

He Bai (Richard)

200 Ring Rd, Waterloo

ON, Canada N2L 3G1

☎ (+1) 548 333 6369

✉ he.bai@uwaterloo.ca

🌐 <https://uwaterloo.ca/scholar/h32bai>

Education

- Start from **University of Waterloo**,
09/2019 *Computer Science*, PhD,
Supervisor: Ming Li.
- 09/2017– **Institute of Automation, Chinese Academy of Sciences**,
07/2019 *Pattern Recognition and Intelligent System*, Master of Science,
Supervisor: Yu Zhou, Chengqing Zong.
- 09/2012– **Harbin Institute of Technology**,
07/2016 *Control Science and Engineering*, Bachelor of Engineering.

Research Experience

- May 2022– **ML Research Intern**, APPLE, SCV.
Aug. 2022 ○ LLM pretraining.
- May 2021– **Research Intern**, MICROSOFT RESEARCH, MONTREAL.
Jul. 2021 ○ Language model trained with WordNet knowledge.
- Jul. 2018– **NLP Research Intern**, IBM RESEARCH, BEIJING.
Oct. 2018 ○ Chinese Openai-GPT model pre-training.

Publications

- He Bai, Renjie Zheng, Xintong Li, Junkun Chen, Liang Huang. *A³T: Alignment-Aware Acoustic and Text Pretraining for Speech Synthesis and Editing*. *ICML 2022 (spotlight paper)*.
- He Bai, Tong Wang, Alessandro Sordani, Peng Shi. *Better Language Model with Hypernym Class Prediction*. *ACL 2022 (full paper)*.
- Peng Shi, Rui Zhang, He Bai, Jimmy Lin. *Cross-Lingual Training with Dense Retrieval for Document Retrieval*. *EMNLP-MSR 2021 (workshop paper)*.
- He Bai, Peng Shi, Jimmy Lin, Luchen Tan, Kun Xiong, Wen Gao, Jie Liu, Ming Li. *Semantics of the Unwritten: The Effect of End of Paragraph and Sequence Tokens on Text Generation*. *ACL-SRW 2021 (workshop paper)*.
- He Bai, Peng Shi, Jimmy Lin, Yuqing Xie, Luchen Tan, Kun Xiong, Wen Gao, Ming Li. *Segatron: Segment-aware Transformer for Language Modeling and Understanding*. *AAAI 2021 (full paper)*.
- Peng Shi, He Bai, Jimmy Lin. *Cross-Lingual Training of Neural Models for Document Ranking*. *Findings of EMNLP 2020 (short paper)*.
- He Bai, Yu Zhou, Jiajun Zhang and Chengqing Zong. *Memory Consolidation for Contextual Spoken Language Understanding with Dialogue Logistic Inference*. *ACL 2019 (short paper)*.
- He Bai, Yu Zhou, Jiajun Zhang, Liang Zhao, Mei-Yuh Hwang and Chengqing Zong. *Source Critical Reinforcement Learning for Transferring Spoken Language Understanding to a New Language*. *COLING 2018 (full paper)*.

Program Committees and Peer Review

- 2022 **Outstanding Reviewer (top 10%) of the International Conference of Machine Learning (ICML)**.
- 2021-now **Reviewer of Association for Computational Linguistics Rolling Review (ARR)**.

2020, 2022 **Reviewer of the International Conference on Computational Linguistic (COLING).**
2022 **Reviewer of the Workshop on Multilingual Information Access (MIA).**
2021,2020 **Reviewer of the Annual Meeting of the Association for Computational Linguistics (ACL).**
2022,2019 **Reviewer of the Conference on Empirical Methods in Natural Language Processing (EMNLP).**
2020 **Reviewer of the AAAI Conference on Artificial Interlligence (AAAI).**

Interests

Research Sequence Modeling, Unsupervised Learning, Multilingual NLP
Personal Yoga, Cooking