

Jie Cao

Ph.D. Student | School of Computing | University of Utah
<https://mlciv.com> | jcao@cs.utah.edu | (801) 448-5203

EDUCATION

2015 – Present	Doctor of Philosophy COMPUTER SCIENCE <i>University of Utah</i>
2009 – 2012	Master of Science COMPUTER SCIENCE <i>Huazhong University of Science and Technology (HUST)</i>
2005 – 2009	Bachelor of Engineering INFORMATION SECURITY <i>Huazhong University of Science and Technology (HUST)</i>

RESEARCH EXPERIENCE

NLP LAB, UNIVERSITY OF UTAH, SLC	2015 – PRESENT
– I work with Prof. Vivek Srikumar in Utah NLP Lab. I study methods for predicting structured representations of natural language text, e.g. Semantic Parsing [4], Dialogue System [6, 3, 1].	
Intern at AWS AI, AMAZON LEX, Remote	SUMMER 2020
– Our paper on schema-guided dialog got accepted to NAACL 2021 [1]	
Intern at AWS AI, AMAZON LEX, Seattle	SUMMER 2019
– Cross framework meaning representation parsing, ranked 1st on AMR on CoNLL 2019 shared task [4]	
Intern at TENCENT, WECHAT AI, Palo Alto	SUMMER 2018
– End-to-End Goal-oriented Dialogue System [6], ranked 2nd in DSTC7 track1	
– Rhetorically Controlled Poetry Generation[5] got accepted to ACL 2019	
CGCL LAB, HUST, Wuhan	2008 – 2012
– When I was a master student at HUST, I worked with Prof. Xia Xie and Prof. Hai Jin. My research are around large scale statistical computing, programming language [7, 8], system virtualization [10, 9]	

WORK EXPERIENCE

Assitant Researcher at SOHU RDC LAB, Beijing	2014 – 2015
– Hadoop, Spark, Data migration, Data security, Distributed machine learning.	
Senior Software Engineer at ZUN CLUB, Beijing	2013 – 2014
– Heterogeneous data intergration, Hotel recommendation system.	
Software Engineer at BAIDU, Beijing	2012 – 2014
– Baidu Voice Assistant, Query analysis, Dialogue	
– Mobile Search Anti-Attack Ecosystem, Speed optimization	
Intern at Alibaba, Hangzhou	2010 – 2011
– KV Storage, MySQL, Database Replication, Real-time Computing, Distributed Pub/Sub Data Pipeline.	

TEACHING & MENTORING

2019-2020	Mentor	Bachelor Thesis (Tarun Sunkaraneni), U of Utah
Spring 2018	TA	Structured Prediction, U of Utah
Fall 2016, 2018	TA	Machine Learning, U of Utah
2006-2008	Leader	Algorithm&Game Team, Unique Studio, HUST

HONORS & AWARDS

2019	CoNLL Shared Task, MRP, Top 1 on AMR
2019	DSTC7, Top 2 on track 1
2010	VMware Cloud Computing Innovation Cup, Top 50
2009	Google Android Innovative Idea Sharing Award
2007	“Computer World” Magazine Scholarship
2007	Microsoft Imagine Cup
	– Algorithm Challenge, Top 50 World-Wide
	– Visual Gaming Contest, Top 2 in China, 18th World Final
2006	HUST ACM Programming Contest, Top 3

ACADEMIC SERVICES

PC Member / Reviewers for ACL’20, 21, EACL’20, EMNLP’20, NAACL’21, CoNLL’19, COLING’20, AAAI’19, 20, MRP’2019, etc.

PUBLICATIONS

- [1] **Jie Cao** and Yi Zhang. “A Comparative Study on Schema-Guided Dialogue State Tracking”. In: *Proceedings of NAACL 2021, To Appear*. 2021.
- [2] Debjyoti Paul*, Jie Cao*, Feifei Li, and Vivek Srikumar. “Database Workload Characterization with Query Plan Encoders”. In: *arXiv preprint arXiv:2105.12287* (2021).
- [3] **Jie Cao**, Michael Tanana, Zac Imel, Eric Poitras, David Atkins, and Vivek Srikumar. “Observing Dialogue in Therapy: Categorizing and Forecasting Behavioral Codes”. In: *Proceedings of ACL 2019*. Florence, Italy, 2019.
- [4] **Jie Cao**, Yi Zhang, Adel Youssef, and Vivek Srikumar. “Amazon at MRP 2019: Parsing Meaning Representations with Lexical and Phrasal Anchoring”. In: *Proceedings of the Shared Task on MRP at the CoNLL 2019*. 2019.
- [5] Zhiqiang Liu, Zuohui Fu, **Jie Cao**, Gerard de Melo, Yik-Cheung Tam, Cheng Niu, and Jie Zhou. “Rhetorically Controlled Encoder-Decoder for Modern Chinese Poetry Generation”. In: *Proceedings of ACL 2019*. Florence, Italy, 2019.
- [6] Shuo Sun*, Yik-Cheung Tam*, **Jie Cao***, Can-Xiang Yan, Zuohui Fu, Cheng Niu, and Jie Zhou. “End-to-end Gated Self-attentive Memory Network for Dialog Response Selection”. In: *Proceedings of AAAI, DSTC7 Workshop*. 2019.
- [7] Xijiang Ke, Hai Jin, Xia Xie, and **Jie Cao**. “A distributed SVM method based on the iterative MapReduce”. In: *International Conference on Semantic Computing (ICSC)*. IEEE. 2015.
- [8] **Jie Cao***, Xia Xie*, Hai Jin, Xijiang Ke, and Wenzhi Cao. “JRBridge: A framework of large-scale statistical computing for R”. In: *Asia-Pacific on Services Computing Conference (APSCC)*. IEEE. 2012.
- [9] **Jie Cao**, Xia Xie, and Jin Hai. “A real-time scheduling system on embedded virtualization”. Patent ZL201110410689.1 (CN). 2011.
- [10] **Jie Cao**. “Cache Performance Evaluation on Xen Virtualization Platform”. In: *Undergraduate Thesis*. 2009.