

# LINGJUN ZHAO

(1) 917-478-7713   ◇   lingjun.zhao@hotmail.com  
Cambridge MA

## EDUCATION

---

### University of Maryland, College Park

*Jan 2021 - present*

Ph.D. student in Computer Science

### Columbia University, New York

*Sep 2016 - Dec 2017*

M.S. in Computer Science · GPA: 3.5/4.0

*Excelled in Selected Coursework:* Natural Language Processing · Advanced Machine Learning · Deep Learning for Computer Vision · Advanced Database Systems

### Sun Yat-Sen University, Guangzhou, China

*Sep 2012 - Jun 2016*

B.S. in Computational Mathematics · GPA: 3.6/4.0

Visiting student at University of California, Berkeley

*Fall 2014*

*Excelled in Selected Coursework:* Multivariate Statistics · Mathematical Optimization · Numerical Linear Algebra · Real Analysis · Digital Image Processing

## EXPERIENCE

---

### Staff Scientist II

Mar 2018 - Dec 2020

*Raytheon BBN Technologies*

*Cambridge, MA*

- Designed and built a character-level CNN context-aware lexical translation model, effective for the probabilistic cross-lingual information retrieval (CLIR) in the IARPA MATERIAL program
- Designed and built a weakly supervised attentional model for CLIR, trained with samples extracted from parallel translation sentences to estimate CLIR relevance
- Developed a neural CLIR model using relevance outputs from the unsupervised probabilistic CLIR
- Developed a CLIR diagnostic tool and lexical translation evaluation metrics for CLIR
- Trained and tuned a siamese neural model for the event mention retrieval
- Developed a probabilistic event IR model for the IARPA BETTER program evaluation

### Data Science Intern

July 2017 - Aug 2017

*Pacific Northwest National Laboratory*

*Richland, WA*

- Developed neural relation extraction attentional models with LSTM and CNN for feature extraction
- Evaluated to have competitive results on both SemEval-2017 Task 10 and SemEval-2010 Task 8

### Research Assistant

Sep 2016 - Dec 2016

*Columbia University, Advisor: John Kender*

*New York, NY*

- Applied Viola-Jones face detection to generate upper body bounding boxes for upper body pose estimation in TED Talk videos. Derived upper bounding box scaling and filtered interesting frames
- Approximated the exponential distribution and matrix symmetry of interesting frames duration

### Bachelor's Thesis Project

Nov 2015 - May 2016

*Sun Yat-Sen University, Advisor: Guocan Feng*

*Guangzhou, China*

- Designed and built a CNN human face tracking model, applied to public surveillance videos

## PUBLICATIONS

---

B. Min, Y. Chan, **L. Zhao**. “Towards Few-Shot Event Mention Retrieval: An Evaluation Framework and A Siamese Network Approach.” LREC 2020. (link)

Z. Jiang, A. El-Jaroudi, W. Hartmann, D. Karakos, **L. Zhao**. “Cross-lingual Information Retrieval with BERT.” LREC 2020 Cross-Language Search and Summarization of Text and Speech Workshop. (link)

**L. Zhao**, R. Zbib, Z. Jiang, D. Karakos, Z. Huang. “Weakly Supervised Attentional Model for Low Resource Ad-hoc Cross-lingual Information Retrieval.” EMNLP 2019 Deep Learning Approaches for Low-Resource NLP Workshop. (link)

R. Zbib, **L. Zhao**, D. Karakos, W. Hartmann, J. DeYoung, Z. Huang, Z. Jiang, N. Rivkin, L. Zhang, R. Schwartz, J. Makhoul. “Neural-Network Lexical Translation for Cross-lingual IR from Text and Speech.” SIGIR 2019. (link)

## AWARDS

---

- e Prize & First Prize (top 3%), Asia Student Supercomputer Challenge, 2015
- Scholarship for Outstanding Students (top 15%), Sun Yat-Sen University, 2015
- Second Place (top 3%), Mathematical Contest in Modeling, Sun Yat-Sen University, 2012

## SERVICES

---

- Volunteer of Computer Science graduate admissions, University of Maryland, 2020
- Reviewer of ACM TALLIP 2019, Secondary Reviewer of EMNLP 2018
- Director of Computational Science Society, Sun Yat-Sen University, 2013

## SKILLS

---

<b>Programming Languages</b>	PYTHON, C++, C, MATLAB, PERL
<b>Machine Learning Libraries</b>	Tensorflow, Keras, Pytorch, scikit-learn, NLTK
<b>Other Tools</b>	Git, Vim, Jupyter, CUDA, NumPy, LATEX