

NING SHI

Homepage | GitHub | Google Scholar | LinkedIn | Twitter | YouTube

Phone: +1 (587) 596-2062 | Email: ning.shi@ualberta.ca

Address: Edmonton, Alberta, Canada, T5J 0K5

RESEARCH INTERESTS

My research interests lie in the broad area of computational linguistics, natural language processing (NLP), reinforcement learning, and their applications. My current Ph.D. thesis aims to construct a unified framework for NLP tasks grounded in the theory of sentential semantics, enhancing model adaptability, interpretability, and fairness across tasks and languages for social good. More information can be found on my Homepage¹.

EDUCATION

University of Alberta Doctor of Philosophy in Computing Science Committee Members: Levi Santana De Lelis, Alona Fyshe, Grzegorz Kondrak	<i>Sep. 2022 - (May 2026)</i> <i>GPA: 4.0/4.0</i>
Georgia Institute of Technology Master of Science in Computer Science	<i>Aug. 2019 - May 2022</i> <i>GPA: 4.0/4.0</i>
Syracuse University Master of Science in Applied Data Science	<i>Oct. 2018 - May 2021</i> <i>GPA: 4.0/4.0</i>
New York University Master of Science in Management and Systems	<i>Sep. 2016 - May 2018</i> <i>GPA: 3.9/4.0</i>
Donghua University Bachelor of Management in Electronic Commerce	<i>Sep. 2010 - Jun. 2014</i>

PUBLICATIONS

My research has led to multiple peer-reviewed publications at top conferences, including ACL, EMNLP, EACL, *SEM, and Interspeech, as well as renowned workshops such as BlackBoxNLP and SemEval. For a full list of publications, please refer to my Google Scholar.

Peer-reviewed

- [C.1] **Ning Shi**, David Basil, Bradley Hauer, Noshin Nawal, Jai Riley, Daniela Teodorescu, John Zhang, and Grzegorz Kondrak. 2024. UAlberta at SemEval-2025 Task 2: Prompting and Ensembling for Entity-Aware Translation. In Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025), Vienna, Austria. Association for Computational Linguistics.
(🏆 **1st Place on Track COMET Score**)
- [C.2] Jai Riley, Bradley M. Hauer, Nafisa Sadaf Hriti, Guoqing Luo, Amir Reza Mirzaei, Ali Rafei, Hadi Sheikhi, Mahvash Siavashpour, Mohammad Tavakoli, **Ning Shi** and Grzegorz Kondrak. 2025. Semi-Automated Construction of Sense-Annotated Datasets for Practically Any Language. In Proceedings of the 31st International Conference on Computational Linguistics, pages 6270–6284, Abu Dhabi, UAE. Association for Computational Linguistics.
- [C.3] Senyu Li, Bradley Hauer, **Ning Shi**, and Grzegorz Kondrak. 2024. Translation-based Lexicalization Generation and Lexical Gap Detection: Application to Kinship Terms. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 6891–6900, Bangkok, Thailand. Association for Computational Linguistics.

¹sites.google.com/ualberta.ca/shining

- [C.4] **Ning Shi**, Bradley Hauer, and Grzegorz Kondrak. 2024. Lexical Substitution as Causal Language Modeling. In Proceedings of the 13th Joint Conference on Lexical and Computational Semantics (*SEM 2024), pages 120–132, Mexico City, Mexico. Association for Computational Linguistics.
- [C.5] **Ning Shi**, Bradley Hauer, Jai Riley, and Grzegorz Kondrak. 2024. Paraphrase Identification via Textual Inference. In Proceedings of the 13th Joint Conference on Lexical and Computational Semantics (*SEM 2024), pages 133–141, Mexico City, Mexico. Association for Computational Linguistics.
- [C.6] Logan Woudstra, Moyo Dawodu, Frances Igwe, Senyu Li, **Ning Shi**, Bradley Hauer, and Grzegorz Kondrak. 2024. Identifying Emotional and Polar Concepts via Synset Translation. In Proceedings of the 13th Joint Conference on Lexical and Computational Semantics (*SEM 2024), pages 142–152, Mexico City, Mexico. Association for Computational Linguistics.
- [C.7] **Ning Shi**, Senyu Li, Guoqing Luo, Amirreza Mirzaei, Ali Rafiei, Jai Riley, Hadi Sheikhi, Mahvash Siavashpour, Mohammad Tavakoli, Bradley Hauer, and Grzegorz Kondrak. 2024. UAlberta at SemEval-2024 Task 1: A Potpourri of Methods for Quantifying Multilingual Semantic Textual Relatedness and Similarity. In Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024), pages 1798–1805, Mexico City, Mexico. Association for Computational Linguistics.
(🏆 **1st Place on Track A English** among all submitted systems)
- [C.8] Xiang Zhang, Senyu Li, Bradley Hauer, **Ning Shi**, and Grzegorz Kondrak. 2023. Don’t Trust ChatGPT when your Question is not in English: A Study of Multilingual Abilities and Types of LLMs. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, pages 7915–7927, Singapore. Association for Computational Linguistics.
- [C.9] Yangyi Chen, Hongcheng Gao, Ganqu Cui, Lifan Yuan, Dehan Kong, Hanlu Wu, **Ning Shi**, Bo Yuan, Longtao Huang, Hui Xue, Zhiyuan Liu, Maosong Sun, and Heng Ji. 2023. From Adversarial Arms Race to Model-centric Evaluation: Motivating a Unified Automatic Robustness Evaluation Framework. In Findings of the Association for Computational Linguistics: ACL 2023, pages 9607–9632, Toronto, Canada. Association for Computational Linguistics.
- [C.10] Michael Ogezi, Bradley Hauer, Talgat Omarov, **Ning Shi**, and Grzegorz Kondrak. 2023. UAlberta at SemEval-2023 Task 1: Context Augmentation and Translation for Multilingual Visual Word Sense Disambiguation. In Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023), pages 2043–2051, Toronto, Canada. Association for Computational Linguistics.
- [C.11] Xiang Zhang, **Ning Shi**, Bradley Hauer, and Grzegorz Kondrak. 2023. Bridging the Gap Between BabelNet and HowNet: Unsupervised Sense Alignment and Sememe Prediction. In Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, pages 2789–2798, Dubrovnik, Croatia. Association for Computational Linguistics.
- [C.12] **Ning Shi**, Boxin Wang, Wei Wang, Xiangyu Liu, and Zhouhan Lin. 2022. Revisit Systematic Generalization via Meaningful Learning. In Proceedings of the Fifth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP, pages 62–79, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.
- [C.13] **Ning Shi**, Bin Tang, Bo Yuan, Longtao Huang, Yewen Pu, Jie Fu, and Zhouhan Lin. 2022. Text Editing as Imitation Game. In Findings of the Association for Computational Linguistics: EMNLP 2022, pages 1583–1594, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.
- [C.14] Zihan Zhang, Jinfeng Li, **Ning Shi**, Bo Yuan, Xiangyu Liu, Rong Zhang, Hui Xue, Donghong Sun, and Chao Zhang. 2022. RoChBert: Towards Robust BERT Fine-tuning for Chinese. In Findings of the Association for Computational Linguistics: EMNLP 2022, pages 3502–3516, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.
- [C.15] **Ning Shi**, Wei Wang, Boxin Wang, Jinfeng Li, Xiangyu Liu, and Zhouhan Lin. 2021. Incorporating External POS Tagger for Punctuation Restoration. In Proceedings of Interspeech 2021, pages 1987–1991,

doi: 10.21437/Interspeech.2021-1708.

- [C.16] Wei Wang, Boxin Wang, **Ning Shi**, Jinfeng Li, Bingyu Zhu, Xiangyu Liu, and Rong Zhang. 2021. Counterfactual Adversarial Learning with Representation Interpolation. In Findings of the Association for Computational Linguistics: EMNLP 2021, pages 4809–4820, Punta Cana, Dominican Republic. Association for Computational Linguistics.
- [C.17] **Ning Shi**, Ziheng Zeng, Haotian Zhang, and Yichen Gong. 2020. Recurrent Inference in Text Editing. In Findings of the Association for Computational Linguistics: EMNLP 2020, pages 1758–1769, Online. Association for Computational Linguistics.
- [C.18] Keran Zhao, Yuheng Hu, Yingda Lu, and **Ning Shi**. 2019. The Impact of Emotion on Social Media Content Consumption: Evidence from Video Analysis on Live Streaming Platforms. 2019 INFORMS Annual Meeting, Seattle, WA.

Preprint

- [P.1] **Ning Shi**, Zijun Wu. 2024. Action Controlled Paraphrasing. arXiv:2405.11277
- [P.2] Xiang Zhang, Senyu Li, **Ning Shi**, Bradley Hauer, Zijun Wu, Grzegorz Kondrak, Muhammad Abdul-Mageed, Laks VS Lakshmanan. 2024. Cross-Modal Consistency in Multimodal Large Language Models. arXiv:2411.09273.
- [P.3] Michael Ogezi, **Ning Shi**. 2024. Optimizing Negative Prompts for Enhanced Aesthetics and Fidelity in Text-To-Image Generation. arXiv:2403.07605
- [P.4] Zekun Wang, Ge Zhang, Kexin Yang, **Ning Shi**, Wangchunshu Zhou, Shaochun Hao, Guangzheng Xiong, Yizhi Li, Mong Yuan Sim, Xiuying Chen, Qingqing Zhu, Zhenzhu Yang, Adam Nik, Qi Liu, Chenghua Lin, Shi Wang, Ruibo Liu, Wenhui Chen, Ke Xu, Dayiheng Liu, Yike Guo, Jie Fu. 2023. Interactive Natural Language Processing. arXiv preprint arXiv:2305.13246.

In Submission

- [S.1] **Ning Shi**, Bradley Hauer, John Zhang, and Grzegorz Kondrak. 2025. Contradiction Detection at the Lexical Level.
- [S.2] Zekun Wang, Kang Zhu, Chunpu Xu, Wangchunshu Zhou, Jiaheng Liu, Yibo Zhang, Jiashuo Wang, **Ning Shi**, Ge Zhang, Siyu Li, Yizhi Li, Yuanxing Zhang, Ke Xu, Wenhao Huang, Jie Fu. 2024. MIO: A Foundation Model on Multimodal Tokens.

EXPERIENCE

Detailed descriptions are provided for select positions. Full details available upon request.

Association for Computational Linguistics (ACL) Edmonton, Canada (Remote)
ACL Rolling Review (ARR) Support Team Member *Oct. 2023 - Present*

- Managed media operations, including X/Twitter, LinkedIn, Facebook, and Bluesky.
- Collaborated with editors and conference committees to ensure and enhance the review process.

Alberta Machine Intelligence Institute (Amii) Edmonton, Canada
Graduate Research Assistant, Supervised by Prof. Grzegorz Kondrak *Sep. 2022 - Present*

- Researched computational linguistics and natural language processing, focusing on sentential semantics.
- Published 10 papers in peer-reviewed conferences and workshops.

University of Alberta Edmonton, Canada
Graduate Teaching Assistant / Assistant Head *Sep. 2022 - Present*

- CMPUT 174 - Introduction to the Foundations of Computation I
- CMPUT 261 - Introduction to Artificial Intelligence
- CMPUT 331 - Computational Cryptography

- CMPUT 399 - Topics in Computing Science: Emotion Identification at the Word Sense Level
- CMPUT 603 - Teaching and Research Methods
- CMPUT 650 - Topics in Artificial Intelligence: Computational Semantics

CIFAR-OSMO AI4Good Lab²

Graduate Teaching Assistant

Edmonton, Canada

Apr. 2024 - Jun. 2023

- Facilitated daily sessions, enhancing material comprehension and providing instructional support.
- Delivered recap lectures and advanced machine learning skill-development sessions.
- Supervised 2 teams, 10 trainees, in learning and developing machine learning prototypes for social good.

Computing Research Association (CRA)

Undergraduate Research to PhD Program (UR2PhD) Research Mentor

Edmonton, Canada

Sep. 2023 - Dec. 2023

- Research Topic#1 - Understanding Multilingual Emotional Expressions
- Research Topic#2 - Identifying Emotional and Polar Concepts via Synset Translation
- Supervised 3 undergraduates, leading to a paper accepted by *SEM 2024.

Beijing Academy of Artificial Intelligence (BAAI)

Graduate Research Intern, Advised by Jie Fu

Beijing, China (Remote)

Aug. 2022 - Aug. 2023

- Investigated advanced large language models on interactive natural language processing.
- Collaborated with a team of researchers, leading to two papers currently under review.

Alibaba Artificial Intelligence Governance Laboratory (AAIG)

Senior Algorithm Engineer (Algorithm Engineer II)

Hangzhou, China

Nov. 2020 - Jul. 2022

- The owner of a proactive chatbot service for violation detection.
- Explored cool applications of NLP in security.

Shining Lab

Founder, Advised by Zhouhan Lin

Shanghai, China

Sep. 2019 - Present

- Directed pioneering research projects in the field of natural language processing.
- Multiple papers accepted to top conferences (e.g., ACL, EMNLP, Interspeech).

Learnable, Inc.

Machine Learning Engineer

Boston, United States

Sep. 2018 - Sep. 2020

Pythagoras Investment Management LLC

Algorithm Engineer Intern

New York City, United States

Dec. 2017 - May 2018

China Net Cloud (Shanghai) Co., Ltd

Project Manager

Shanghai, China

Mar. 2014 - Jun. 2016

TALKS

Computational Semantics via Lexical Concepts

Alberta Machine Intelligence Institute (Amii) AI Seminar

Sep. 20th, 2024

Inspiring Natural Language Processing with Human Cognition

Alberta Machine Intelligence Institute (Amii) AI Seminar

Dec. 23rd, 2022

Text Editing as Imitation Game

the Second GEM in EMNLP 2022

Dec. 7th 2022

Recurrent Inference in Text Editing

the Fourth SPNLP in EMNLP 2020

Nov. 20th 2020

SERVICES

²mila.quebec/en/news/ai4goodlab

Volunteer: ARR Support Team, ACL 2023

Reviewer

Journal: Electronic Commerce Research, Information Processing and Management

Conference: AAAI 2025, ACL 2023-2025, AISTATS 2025, ARR 2021-2025, COLM 2024-2025, EMNLP 2022-2025, ICLR 2024-2025, ICML 2022-2025, NeurIPS 2022-2025, RLC 2024-2025

Workshop: ENLSP 2023-2025, SemEval 2023-2025, WiNLP 2024

Secondary Reviewer

Journal: Transactions of the Association for Computational Linguistics

Conference: Canadian AI 2023-2025, COLING 2025, LREC-COLING 2024, EMNLP-IJCNLP 2019

Workshop: CAWL 2023, SemEval 2024

AWARDS

Graduate Student Teaching Award

University of Alberta

Apr. 8th, 2025

Upper Bound Talent Bursary

Alberta Machine Intelligence Institute (Amii)

Jan. 29th, 2025

Alberta Graduate Excellence Scholarship

Alberta, Canada

Dec. 6th, 2024

Reverse EXPO Best Poster Award

Department of Computing Science, University of Alberta

Sep. 6th, 2024

Computing Science Graduate Award for Parents

Department of Computing Science, University of Alberta

Sep. 3rd, 2024

Upper Bound Talent Bursary

Alberta Machine Intelligence Institute (Amii)

Feb. 29th, 2024

Alberta Graduate Excellence Scholarship

Alberta, Canada

Nov. 29th, 2023

GSA Academic Travel Grant

Graduate Students' Association, University of Alberta

Jul. 27th, 2023

Full Doctoral Scholarship

Department of Computing Science, University of Alberta

Sep. 2022 - (May 2026)

Graduate Research Assistantship Fellowship

Department of Computing Science, University of Alberta

Sep. 2022 - (May 2026)