Building Autonomous Agents

[Frank F. Xu](#)*, Shuyan Zhou*, Hao Zhu, Xuhui Zhou, Robert Lo, Abishek Sridhar, Xianyi Cheng, Yonatan Bisk, Daniel Fried, Uri Alon, Graham Neubig.

In International Conference on Learning Representations (ICLR), 2024

Why do Nearest Neighbor Language Models Work?

[Frank F. Xu](#), Uri Alon, Graham Neubig.

In International Conference on Machine Learning (ICML), 2023

Hierarchical Prompting Assists Large Language Model on Web Navigation

Abishek Sridhar, Robert Lo, [Frank F. Xu](#), Hao Zhu, Shuyan Zhou.

In Findings of the Association for Computational Linguistics: EMNLP, 2023

Active Retrieval Augmented Generation

[Frank F. Xu](#)*, Zhengbao Jiang*, Luyu Gao*, Zhiqing Sun*, Qian Liu, Jane Dwivedi-Yu, Yiming Yang, Jamie Callan, Graham Neubig.

In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023

A Systematic Evaluation of Large Language Models of Code

[Frank F. Xu](#), Uri Alon, Graham Neubig, Vincent J Hellendoorn.

In Deep Learning For Code (DL4C) Workshop @ ICLR, 2022

Also in *The 6th Annual Symposium on Machine Programming (MAPS) @ PLDI*, 2022

DocCoder: Generating Code by Retrieving and Reading Docs

Shuyan Zhou, Uri Alon, [Frank F. Xu](#), Zhengbao Jiang, Graham Neubig.

In International Conference on Learning Representations (ICLR), 2023

MCoNaLa: A Benchmark for Code Generation from Multiple Natural Languages

Zhiruo Wang, Grace Cuenca, Shuyan Zhou, [Frank F. Xu](#), Graham Neubig.

In The 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2023

Neuro-Symbolic Language Modeling with Automaton-augmented Retrieval

Uri Alon, [Frank F. Xu](#), Junxian He, Sudipta Sengupta, Dan Roth, Graham Neubig.

In International Conference on Machine Learning (ICML), 2022

Capturing Structural Locality in Non-parametric Language Models

Frank F. Xu, Junxian He, Graham Neubig, Vincent J. Hellendoorn.

In International Conference on Learning Representations (ICLR), 2022

In-IDE Code Generation from Natural Language: Promise and Challenges

Frank F. Xu, Bogdan Vasilescu, Graham Neubig.

In ACM Transactions on Software Engineering and Methodology (TOSEM), 2021

Learning Structural Edits via Incremental Tree Transformations

Ziyu Yao, Frank F. Xu, Pengcheng Yin, Huan Sun, Graham Neubig.

In International Conference on Learning Representations (ICLR), 2021

Data-to-Text Generation with Style Imitation

Shuai Lin, Wentao Wang, Zichao Yang, Xiaodan Liang, Frank F. Xu, Eric Xing, Zhiting Hu.

In Findings of the Association for Computational Linguistics: EMNLP, 2020

A Benchmark for Structured Procedural Knowledge Extraction from Cooking Videos

Frank F. Xu, Lei Ji, Botian Shi, Junyi Du, Graham Neubig, Yonatan Bisk, Nan Duan.

In Proceedings of the First International Workshop on Natural Language Processing Beyond Text @ EMNLP, 2020

Minimally Supervised Categorization of Text with Metadata

Yu Zhang, Yu Meng, Jiaxin Huang, Frank F. Xu, Xuan Wang, Jiawei Han.

In Proceedings of 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2020

Incorporating External Knowledge through Pre-training for Natural Language to Code Generation

Frank F. Xu*, Zhengbao Jiang*, Pengcheng Yin, Bogdan Vasilescu, Graham Neubig.

In Proceedings of 58th Annual Meeting of the Association for Computational Linguistics (ACL), 2020

How Can We Know What Language Models Know?

Frank F. Xu*, Zhengbao Jiang*, Jun Araki, Graham Neubig.

In Transactions of the Association for Computational Linguistics (TACL), 2020

A Benchmark for Structured Procedural Knowledge Extraction from Cooking Videos

Frank F. Xu, Lei Ji, Botian Shi, Junyi Du, Graham Neubig, Yonatan Bisk, Nan Duan.

In Proceedings of the First International Workshop on Natural Language Processing Beyond Text @ EMNLP, 2020

Parsimonious Morpheme Segmentation with an Application to Enriching Word Embeddings

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

In Proceedings of IEEE International Conference on Big Data, 2019

HiGitClass: Keyword-Driven Hierarchical Classification of GitHub Repositories

Yu Zhang, Frank F. Xu, Sha Li, Yu Meng, Xuan Wang, Qi Li, Jiawei Han.

In Proceedings of IEEE International Conference on Data Mining (ICDM), 2019

StateLens: A Reverse Engineering Solution for Making Existing Dynamic Touchscreens Accessible

Anhong Guo, Junhan Kong, Michael Rivera, Frank F. Xu, Jeffrey P. Bigham.

In Proceedings of 32nd Annual ACM Symposium on User Interface Software and Technology (UIST), 2019

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 LocatedNear 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

Zequiu 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 Located Near 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

WORKSHOP ORGANIZER

CMU Agent Workshop 2024 (<https://cmu-agent-workshop.github.io/>)

PROGRAM COMMITTEE MEMBER

First Workshop on Natural Language Processing for Programming at ACL-IJCNLP 2021 (NLP4Prog)

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

CONFERENCE REVIEWER

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 2018-2022)

Conference on Empirical Methods in Natural Language Processing (EMNLP 2021-2024)

The Conference on Neural Information Processing Systems (NeurIPS 2021-2024)

The International Conference on Learning Representations (ICLR 2022-2024)

ACL Rolling Review (monthly since July 2021)

JOURNAL REVIEWER

IEEE Transactions on Artificial Intelligence

IEEE Transactions on Cybernetics

Invited Talks

May 2024 Language Models as Agents

*CMU Agent
Workshop
JetBrains s.r.o.*

Jan 2022 In-IDE Code Generation from Natural Language: Promise and Challenges