

# Meng Cao

☎ +1-438-722-6288    ✉ meng.cao@mail.mcgill.ca    🌐 mcao516    🔗 mcao516.github.io

## EDUCATION

---

**McGill University / Mila - Québec AI Institute** Jan. 2021 – Present  
Ph.D. in Computer Science GPA: 4.0/4.0  
Advisor: Jackie Chi Kit Cheung

**McGill University / Mila - Québec AI Institute** Sept. 2019 – Jan. 2021  
M.Sc. in Computer Science GPA: 4.0/4.0  
Advisor: Jackie Chi Kit Cheung

**Northeastern University (China)** Sept. 2015 – June 2019  
Major in Software Engineering GPA: 3.96/4.0  
Ranking: 1/59

## RESEARCH INTERESTS

---

Natural Language Processing, Text Summarization, Reinforcement Learning, Machine Learning

## PUBLICATIONS

---

- [1] Learning with Rejection for Abstractive Text Summarization  
**Meng Cao**, Yue Dong, Jingyi He and Jackie C. K. Cheung  
*EMNLP 2022 (To appear)*
- [3] Analyzing Multi-Sentence Aggregation in Abstractive Summarization via the Shapley Value  
Jingyi He, **Meng Cao** and Jackie C. K. Cheung  
*Findings of EMNLP 2022 (To appear)*
- [4] Hallucinated but Factual! Inspecting the Factuality of Hallucinations in Abstractive Summarization  
**Meng Cao**, Yue Dong and Jackie C. K. Cheung.  
*In Annual Meeting of the Association for Computational Linguistics (ACL), 2022*
- [5] A Survey on Neural Abstractive Summarization Methods and Factual Consistency of Summarization  
**Meng Cao**  
*ArXiv Preprint, 2021*
- [5] Factual Error Correction for Abstractive Summarization Models  
**Meng Cao**, Yue Dong, Jiapeng Wu and Jackie C. K. Cheung  
*In proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020*
- [6] TeMP: Temporal Message Passing for Temporal Knowledge Graph Completion  
Jiapeng Wu, **Meng Cao**, Jackie Chi Kit Cheung, William L. Hamilton  
*In proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020*
- [7] Referring Expression Generation Using Entity Profiles  
**Meng Cao** and Jackie Chi Kit Cheung  
*In proceedings of the Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019*
- [8] Transfer Learning for Cross-Domain Sequence Tagging Tasks  
**Meng Cao**, Chaohe Zhang, Dancheng Li, Qingping Zheng and Ling Luo  
*Future of Information and Communications Conference (FICC) 2018, San Francisco, March 14-15, 2019*

## INDUSTRY EXPERIENCES

---

### Microsoft Research Lab – Montréal

Machine Learning Research Intern

Apr. 2022 – Present

Advisor: Samira Shabanian & Mehdi Fatemi

- Mitigate toxicity in language model using reinforcement learning approach.

### Huawei Noah's Ark Lab – Canada, Montreal

Applied Machine Learning Research Intern

May 2021 – March 2022

Advisor: Mehdi Rezagholizadeh

- Knowledge distillation for text generation tasks
- Knowledge distillation with cluster labels

### Borealis AI – Canada, Toronto

Research Intern, Alan Team

May 2020 – August 2020

Advisor: Yanshuai Cao

- Build a distributed data-parallel Text-to-SQL system.
- Applied meta-learning algorithms (MAML, Reptile, MetaReg) in cross-database semantic parsing task.
- Achieved promising results in the zero-shot domain transfer setting.

### IBM Research – China, Beijing

Research Assistant Intern, Information Analytics Team

September 2018 – June 2019

Advisor: Shiwan Zhao

- Chinese word segmentation: responsible for training Chinese word segmentation model. Proposed a novel segmentation model based on n-gram and information entropy to segment the input sentences.
- Neural Architecture Search: PyTorch implementation of Google's ENAS neural architecture algorithm.
- Migrated a neural network model which is implemented in Keras into Java environment using Deeplearning4j.

## PRESENTATIONS AND TALKS

---

Factual Error Correction for Abstractive Summarization Models (Poster)  
EMNLP 2020

Nov., 2020  
Online

Referring Expression Generation Using Entity Profiles (Poster)  
EMNLP 2019

Nov., 2019  
Hong Kong, China

Transfer Learning for Cross-Domain Sequence Tagging Tasks (Oral)  
FICC 2018

March, 2019  
San Francisco, USA

## SCHOLARSHIPS

---

### China National Scholarship (¥ 16,000)

2015-2016, 2017-2018

The Chinese Government, among top-0.2% all undergraduates in China

### The Bao Gang Education Scholarship (¥ 10,000)

September 2018

China BAOWU Steel Group, selected rate: 4/20000 in NEU

### The First-Class Scholarship in Northeastern University (¥ 9,000)

2015-2016, 2016-2017, 2017-2018

Northeastern University, among top-3% undergraduates in NEU

## AWARDS

---

### "Top Ten Undergraduates" in Northeastern University

May 2018

Northeastern University, among top-0.2% undergraduates in NEU

### The Excellent Performance Award for IBM Undergraduate Innovation Lab Program

April 2017

IBM Research China, Beijing

### The First Prize in the 8th National College Students Mathematics Competition

October 2016

Chinese Mathematical Society

### The First Prize in the 9th National College Students Mathematics Competition

October 2017

Chinese Mathematical Society

<b>The Meritorious Winner for the Mathematical Contest in Modeling (MCM/ICM), 2018</b>	February 2018
COMAP, the Consortium for Mathematics and Its Applications (USA)	
<b>The Honorable Mention for the Mathematical Contest in Modeling (MCM/ICM), 2017</b>	February 2017
COMAP, the Consortium for Mathematics and Its Applications (USA)	

## RESEARCH FUNDING

---

<b>Mitacs Globalink Research Internship (\$19,500)</b>	June 2018 – September 2018
Mitacs Globalink summer research internship program (\$4,500) & Mitacs graduate student funding (\$15,000)	

## SKILLS

---

**Programming Languages:** skilled in: Java, Python (especially NumPy, PyTorch); familiar with: Tensorflow, JavaScript, SQL

**Extensive knowledge of:** Natural Language Processing, Abstractive Text Summarization, Machine Learning (generative models), Deep Learning