Curriculum Vitae

Zhewei Sun

zheweisun (at) cs.toronto.edu

http://www.cs.toronto.edu/~zheweisun/

Education

September 2018 —
Present

University of Toronto, Toronto, ON
Ph.D. in Computer Science
Thesis title: *Natural Language Processing for Slang*
Advisor: *Yang Xu*

August 2016 —
May 2018

Georgia Institute of Technology, Atlanta, GA
M.S. in Computer Science
Specialization in Machine Learning

September 2012 —
June 2016

University of Waterloo, Waterloo, ON
B.S. in Computer Science
Minor in Economics

Publications

6. **Toward Informal Language Processing: Knowledge of Slang in Large Language Models**

   Zhewei Sun, Qian Hu, Rahul Gupta, Richard Zemel, and Yang Xu

   To appear in Proceedings of the 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024, Long Paper, Oral).

5. **Tracing Semantic Variation in Slang**

   Zhewei Sun and Yang Xu

   In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022, Long Paper, Oral).

4. **Semantically Informed Slang Interpretation**

   Zhewei Sun, Richard Zemel, and Yang Xu

   In Preceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT 2022, Long Paper, Oral).

3. **A Computational Framework for Slang Generation**

   Zhewei Sun, Richard Zemel, and Yang Xu

Transactions of the Association for Computational Linguistics, 9, 462-478 (TACL 2021, Oral Presentation at ACL 2021).

2. **Slang Detection and Identification**

   Zhengqi Pei, *Zhewei Sun*, and Yang Xu

   In Proceedings of the 2019 Conference on Computational Natural Language Learning (CoNLL 2019, Long Paper, Poster).

1. **Slang Generation as Categorization**

   Zhewei Sun, Richard Zemel, and Yang Xu

   In Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019).

Honors and awards

- Queen Elizabeth II Graduate Scholarships in Science & Technology (2022)
- Cecil and Edna Cotton Scholarship (2015)
- University of Waterloo - President's Research Award (2015)
- University of Waterloo - President's Scholarship of Distinction (2012)
- Bronze Award - Canadian Computing Competition Stage Two (2011)

Experience

September 2018 — Present	**University of Toronto - Computational Linguistics Group** Research Assistant Advisor: *Yang Xu* Project: Natural language processing for informal language, with specific focus on slang.
August 2017 — May 2018	**Georgia Institute of Technology - Computational Linguistics Lab** Graduate Research Assistant Advisor: *Jacob Eisenstein* Project: Discourse neural machine translation and language modelling.
August 2016 — December 2016	**Georgia Institute of Technology - Sonification Lab** Research and Development Studio Advisor: *Bruce Walker* Project: Heads-up displays in driving simulators.

September 2015 — *December 2015*	**University of Waterloo - Department of Computer Science** CS Undergraduate Research Assistant Advisor: *Justin W. L. Wan* Project: Computational finance on GPU systems.
May 2013 — *August 2013*	**IBM Canada** Software Enginnering Intern Supervisor: *Christopher Kaufman* Project: Score-carding feature development for the Cognos TM1 business analytics suite.

Teaching

September 2018 — *Present*	**University of Toronto - Department of Computer Science** Teaching Assistant Courses: • CSC485/2501 - Computational Linguistics (Fall 2018, Fall 2019, Fall 2020, Fall 2021, Fall 2022, Fall 2023). • CSC401/2511 - Natural Language Computation (Winter 2019, Winter 2020, Winter 2021). • CSC384 - Introduction to Artificial Intelligence (Summer 2019, Summer 2020, Summer 2021, Summer 2022, Summer 2023).
August 2017 — *May 2018*	**Georgia Institute of Technology - School of Interactive Computing** Graduate Teaching Assistant Courses: • CS7646 - Machine Learning for Trading (Fall 2017). • CS7650 - Natural Language Processing (Spring 2018).

Invited talks & posters

February 2021	**Vector Institute Research Symposium 2021** Title: A Computational Framework for Slang Generation
September 2020	**Vector Institute Natural Language Processing Symposium** Title: Informal Natural Language Processing: The Case of Slang

| *July 2020* | **The 42nd Annual Meeting of the Cognitive Science Society** |
| | Title: The Emergence and Propagation of Online Slang |

| *October 2019* | **Vector Institute Natural Language Processing Symposium** |
| | Title: Informal Natural Language Processing: The Case of Slang |

| *September 2019* | **Google Brain Toronto** |
| | Title: Informal Natural Language Processing: The Case of Slang |

Students mentored

- Zhenyu Wang (May 2022 - August 2022)
- Sherry Yuan (May 2020 - August 2020)
- Yishan Jiang (September 2019 - April 2020)
- Zhengqi Pei (September 2018 - August 2019)

Open-source software

- SlangGen. Natural language processing pipeline to generate novel slang usages based on historical slang usages from dictionaries.

  `https://github.com/zhewei-sun/slanggen`

- CatGO. A light-weight Python implementation of popular categorization models with automatic parameter optimization.

  `https://github.com/zhewei-sun/CatGO`

Technical experience

Software

Proficient: Python, Java, C, C++, Numpy, Scipy, PyTorch, Transformers, Matlab, Dynet
Familiar: C#, XAML, JavaScript, CUDA, Scheme, R, Tensorflow, OpenNMT

Languages

English (Native), Mandarin (Native), Japanese (intermediate)