

# Li “Harry” Zhang

zharry.com  
Harry.Zhang@drexel.edu

Last updated: Sept 2025

## RESEARCH INTERESTS

Artificial Intelligence, Machine Learning, Natural Language Processing  
Large Language Models, Planning, Reasoning, Agents, Formal Methods, etc.

## ACADEMIC AFFILIATIONS

**Drexel University**, Philadelphia, PA Dec 2024 – Present  
Tenure-Track Assistant Professor  
**University of Pennsylvania**, Philadelphia, PA Aug 2019 – Aug 2024  
Ph.D. Computer and Information Science GPA: 3.96/4.00  
Adviser: Prof. Chris Callison-Burch  
Thesis: *Structured Event Reasoning with Large Language Models*  
Committee: Prof. Dan Roth (chair), Prof. Rada Mihalcea, Prof. Graham Neubig,  
Prof. Mark Yatskar, Dr. Marianna Apidianaki  
**University of Michigan**, Ann Arbor, MI Aug 2015 – Dec 2018  
B.S.E. Computer Science, summa cum laude GPA: 3.82/4.00  
Mentors: Prof. Rada Mihalcea and Prof. Dragomir Radev

## PUBLICATIONS

28 papers published in top NLP/AI conferences and workshops  
▷ 13 first-authored by self ∪ 8 first-authored by advised or mentored students.  
Total citations: 3000+; h-index: 15+  
(\*Equal contribution; ^Mentored students)  
[35] L. Gong, W. Zhu, J. Thomason and **L. Zhang**. Zero-Shot Iterative Formalization and Planning in Partially Observable Environments. Preprint.  
[34] P. Kagitha and **L. Zhang**. Addressing the Challenges of Planning Language Generation. Preprint.  
[33] R. Amonkar, M. Lai, R. Le Bras and **L. Zhang**. Are LLMs Better Formalizers than Solvers on Complex Problems? Preprint.  
[32] Y. Yuan, M. He, A. Shahid, J. Huang, Z. Li, **L. Zhang**. TurnaboutLLM: A Deductive Reasoning Benchmark from Detective Games. In EMNLP 2025.  
[31] W. Hu, J. Duan, C. Wei, **L. Zhang**, Y. Zhang and K. Xu. DynaCode: A Dynamic Complexity-Aware Code Benchmark for Evaluating Large Language Models in Code Generation. In Findings of ACL 2025.  
[30] C. Huang and **L. Zhang**. On the Limit of Language Models as Planning Formalizers. In ACL 2025.  
[29] **L. Zhang**, P. Jansen, P. Clark, C. Callison-Burch and N. Tandon. PDDLEGO: Iterative Planning in Textual Environments. In \*SEM 2024.  
[28] T. Zhang\*, **L. Zhang\***, Z. Hou^, Z. Wang^, Y. Gu, P. Clark, C. Callison-Burch and N. Tandon. PROC2PDDL: Open-Domain Planning Representations from Texts. In the 2<sup>nd</sup> Natural Language Reasoning and Structured Explanations Workshop at ACL 2024.  
[27] Q. Lyu, K. Shridhar, C. Malaviya, **L. Zhang**, Y. Elazar, N. Tandon, M. Apidianaki, M. Sachan and C. Callison-Burch. Calibrating Large Language Models with Sample Consistency. In AAAI 2025; **Area Chair Award**.  
[26] Y. Lal, **L. Zhang**, F. Brahman, B. Majumder, Peter Clark and N. Tandon. One Size Does Not Fit All: Customizing Open-Domain Procedures. In Findings of ACL 2024.

- [25] B. Majumder, B. Dalvi, P. Jansen, O. Tafjord, N. Tandon, **L. Zhang** and C. Callison-Burch, Peter Clark. CLIN: A Continually Learning Language Agent for Rapid Task Adaptation and Generalization. In COLM 2024.
- [24] Z. Hou<sup>^</sup>, **L. Zhang** and C. Callison-Burch. *Choice-75: A Dataset on Decision Branching in Script Learning*. In LREC-COLING 2024.
- [23] **L. Zhang**, H. Xu<sup>^</sup>, A. Kommula, N. Tandon and C. Callison-Burch. *OpenPI2.0: An Improved Dataset for Entity Tracking in Texts*. In EACL 2024.
- [22] **L. Zhang**<sup>\*</sup>, L. Dugan<sup>\*</sup>, H. Xu<sup>^</sup> and C. Callison-Burch. *Exploring the Curious Case of Code Prompts*. In preprint. In the 1<sup>st</sup> Natural Language Reasoning and Structured Explanations Workshop at ACL 2023.
- [21] T. Zhang<sup>^</sup>, I. Tham, Z. Hou<sup>^</sup>, Jia. Ren, L. Zhou, H. Xu<sup>^</sup>, **L. Zhang**, L. Martin, R. Dror, S. Li, H. Ji, M. Palmer, S. Brown, R. Suchocki, C. Callison-Burch. *Human-in-the-Loop Schema Induction*. In preprint; in ACL 2023 Demos.
- [20] Q. Lyu<sup>\*</sup>, S. Havaladar<sup>\*</sup>, A. Stein<sup>\*</sup>, **L. Zhang**, D. Rao, E. Wong, M. Apidianaki and C. Callison-Burch. *Faithful Chain of Thought Reasoning*. In IJCNLP-AAACL 2023.
- [19] **L. Zhang**<sup>\*</sup>, H. Xu<sup>^</sup>, Y. Yang, S. Zhou, W. You, M. Arora and C. Callison-Burch. *Causal Reasoning of Entities and Events in Procedural Texts*. In Findings of EACL 2023.
- [18] **L. Zhang** and C. Callison-Burch. *Language Models are Drummers: Drum Composition with Natural Language Pre-Training*. In 1<sup>st</sup> Workshop on Creative AI across Modalities at AAAI 2023.
- [17] Y. M. Cho<sup>^</sup>, **L. Zhang** and C. Callison-Burch. *Unsupervised Entity Linking with Guided Summarization and Multiple Choice Selection*. In EMNLP 2022.
- [16] S. Gehrmann, ..., **L. Zhang**, ..., H. Zhu, S. Brahma, Y. Li, ... *GEMv2: Multilingual NLG Benchmarking in a Single Line of Code*. In EMNLP 2022.
- [15] A. Srivastava, ..., **L. Zhang**, Q. Lyu and C. Callison-Burch, ... *Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models*. In TMLR.
- [12] Q. Lyu, H. Zheng, D. Li, **L. Zhang**, M. Apidianaki, and C. Callison-Burch. *Is "my favorite new movie" my favorite movie? Probing the Understanding of Recursive Noun Phrases*. In NAACL 2022.
- [11] **L. Zhang**, I. Jindal and Y. Li. *Label Definitions Improve Semantic Role Labeling*. In NAACL 2022.
- [10] **L. Zhang**<sup>\*</sup>, S. Zhou<sup>\*</sup>, Q. Lyu, Y. Yang, G. Neubig and C. Callison-Burch. *Show Me More Details: Discovering Event Hierarchies from WikiHow*. In ACL 2022.
- [9] Y. Yang, A. Panagopoulou, Q. Lyu, **L. Zhang**, M. Yatskar and C. Callison-Burch. *Visual Goal-Step Inference using wikiHow*. In EMNLP 2021; presented at the 2nd Workshop on Advances in Language and Vision Research at NAACL 2021.
- [8] **L. Zhang**<sup>\*</sup>, Q. Lyu<sup>\*</sup> and C. Callison-Burch. *Goal-Oriented Script Construction*. In INLG 2021.
- [7] **L. Zhang**, Q. Lyu and C. Callison-Burch. *Intent Detection with WikiHow*. In AACL-IJCNLP 2020.
- [6] **L. Zhang**<sup>\*</sup>, Q. Lyu<sup>\*</sup> and C. Callison-Burch. *Reasoning about Goals, Steps, and Temporal Ordering with WikiHow*. In EMNLP 2020; presented at Workshop on Enormous Language Models at ICLR 2021.
- [5] **L. Zhang**, H. Zhu, S. Brahma and Y. Li. *Small but Mighty: New Benchmarks for Split and Rephrase*. In EMNLP 2020.
- [4] **L. Zhang**, S. R. Wilson and R. Mihalcea. *Multi-Label Transfer Learning for Semantic Similarity*. In \*SEM 2019 and presented at NAACL 2019.
- [1] C. Finegan-Dollak, J. K. Kummerfeld, **L. Zhang**, K. R. D. Ramanathan, S. Sadasivam, R. Zhang and D. Radev. *Improving Text-to-SQL Evaluation Methodology*. In ACL 2018.

<b>External Funding</b>	<b>Decision-Embedded Deep Learning for Transit Systems (\$432,572)</b> 2024 - 2027 <i>NSF</i> • Inherited from Dr. Kaidi Xu.	
	<b>Alexa Prize TaskBot Challenge (\$250,000)</b> 2021 - 2022 <i>Amazon</i> • Primarily authored, applied, and received a stipend award of \$250,000 to lead University of Pennsylvania's effort in the Alexa Prize TaskBot Challenge 2021.	
<b>INDUSTRY EXPERIENCE</b>	<b>Research Intern</b> Apr 2023 – Dec 2023 <i>Allen Institute for Artificial Intelligence (AI2)</i> Seattle, WA	
	<b>Research Intern</b> Apr 2019 – Jun 2019; May 2021 – Aug 2021 <i>IBM Research</i> San Jose, CA	
	<b>Software Engineer Intern</b> May 2017 – Aug 2017 <i>Goldman Sachs Group, Inc.</i> Jersey City, NJ	
<b>ACADEMIC SERVICE</b>	<b>Chair</b> • Area Chair: AAACL 2025, EMNLP 2025, ACL 2025, ACL 2024, EMNLP 2024, COLING 2024	
	• Program Chair: Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL) 2023	
	• Program Chair: 1 <sup>st</sup> Workshop on Data Science with Human in the Loop at EMNLP 2022 2022	
	• Session Chair: Asia-Pacific Chapter of the Association of Computational Linguistics (AAACL) 2020	
	<b>Reviewer</b> • Association of Computational Linguistics (ACL) recurring	
	• North American Chapter of ACL (NAACL) recurring	
	• Empirical Methods in Natural Language Processing (EMNLP) recurring	
	• Association for the Advancement of Artificial Intelligence (AAAI) recurring	
	• Conference on Language Modeling (COLM) recurring	
	• International Conference on Language Resources and Evaluation (LREC) recurring	
<b>TEACHING</b>	<b>Instructor — Applied Natural Language Processing</b> Apr 2025 – Jun 2025 <u>CS T780: The graduate level NLP course</u> Drexel University	
	<b>Teaching Assistant — Computational Linguistics</b> Jan 2020 – Dec 2020 <u>CIS 530: The graduate level NLP course</u> University of Pennsylvania	
	<b>Teaching Assistant — Natural Language Processing</b> Sept 2018 – Dec 2018 <u>EECS 595: The graduate level NLP course</u> University of Michigan	
	<b>Teaching Assistant — Programming and Data Structures</b> Sept 2016 – Apr 2017 <u>EECS 280: An introductory programming course</u> University of Michigan	
	<b>Tutor — Elementary Chemistry</b> Sept 2016 – Dec 2016 <u>Science Learning Center</u> University of Michigan	

## **ADVISING**

### **PhD Students**

Cassie Huang

Jan 2025 – present

Ceyhun Efe Kayan

Sep 2025 – present

### **Master Students**

Prabhu Prakash Kagitha

Feb 2025 – present

Chimezie Maduno

May 2025 – present

### **Undergraduate Students**

Rikhil Amonkar

Feb 2025 – present

Stuti Mohan

Jun 2025 – present

### **Interns and Visiting Students**

Muyu He

Mar 2025 – present

Yuan Yuan

Sep 2024 – present

Renxiang Wang

Apr 2025 – present

### **Alumni and Past Students**

Krystal Gong, Tianyi Zhang, Hainiu Xu, Zhaoyi Hou, Young-Min Cho