EDUCATION

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

RESEARCH EXPERIENCE

UTAH NLP GROUP, UNIVERSITY OF UTAH, Salt Lake City   2015 - PRESENT
    Research Space

– I work with Prof. Vivek Srikumar in Utah NLP Group. 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 accetped 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 somputing, 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.

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" Magzine 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

PUBLICATIONS

[1] **Jie Cao** and Yi Zhang. "A Comparative Study on Schema-Guided Dialogue State Tracking". In: *Proceedings of NAACL 2021*. 2021.

[2] Debjyoti Paul*, **Jie Cao***, Feifei Li, and Vivek Srikumar. "Database Workload Characterization with Query Plan Encoders". In: *arXiv preprint arXiv:2105.12287(To Appear in VLDB'22)* (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 embeded virtualization". Patent ZL201110410689.1 (CN). 2011.

[10] **Jie Cao**. "Cache Performance Evaluation on Xen Virtualization Platform". In: *Undergraduate Thesis*. 2009.

TALKS

Slides for selected talks can be found in my slideshare site. https://www.slideshare.net/jiessiecao/

– 05/2021, Invited Talk on CS 501 Seminar:Contemporary Issues at University of Idaho, 'Inductive Biases in Deep Linguistic Structured Prediction'

– 06/2021, Conference Presentation on NAACL'2021, 'A Comparative Study on Schema-Guided Dialogue State racking'

– 12/2020, Talk on EMNLP'2020 Watch Party@Amazon Lex, 'Task-oriented Conversational Semantic Parsing'

– 11/2019, Conference Presentation on MRP 2019, 'Amazon at MRP 2019: Parsing Meaning Representations with Lexical and Phrasal Anchoring'

– 01/2016, Seminar Presentation on Spring'2016 NLP seminer at University of Utah, 'Combinatory Categorical Grammar'

– 11/2015, Seminar Presentation on Fall'2015 Deep Learning seminer at University of Utah, 'Parsing Natural Language Scense and Natural Language with Recursive Neural Networks'

– 11/2015, Seminar Presentation on Fall'2015 Data Science Seminer at University of Utah, 'Challenges on Data Parallelism and Model Parallelism'

– 12/2015, Winner Presentation on Fall'2015 NLP Class at University of Utah, 'Talking Geckos QA System', our QA system ranked 1st in QA competition.

SOFTWARE

Other code for papers can be found in https://github.com/utahnlp and https://github.com/mlciv

– LAPA-MRP, our open-sourced system submission to MRP 2019 shared task at CoNLL 2019. Our model ranked 1st in the AMR subtask, 5th in UCCA, 6th in PSD and 7th in DM.

– dt-proxy, A nodejs-based token proxy to support distributed data copy between secure and insecure hadoop clusters.

– alred-workflows, A home-brew repo for productive alfred workflows, including propbank query, display rotation, and so on.

– renjin, a JVM-based interpreter for the R language. https://github.com/bedatadriven/renjin

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

ACADEMIC SERVICE

Program Committee Member

– Conference of the Association for the Advancement of Artificial Intelligence (AAAI), 2020–22;

– Annual meeting of the Association of Computational Linguistics (ACL), 2020-22;

– International Conference on Computational Linguistics (COLING), 2020;

– Conference of the European Chapter of the ACL (EACL), 2021;

– Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020-2022;

– North American Chapter of the ACL (NAACL), 2021;

– Shared Task on Cross-Framework Meaning Representation Parsing at the 2019 Conference on Natural Language Learning (MRP'2019@CoNLL).

Reviwer

– ACL Rolling Review, 2021-22