

Zhengbao Jiang

Education

- 2018.8–2024.5 **Ph.D.**, *Carnegie Mellon University*, Language Technologies Institute, *GPA: 4.1/4.3*.
(expected) Advisor: Graham Neubig.
Area: Natural Language Processing, Information Retrieval, Machine Learning.
- 2015.9–2018.6 **M.S.**, *Renmin University of China*, School of Information, *GPA: 3.85/4*.
Advisor: Zhicheng Dou, Wayne Xin Zhao, Jian-Yun Nie, Ji-Rong Wen.
Area: Information Retrieval, Natural Language Processing, Data Mining.
- 2011.9–2015.6 **B.Eng.**, *Renmin University of China*, School of Information, *GPA: 3.81/4*.
Advisor: Zhicheng Dou, Ji-Rong Wen.
Major: Computer Science.

Research Interests

- RAG [FLARE](#) [ReAtt](#), Nonparametric LLMs through retrieval-augmented generation.
- Alignment [pre-instruction-tuning](#) [QA calibration](#), Alignment/Calibration for factuality and truthfulness.
- Prompting [LPAQA](#) [X-FACTR](#), Eliciting factual knowledge from LMs through prompt learning/engineering.
- Reasoning [OmniTab](#) [multihop QA](#), Improving reasoning capacity through synthetic pre-training.

Experience

- 2023.6–2023.12 **Research Intern**, *FAIR at Meta*.
Mentor: Srini Iyer, Scott Wen-tau Yih, Xi Victoria Lin, Pedro Rodriguez.
Research: knowledge updates on LLMs.
- 2022.6–2022.12 **Research Intern**, *FAIR at Meta*.
Mentor: Fabio Petroni, Jane Yu, Timo Schick, Patrick Lewis.
Research: unified retrieval-augmented LLMs.
- 2021.6–2021.12 **Research Intern**, *Microsoft Azure AI*.
Mentor: Yi Mao, Pengcheng He, Weizhu Chen.
Research: omnivorous pre-training for table-based QA to improve both grounding and reasoning.
- 2020.6–2020.12 **Research Intern**, *Amazon*.
Mentor: Jialong Han, Xin Luna Dong.
Research: open information extraction and integration.
- 2017.10–2018.3 **Research Intern**, *Alibaba AI Labs*.
Mentor: Zaiqing Nie.
Research: probabilistic word embeddings and text generation.
- 2017.7–2017.10 **Research Intern**, *Microsoft Research Asia*.
Mentor: Zaiqing Nie.
Research: multi-sense word embeddings.
- 2014.2–2014.12 **Software Engineer**, *OfferCalendar*.
Full-stack developer for a startup focusing on job search.

Publication (* indicates equal/lead contribution)

- [arXiv'24](#) **Instruction-tuned Language Models are Better Knowledge Learners.**
Zhengbao Jiang, Zhiqing Sun, Weijia Shi, Pedro Rodriguez, Chunting Zhou, Graham Neubig, Xi Victoria Lin, Scott Wen-tau Yih, Srinivasan Iyer

- [EMNLP'23](#) **Active Retrieval Augmented Generation.**
Zhengbao Jiang*, Frank F. Xu*, Luyu Gao*, Zhiqing Sun*, Qian Liu, Jane Dwivedi-Yu, Yiming Yang, Jamie Callan, Graham Neubig
- [arXiv'23](#) **From Zero to Hero: Examining the Power of Symbolic Tasks in Instruction Tuning.**
Qian Liu*, Fan Zhou*, Zhengbao Jiang, Longxu Dou, Min Lin
- [EMNLP'22](#) **Retrieval as Attention: End-to-end Learning of Retrieval and Reading within a Single Transformer.**
oral
Zhengbao Jiang*, Luyu Gao*, Jun Araki, Haibo Ding, Zhiruo Wang, Jamie Callan, Graham Neubig
- [arXiv'23](#) **GPTScore: Evaluate as You Desire.**
Jinlan Fu, See-Kiong Ng, Zhengbao Jiang, Pengfei Liu
- [ICLR'23](#) **PEER: A Collaborative Language Model.**
Timo Schick, Jane Dwivedi-Yu, Zhengbao Jiang, Fabio Petroni, Patrick Lewis, Gautier Izacard, Qingfei You, Christoforos Nalmpantis, Édouard Grave, Sebastian Riedel
- [arXiv'23](#) **EditEval: An Instruction-Based Benchmark for Text Improvements.**
Jane Dwivedi-Yu, Timo Schick, Zhengbao Jiang, Maria Lomeli, Patrick Lewis, Gautier Izacard, Édouard Grave, Sebastian Riedel, Fabio Petroni
- [ICLR'23](#) **DocPrompting: Generating Code by Retrieving the Docs.**
Shuyan Zhou, Uri Alon, Frank F. Xu, Zhiruo Wang, Zhengbao Jiang, Graham Neubig
- [NAACL'22](#) **OmniTab: Pretraining with Natural and Synthetic Data for Few-shot Table-based Question Answering.**
oral
Zhengbao Jiang, Yi Mao, Pengcheng He, Graham Neubig, Weizhu Chen
- [ACM CSUR'22](#) **Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing.**
Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, Graham Neubig
- [COLING'22](#) **Understanding and Improving Zero-shot Multi-hop Reasoning in Generative Question Answering.**
oral
Zhengbao Jiang, Jun Araki, Haibo Ding, Graham Neubig
- [NAACL'22](#) **Table Retrieval May Not Necessitate Table-specific Model Design.**
SUKI workshop
Zhiruo Wang, Zhengbao Jiang, Eric Nyberg, Graham Neubig
- [ACL'21](#) **CoRI: Collective Relation Integration with Data Augmentation for Open Information Extraction.**
Zhengbao Jiang, Jialong Han, Bunyamin Sisman, Luna Xin Dong
- [TACL'21](#) **How Can We Know When Language Models Know? On the Calibration of Language Models for Question Answering.**
Zhengbao Jiang, Jun Araki, Haibo Ding, Graham Neubig
- [NAACL'21](#) **GSum: A General Framework for Guided Neural Abstractive Summarization.**
Zi-Yi Dou, Pengfei Liu, Hiroaki Hayashi, Zhengbao Jiang, Graham Neubig
- [EMNLP'20](#) **X-FACTR: Multilingual Factual Knowledge Retrieval from Pretrained Language Models.**
Zhengbao Jiang*, Antonios Anastasopoulos*, Jun Araki, Haibo Ding, Graham Neubig
- [TACL'20](#) **How Can We Know What Language Models Know? .**
oral
Zhengbao Jiang*, Frank F. Xu*, Jun Araki, Graham Neubig
- [ACL'20](#) **Generalizing Natural Language Analysis through Span-relation Representations.**
Zhengbao Jiang, Wei Xu, Jun Araki, Graham Neubig
- [ACL'20](#) **Incorporating External Knowledge through Pre-training for Natural Language to Code Generation.**
short
Frank F. Xu*, Zhengbao Jiang*, Pengcheng Yin, Bogdan Vasilescu, Graham Neubig
- [ECML'20](#) **Graph-Revised Convolutional Network .**
Donghan Yu, Ruohong Zhang, Zhengbao Jiang, Yuexin Wu, Yiming Yang

- [AKBC'20](#) **Learning Relation Entailment with Structured and Textual Information.**
Zhengbao Jiang, Jun Araki, Donghan Yu, Ruohong Zhang, Wei Xu, Yiming Yang, Graham Neubig
- [ACL'19](#) **Improving Open Information Extraction via Iterative Rank-Aware Learning.**
short Zhengbao Jiang, Pengcheng Yin, Graham Neubig
- [CIKM'18](#) **Personalizing Search Results Using Hierarchical RNN with Query-aware Attention.**
Songwei Ge, Zhicheng Dou, Zhengbao Jiang, Jian-Yun Nie, Ji-Rong Wen
- [TKDE'18](#) **Supervised Search Result Diversification via Subtopic Attention.**
Zhengbao Jiang, Zhicheng Dou, Wayne Xin Zhao, Jian-Yun Nie, Ming Yue, Ji-Rong Wen
- [SIGIR'17](#) **Learning to Diversify Search Results via Subtopic Attention.**
Zhengbao Jiang, Ji-Rong Wen, Zhicheng Dou, Wayne Xin Zhao, Jian-Yun Nie, Ming Yue
- [TKDE'17](#) **Generating Query Facets Using Knowledge Bases.**
Zhengbao Jiang, Zhicheng Dou, Ji-Rong Wen
- [TKDE'16](#) **Automatically Mining Facets for Queries from Their Search Results.**
Zhicheng Dou, Zhengbao Jiang, Sha Hu, Ji-Rong Wen, Ruihua Song

Skills

Programming Python, JavaScript, Java, C#, C++, C, HTML, Matlab
Language Chinese: native
English: fluent, TOEFL 104 (R28, L25, S26, W25) GRE 326 (V156, Q170) + 3.5

Awards

- 2017 **Wu Yuzhang Scholarship, 10/24000.**
Summa cum laude of Renmin University of China
- 2016, 2017 **National Scholarship for Graduates.**
- 2016 **Kaggle Avito Duplicate Ads Detection, 16th/548.**
Detect duplicate ads using both text and image information.
- 2015 **Excellent Graduate Student Award of Renmin University of China.**
- 2014 **Kwang-Hua Scholarship.**
- 2013 **First-Class Scholarship of Renmin University of China.**

Professional Activities

Journal Reviewer TKDE, TIST
Conf Reviewer SIGIR'16, SIGIR'17, CIKM'17, WSDM'17, AAAI'18, SIGMOD'18, ACL'21, EMNLP'21, CIKM'21, WSDM'22, ACL'22, EMNLP'22, COLING'22, ACL'23, ARR