

# ZORA (ZHIRUO) WANG

✉ zhiruow@cs.cmu.edu | 🌐 zorazrw.github.io | 📄 google scholar | 📧 zorazrw

## EDUCATION

<b>Carnegie Mellon University, Language Technologies Institute</b> PhD of Language Technologies Advisor: Daniel Fried, Graham Neubig	<b>Pittsburgh, PA</b> Aug 2023 - Present
<b>Carnegie Mellon University, Language Technologies Institute</b> Master of Language Technologies (MLT), GPA: 4.02 / 4.3 Advisors: Graham Neubig, Eric Nyberg	<b>Pittsburgh, PA</b> Aug 2021 - Aug 2023
<b>Beijing Normal University, School of Mathematical Sciences</b> B.S. in Mathematics & Applied Mathematics, major GPA: 3.98 / 4.0 ◦ Exchange, Chinese University of Hong Kong, full scholarship ◦ Summer School, The College of William & Mary, full scholarship	<b>Beijing, China</b> Sep 2016 - Jun 2020 Spring 2019 Summer 2017

## EXPERIENCE

<b>Stanford University</b> <i>Visiting Student Researcher</i> , Host: Diyi Yang ◦ Study the impact of language model supported computer-use agents on the human workforce ◦ Develop a workflow induction toolkit to inspect and analyze human and agent worker activities	<b>Palo Alto, CA</b> May 2025 - Aug 2025
<b>Microsoft, Github Next Team</b> <i>Consultant</i> , Host: Eddie Aftandilian ◦ Consult on code language model training and robust open-domain evaluation techniques	<b>Remote</b> Aug 2023 - Dec 2023
<b>Amazon Alexa AI, Conversational Modeling Team</b> <i>Research Intern</i> , Manager: Dilek Hakkani-Tur ◦ Proposed a model self-talk schema for knowledge-guided dialogue generation, further bootstrapping large-scale data for adaptive training, under a special (book) domain and low-resource setting	<b>Sunnyvale, CA</b> May 2022 - Aug 2022
<b>Microsoft Research (Asia), Data Knowledge Intelligence Group</b> <i>Research Intern</i> → <i>Assistant Researcher</i> , Manager: Dongmei Zhang ◦ Designed and pre-trained a structure-aware table foundation model, <i>TUTA</i> , with tree-based masked attention and adaptive position encodings, and shaped it into Microsoft <i>Word</i> and <i>Excel</i> ◦ Built <i>HiTab</i> benchmark, supporting table QA and data analysis for complex-structured tables	<b>Beijing, China</b> Aug 2020 - Aug 2021

## SELECTED PUBLICATIONS

- [1] *How Do AI Agents Do Human Work? Comparing AI and Human Workflows Across Diverse Occupations*  
**Zora Zhiruo Wang**, Yijiao Shao, Omar Shaikh, Daniel Fried, Graham Neubig, Diyi Yang  
*arXiv Preprint*
- [2] *Inducing Programmatic Skills for Agentic Tasks*  
**Zora Zhiruo Wang**, Apurva Gandhi, Graham Neubig, Daniel Fried  
*the Conference on Language Modeling (CoLM)*, 2025
- [3] *Agent Workflow Memory*  
**Zora Zhiruo Wang**, Jiayuan Mao, Daniel Fried, Graham Neubig  
*the International Conference on Machine Learning (ICML)*, 2025

- [4] *TheAgentCompany: Benchmarking LLM Agents on Consequential Real World Tasks*  
Frank F. Xu, ... , **Zora Zhiruo Wang**, ..., Shuyan Zhou, Graham Neubig  
*Neural Information Processing Systems (NeurIPS)*, 2025
- [5] *Benchmarking Failures in Tool-Augmented Language Models*  
Eduardo Treviño\*, Hugo Contant\*, James Ngai, Graham Neubig, **Zora Zhiruo Wang<sup>m</sup>**  
*the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2025
- [6] *CowPilot: A Framework for Autonomous and Human-Agent Collaborative Web Navigation*  
Faria Huq, **Zora Zhiruo Wang<sup>m</sup>**, Frank Xu, Tianyue Ou, Shuyan Zhou, Jeffrey P. Bigham, Graham Neubig  
*the North American Chapter of the Association for Computational Linguistics (NAACL), Demo Track*, 2025
- [7] *AutoPresent: Designing Structured Visuals from Scratch*  
Jiaxin Ge\*, **Zora Zhiruo Wang<sup>\*,m</sup>**, ..., Alane Suhr, Daniel Fried, Graham Neubig, Trevor Darrell  
*the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025
- [8] *CodeRAG-Bench: Can Retrieval Augment Code Generation?*  
**Zora Zhiruo Wang\***, Akari Asai\*, Xiyan Yu, Frank F. Xu, Yiqing Xie, Graham Neubig, Daniel Fried  
*the North American Chapter of the Association for Computational Linguistics (NAACL), Findings*, 2025
- [9] *What Are Tools Anyway? A Survey from the Language Model Perspective*  
**Zora Zhiruo Wang**, Zhoujun Cheng, Hao Zhu, Daniel Fried, Graham Neubig  
*the Conference on Language Modeling (CoLM)*, 2024
- [10] *TroVE: Inducing A Verifiable and Efficient Toolbox for Programmatic Tasks*  
**Zora Zhiruo Wang**, Graham Neubig, Daniel Fried  
*the International Conference on Machine Learning (ICML)*, 2024
- [11] *ECCO: Can We Improve Model-Generated Code Efficiency Without Sacrificing Functional Correctness?*  
Siddhant Waghjale, Vishruth Veerendranath, **Zora Zhiruo Wang<sup>m</sup>**, Daniel Fried  
*the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024
- [12] *API-Assisted Code Generation for Question Answering on Varied Table Structured*  
Yihan Can, Shuyi Chen, Ryan Liu, **Zora Zhiruo Wang<sup>m</sup>**, Daniel Fried  
*the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023
- [13] *StarCoder: may the source be with you!*  
Raymond Li, Loubna Ben Allal, ..., **Zora Zhiruo Wang**, ..., Leandro von Werra, Harm de Vries  
*the Transactions on Machine Learning Research (TMLR)*
- [14] *Execution-Based Evaluation for Open-Domain Code Generation*  
**Zora Zhiruo Wang**, Shuyan Zhou, Daniel Fried, Graham Neubig  
*the Conference on Empirical Methods in Natural Language Processing (EMNLP), Findings*, 2023
- [15] *MCoNaLa: A Benchmark for Code Generation from Multiple Natural Languages*  
**Zora Zhiruo Wang**, Grace Cuenca, Frank F. Xu, Shuyan Zhou, Graham Neubig  
*the European Chapter of the Association for Computational Linguistics (EACL)*, 2023
- [16] *DocPrompting: Generating Code By Retrieving the Docs*  
Shuyan Zhou, Uri Alon, Frank F. Xu, **Zora Zhiruo Wang**, Zhengbao Jiang, Graham Neubig  
*the International Conference on Learning Representations (ICLR)*, 2023
- [17] *HiTab: A Hierarchical Table Dataset for Question Answering and Natural Language Generation*  
Zhoujun Cheng\*, Haoyu Dong\*, **Zora Zhiruo Wang\***, R. Jia, J. Guo, Y. Gao, S. Han, J. Lou, D. Zhang  
*the Association for Computational Linguistics (ACL)*, 2022

## INVITED TALKS

---

“How Do AI Agents Do Human Work”

AI-TIME

Dec 2025

<b>“From Workflows to Tools: Building Adaptive Agents On the Fly”</b> Renmin University of China (RUC)	Jul 2025
<b>“From Workflows to Tools: Building Adaptive Agents On the Fly”</b> Pareto AI	Jun 2025
<b>“From Workflows to Tools: Building Adaptive Agents On the Fly”</b> Orby AI	Jun 2025
<b>“Building TheAgentCompany: Benchmarking Consequential Real-World Tasks and Adaptive Knowledge Learning”</b> AAAI 2025 Workshop on Advancing LLM-Based Multi-Agent Collaboration	Mar 2025
<b>“TheAgentCompany: Benchmarking LLM Agents on Consequential Real-World Tasks”</b> AAAI Web Agent Revolution Workshop	Mar 2025
<b>“From Workflows to Tools: Building Adaptive Agents On the Fly”</b> AAAI Web Agent Revolution Workshop	Mar 2025
<b>“From Workflows to Tools: Building Adaptive Agents On the Fly”</b> Tsinghua University	Dec 2024
<b>“Solving Real-World Tasks via Program Generation”</b> UIUC (University of Illinois Urbana-Champaign) iSE Seminar	Nov 2024
<b>“Agent Workflow Memory”</b> CAMEL-AI	Sep 2024
<b>“What Are Tools Anyway?”</b> FLAME (Foundation and LAnguage Model) seminar	May 2024
<b>“Tool using, learning, and making with LLMs”</b> LLM as Agent Seminar	Feb 2024
<b>“Execution-based Open-Domain Coding”</b> JetBrains Research Team	Apr 2023

## COMMUNITY SERVICES

---

### Workshop & Tutorial Organization

◦ Organizer: “The 4th Deep Learning for Code Workshop” at NeurIPS	Dec 2025
◦ Organizer: “The Third Deep Learning for Code Workshop” at ICLR	Apr 2025
◦ Tutorial Presenter: “Large Language Models for Tabular Data” at SIGIR	Jul 2024
◦ Head Organizer of the 1st CMU Agent Workshop	May 2024
◦ Head Organizer of CMU Student Research Symposium	Aug 2022

### Department Services

◦ Member of Faculty Hiring Student Committee; LTI, CMU	2025
◦ Member & Student Representative of CMU LTI DEI Committee	2024
◦ Chair of Faculty Hiring Student Committee; LTI, CMU	2024

### Paper Review

◦ Program committee for NeurIPS, ICLR, ICML, CoLM; ACL, EMNLP, NAACL, EACL; AAAI
--

## TEACHING EXPERIENCE

---

## Teaching Assistant

- Neural Code Generation, Spring 2024: lead student discussion, advise students research projects
- Advanced NLP, Spring & Fall 2023: head TA; prepare assignments, advise students research projects
- Data Science Seminar, Fall 2021: organize research paper discussion, advise students research proposal

## Guest Lectures

- “Language Model-Based Agents” at Human-Centered NLP
- “Language Model-Based Agents” at Clinical Translation of AI
- “Evaluation & Benchmarks for Code Generation” at Neural Code Generation and Advanced NLP
- “Tool Use and Language Agents” at Advanced NLP

## AWARDS & HONORS

---

Google PhD Fellowship	2025
CMU Presidential Fellowship	2023
China National Scholarship recipient	2019

## MEDIA COVERAGE

---

<b>Incubity</b>	Sep 2024
Agent Workflow Memory: Transforming AI Task Management	
<b>Medium</b>	Sep 2024
Agent Workflow Memory: using workflows to guide LLM Agent generations	
<b>MarkTechPost</b>	Sep 2024
Agent Workflow Memory (AWM): An AI Method for Improving the Adaptability and Efficiency of Web Navigation Agents	
<b>MarkTechPost</b>	Aug 2024
A Reproducible AI Benchmark for Evaluating Program Efficiency via Two Paradigms	
<b>MarkTechPost</b>	Jul 2023
Google Research Introduces SPAE: An AutoEncoder For Multimodal Generation With Frozen Large Language Models (LLMs)	
<b>SyncedReview</b>	May 2023
‘May the Source Be With You!’ –BigCode’s Open-Access StarCoder Outperforms All Existing Open Code LLMs	