

# SHUYAN ZHOU

<https://shuyanzhou.github.io> | [shuyanzh@cs.cmu.edu](mailto:shuyanzh@cs.cmu.edu) | (+1) 412-452-4358

## EDUCATION

Ph.D. in LANGUAGE TECHNOLOGIES

Aug 2020 - May 2024 (expected)

**Carnegie Mellon University**

Advisor: Graham Neubig

M.S in LANGUAGE TECHNOLOGIES

Aug 2018 - Aug 2020

**Carnegie Mellon University**

GPA: 4.00/4.00

Advisor: Graham Neubig

## RESEARCH EXPERIENCE

**Carnegie Mellon University**, Pittsburgh, PA

Aug 2018 - Present

Graduated Research Assistant (Advisor: Graham Neubig)

**X, the moonshot factory (Google[x])**, Mountain View, CA

May 2022 - Dec 2022

Ph.D. residency & Student Researcher (Host: Alex Polozov)

**Microsoft Dynamic 365**, Redmond, WA

May 2020 - Aug 2020

Research Intern (Host: Kaushik Chakrabarti)

**Microsoft Research Asia**, Beijing, China

Jul 2017 - Jun 2018

Research Intern (Host: Chin-Yew Lin)

## PUBLICATIONS

\* denotes equal contributions

- [1] **Shuyan Zhou\***, Frank F. Xu\*, Hao Zhu, Xuhui Zhou, Robert Lo, Abishek Sridhar, Xianyi Cheng, Tianyue Ou, Yonatan Bisk, Daniel Fried, Uri Alon, Graham Neubig. WebArena: A Realistic Web Environment for Building Autonomous Agents (ICLR 2024)
- [2] **Shuyan Zhou**, Uri Alon, Frank F. Xu, Zhengbao Jiang, Graham Neubig. DocPrompting: Generating Code by Retrieving the Docs (ICLR 2023 **spotlight**)
- [3] Luyu Gao\*, Aman Madaan\*, **Shuyan Zhou\***, Uri Alon, Pengfei Liu, Yiming Yang, Jamie Callan, Graham Neubig. PaL: Program-aided Language Models (ICML 2023)
- [4] **Shuyan Zhou\***, Uri Alon\*, Sumit Agarwal, Graham Neubig. CodeBERTScore: Evaluating Code Generation with Pretrained Models of Code (EMNLP 2023, DL4C Workshop **spotlight**)
- [5] Abishek Sridhar, Robert Lo, Frank F. Xu, Hao Zhu, **Shuyan Zhou**. Hierarchical Prompting Assists Large Language Model on Web Navigation (Findings of EMNLP 2023)
- [6] Zhiruo Wang, Shuyan Zhou, Daniel Fried, Graham Neubig. Execution-Based Evaluation for Open-Domain Code Generation (Findings of EMNLP 2023)
- [7] Patrick Fernandes, Aman Madaan, Emmy Liu, António Farinhas, Pedro Henrique Martins, Amanda Bertsch, José GC de Souza, **Shuyan Zhou**, Tongshuang Wu, Graham Neubig, André FT Martins. Bridging the gap: A Survey on Integrating (Human) Feedback for Natural Language Generation (TACL 2023)
- [8] Zhiruo Wang, Grace Cuenca, **Shuyan Zhou**, Frank F. Xu, Graham Neubig MCoNaLa: A Benchmark for Code Generation from Multiple Natural Languages (Findings of EACL 2023)
- [9] Li Zhang, Hainiu Xu, Yue Yang, **Shuyan Zhou**, Weiqiu You, Manni Arora, Chris Callison-Burch. Causal Reasoning of Entities and Events in Procedural Texts (Findings of EACL 2023)

- [10] Aman Madaan, **Shuyan Zhou**, Uri Alon, Yiming Yang, Graham Neubig. Language Models of Code are Few-Shot Commonsense Learners (EMNLP 2022)
- [11] **Shuyan Zhou\***, Li Zhang\*, Yue Yang, Qing Lyu, Ricardo Gonzalez, Pengcheng Yin, Chris Callison-Burch, Graham Neubig. Show Me More Details: Discovering Hierarchies of Procedures from Semi-structured Web Data (ACL 2022)
- [12] **Shuyan Zhou**, Pengcheng Yin, Graham Neubig. Procedures as Programs: Hierarchical Control of Situated Agents through Natural Language (Structured and Unstructured Knowledge Integration Workshop 2022)
- [13] Shruti Rijhwani, **Shuyan Zhou**, Graham Neubig, Jaime Carbonell. Soft Gazetteers for Low-resource Named Entity Recognition (ACL 2020)
- [14] **Shuyan Zhou**, Shruti Rijhwani, John Wieting, Jaime Carbonell, Graham Neubig. Improving Candidate Generation for Low-resource Cross-lingual Entity Linking (TACL 2020)
- [15] **Shuyan Zhou**, Shruti Rijhwani, Graham Neubig. Towards Zero-resource Cross-lingual Entity Linking (EMNLP DeepLo Workshop 2019)
- [16] **Shuyan Zhou**, Xiangkai Zeng, Yingqi Zhou, Antonios Anastasopoulos, Graham Neubig. Improving Robustness of Neural Machine Translation with Multi-task Learning (WMT 2019)
- [17] Feng Nie, **Shuyan Zhou**, Jing Liu, Jinpeng Wang, Chin-Yew Lin, Rong Pan. Aggregated Semantic Matching for Short Text Entity Linking (CoNLL 2018)

## TEACHING

---

### Teaching Assistant

CMU 11711 Advanced NLP	Fall 2021
CMU 11747 Neural Network for Natural Language Processing	Spring 2021

### Predoctoral Student Mentoring

Tianyue Ou	CMU MIIS	Sep 2023 - Present
Zhiruo Wang [6][8]	CMU MLT → CMU Ph.D.	Dec 2021 - Sep 2023
Robert Lo [5]	CMU MSCS → Tesla	Dec 2022 - Dec 2023
Abishek Sridhar [5]	CMU MSCS → Snowflake	Dec 2022 - Dec 2023
Sumit Agarwal	CMU MLT → Palantir Technologies	Sep 2022 - Dec 2022
Grace Cuenca [8]	Princeton Undergrad → Goldman Sachs	May 2021 - Sep 2021
Neha Nishikant	CMU Undergrad → Palantir Technologies	Jan 2021 - May 2021

## PROFESSIONAL SERVICES

---

### Paper Review

• ACL	2021, 2022, 2023
• EMNLP	2022, 2023
• NeurIPS	2023
• ACL Rolling Review	2021, 2022
• AAAI	2022
• EACL	2021 (Outstanding Reviewer)
• Workshop on Foundation Model for Decision Making at NeurIPS	2023
• Workshop on Pretrained Large Model at ICLR	2023
• Workshop on Interactive Learning For Natural Language Processing at NeurIPS	2022
• ACL Student Workshop	2020

### DEI

• Member & Student Representative of CMU LTI DEI Committee	2023
--	------

- Founding Member & Organizer of CMU SCS Graduate Application Support Program  
2021, 2022
- Mentor of CMU Graduate Application Support Program 2021, 2022
- Mentor of CMU LTI Mentoring Program 2021
- Mentor of CMU Undergrad AI Mentoring Program 2021, 2022

## AWARDS & HONORS

---

Center for AI Safety Compute Grant 2023

## INVITED TALKS & INTERVIEWS

---

OSU CSE Nov 2023

**Solving Complex Tasks with AI Agents**

[Dataskptic \(podcast\)](#) Nov 2023

**Language Models of Code are Few-Shot Commonsense Learners**

MiraclePlus (former Y Combinator China) Sep 2023

**Controlling Digital Agents through Natural Language**

AutoGPT (panelist) Aug 2023

**Journey to AI Excellence: Exploring the Intricacies of Benchmarking**

CMU Foundations of Statistical ML Group Mar 2023

**Large Language Models Meet Symbolic Reasoning**

Baidu NLP Feb 2023

**Natural Language Reasoning with Large Models of Code**

## MEDIA COVERAGE

---

[Data Science Central](#) Aug 2023

**Pushing boundaries with Generative AI: How Program-aided Language model (PAL) enhances Large Language Models (LLMs) for superior AI performance**

[MarkTechPost](#) Jul 2023

**CMU Researchers Introduce WebArena: A Realistic and Reproducible Web Environment with 4+ Real-World Web Apps for Benchmarking Useful Agents**

[TheSequence](#) Jul 2023

**Edge 307: Learning About Program-Aided Language Models**

[Non-Brand Data](#) Feb 2023

**Automatic Evaluation Metric for Code with CodeBERTScore**

[MarkTechPost](#) Feb 2023

**CMU Researchers Propose DocPrompting: A Natural Language To Code Generation Approach By Retrieving Code Documentation**

[Prophet-Con](#) Feb 2023

**CMU Researchers Suggest DocPrompting: A Pure Language To Code Technology Method By Retrieving Code Documentation**

[Synched](#) Feb 2023

**CMU & Inspired Cognition's DocPrompting Improves Code Generation by Retrieving Relevant Documentation**