

## Zhengzhong Liu

---

CONTACT INFORMATION	Room 5507 Gates Hillman Complex 5000 Forbes Avenue Pittsburgh, PA, 15213	<i>E-mail:</i> liu@cs.cmu.edu <i>Mobile:</i> 412 980 0646
RESEARCH INTERESTS	I am broadly interested in areas such as Computational Linguistics, Natural Language Processing, Information Retrieval, Information Extraction. I am particularly interested in injecting linguistic and semantic regularities into machine learning models, and frames to support such research.	
EDUCATION	<b>Carnegie Mellon University</b>	<b>Sep, 2014 — now</b>
	- PhD Candidate in Language Technologies - Thesis Topic: Event and Facet Semantics	
	<b>Carnegie Mellon University</b>	<b>2012 — 2014</b>
	- Master in Language Technologies	
RESEARCH EXPERIENCES	<b>The Hong Kong Polytechnic University</b>	<b>2007 — 2012</b>
	- Bachelor of Science in Computing (with First Class Honors) - Bachelor of Business Administration with a Major in Management - Engineering School, Jonköping University (Exchange)	
	<b>Active Interpretation of Disparate Alternatives (AIDA)</b>	<b>September, 2018 — present</b>
	- Advisors: Prof. Eduard Hovy, Prof. Teruko Mitamura - The AIDA project aims at overcome the noisy, conflicting, and potentially intentionally deceptive nature of today's data environment via advanced NLP technologies. Our group focuses on interpreting the alternatives of events and participants from text. We have advanced methods to identify the core events and participatns from news articles.	
WORKING EXPERIENCES	<b>Deep Exploration and Filtering of Text (DEFT)</b>	<b>September, 2012 — 2018</b>
	- Advisors: Prof. Eduard Hovy, Prof. Teruko Mitamura - The DEFT project aims at discovering deeper and implicit information from text using NLP and AI techniques. Our group focuses on understanding event semantic in text. We develop automated systems in detecting various types of relations among event mentions.	
	<b>Entity Linking and Disambiguation for Knowledge Based Population</b>	<b>June, 2011 — October, 2011</b>
	- Advisor: Prof. Qin Lu - Leading the entity linking task and collaborate on the slot filling task. Our team develop methods to link text spans to the appropriate Wikipedia entry and propose a disambiguation method to cluster unknown entities.	
WORKING EXPERIENCES	<b>Web Person Name Disambiguation (Undergraduate dissertation)</b>	<b>June, 2010 — June, 2011</b>
	- Advisor: Prof. Qin Lu - Propose a novel tailored clustering algorithm for the Web Person Name Disambiguation task, which achieved the state-of-the-art performance in the Web Person Name Disambiguation task. - Best undergraduate project award nomination in Department of Computing, the Hong Kong Polytechnic University in academic year 2010/11.	
	<b>Research Scientist at Petuum</b>	<b>May, 2019 — Now</b>
	- Leading the research and development of an advanced Natural Language Processing system.	

## Research Intern at Google

May, 2015 — August, 2015

- Mentors: Edgar González Pellicer *and* Daniel Gillick
- Worked with the Google NLP team the Verb Phrase Ellipsis (VPE) problem, which is used to help improve a task-based dialogue systems.
- Study the multi-stages of the VPE problem, propose several strategies for joint-training and achieve the state-of-the-art performance on two public datasets. - We clean and contribute one of the datasets and release it to the community.

## Google Summer of Code

May, 2012 - August, 2012

- Mentor: Pablo N. Mendes
- Worked on the *DBpedia Spotlight* for DBpedia at Google Summer of Code program 2012
- Developed and delivered a collective entity linking system which considers the coherence of all entities in the document.
- Code base: <https://github.com/hunterhector/dbpedia-spotlight>

## OPEN SOURCE EXPERIENCES

### Forte

- I lead the development of an NLP toolkit named Forte. I designed the system architecture and representations for the system. The project will be released soon.

### Texar

- I contributed to the Texar project, which is a toolkit for Text Generation and other NLP applications. I am involved in the project design and implementations of several fundamental modules.

### OAQA

- I contributed to OAQA (Open Advancement of Question Answering Systems) by helping develop an system to answer World History questions.)

### DBpedia Spotlight

- I contributed to a popular Knowledge base project DBpedia. I developed a collective disambiguation module in project DBpedia Spotlight using Scala and Hadoop.

## HONORS AND AWARDS

### ACL Best System Demonstration Nomination

2019

- Our system Texar is selected to be one of the top five system demonstration at ACL.

### ACL Outstanding Long Paper Award

2016

- Awarded to the 10 best paper on ACL, less than 1% of the submitted paper are selected.

### The Zhi Yuan Scholarship of China Soong Ching Ling Foundation

2007 - 2011

- For excellent mainland China student pursuing education in Hong Kong

## PUBLICATIONS

1. Zhiting Hu, Zichao Yang, Haoran Shi, Bowen Tan, Tiancheng Zhao, Junxian He, Xiaodan Liang, Wentao Wang, Xingjiang Yu, Di Wang, Lianhui Qin, Xuezhe Ma, **Zhengzhong Liu**, Devendra Singh, Wangrong Zhu, and Eric P Xing. Texar: A Modularized, Versatile, and Extensible Toolkit for Text Generation. ACL 2019, *best system demonstration nomination*
2. Zhiting Hu, Zichao Yang, Haoran Shi, Bowen Tan, Tiancheng Zhao, Junxian He, Xiaodan Liang, Wentao Wang, Xingjiang Yu, Di Wang, Lianhui Qin, Xuezhe Ma, **Zhengzhong Liu**, Devendra Singh, Wangrong Zhu, and Eric P Xing. 2018. Texar: A Modularized, Versatile, and Extensible Toolbox for Text Generation. In *Proceedings of Workshop for NLP Open Source Software, ACL'18*.
3. **Zhengzhong Liu**, Teruko Mitamura, and Eduard Hovy. 2018. Graph-Based Decoding for Event Sequencing and Coreference Resolution. In *Proceedings of the 27th International Conference on Computational Linguistics (COLING'18)*.
4. **Zhengzhong Liu**, Chenyan Xiong, Teruko Mitamura, and Eduard Hovy. 2018. Automatic Event Salience Identification. In *Proceedings of the 2018 Conference on Empirical Methods in*

5. Hans Chalupsky, Jun Araki, Eduard Hovy, Andrew Hsi, **Zhengzhong Liu**, Xuezhe Ma, Evangelia Spiliopoulou, and Shuxin Yao. 2018. Multi-lingual Extraction and Integration of Entities, Relations, Events and Sentiments into ColdStart ++ KBs with the SAFT System. In *Proceedings of the Text Analysis Conference 2017 (TAC'17)*.
6. Teruko Mitamura, **Zhengzhong Liu**, and Eduard Hovy. 2018. Events Detection, Coreference and Sequencing: What's next? Overview of the TAC KBP 2017 Event Track. In *Proceedings of the Text Analysis Conference 2017 (TAC'17)*.
7. Chenyan Xiong, **Zhengzhong Liu**, Jamie Callan, and Tie-Yan Liu. 2018. Towards Better Text Understanding and Retrieval through Kernel Entity Saliency Modeling. In *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'18)*.
8. Chenyan Xiong, **Zhengzhong Liu**, Jamie Callan, and Eduard Hovy. 2017. JointSem: Combining Query Entity Linking and Entity based Document Ranking. In *Proceedings of the 26th ACM International Conference on Information and Knowledge Management (CIKM'17)*.
9. Keyang Xu, **Zhengzhong Liu**, and Jamie Callan. 2017. De-duping URLs with Sequence-to-Sequence Neural Networks. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'17)*.
10. Teruko Mitamura, **Zhengzhong Liu**, and Eduard Hovy. 2017. Overview of TAC-KBP 2016 Event Nugget Track. In TAC 2016. **Zhengzhong Liu**, Jun Araki, Teruko Mitamura, and Eduard Hovy. 2016. CMU-LTI at KBP 2016 Event Nugget Track. In *Proceedings of the Text Analysis Conference 2016 (TAC'16)*.
11. Zhiting Hu, Xuezhe Ma, **Zhengzhong Liu**, Eduard Hovy, and Eric Xing. 2016. Harnessing Deep Neural Networks with Logic Rules. In *Proceeding of 2016 Annual Meeting of the Association for Computational Linguistics (ACL'16)*.
12. Xuezhe Ma, **Zhengzhong Liu**, and Eduard Hovy. 2016. Unsupervised Ranking Model for Entity Coreference Resolution. In *The 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL'16)*.
13. **Zhengzhong Liu**, Edgar González, and Dan Gillick. 2016. Exploring the steps of Verb Phrase Ellipsis. In *NAACL 2016 Workshop on Coreference Resolution Beyond OntoNotes (CORBON'16)*.
14. **Zhengzhong Liu**, Jun Araki, Teruko Mitamura, and Eduard Hovy. 2016. CMU-LTI at KBP 2016 Event Nugget Track. In *Proceedings of the Text Analysis Conference 2016 (TAC'16)*.
15. **Zhengzhong Liu**, Jun Araki, Dheeru Dua, Teruko Mitamura, and Eduard Hovy. 2015. CMU-LTI at KBP 2015 Event Track. In *Proceedings of the Text Analysis Conference 2015 (TAC'15)*.
16. Teruko Mitamura, **Zhengzhong Liu**, and Eduard Hovy. 2016. Overview of TAC KBP 2015 Event Nugget Track. In *Proceedings of the Text Analysis Conference 2015 (TAC'15)*.
17. **Zhengzhong Liu**, Teruko Mitamura, and Eduard Hovy. 2015. Evaluation Algorithms for Event Nugget Detection: A Pilot Study. In *Proceedings of the 3rd Workshop on EVENTS, NAACL-HLT'15*, pages 53–57.
18. Di Wang, Lenoid Boytsov, Jun Araki, Alkesh Patel, Jeff Gee, **Zhengzhong Liu**, Eric Nyberg, and Teruko Mitamura. 2014. CMU Multiple-choice Question Answering System at NTCIR-11 QA-Lab. In *Proceedings of the 11th NTCIR Conference (NTCIR'14)*.
19. **Zhengzhong Liu**, Jun Araki, Eduard Hovy & Teruko Mitamura. 2014. Supervised Within-Document Event Coreference using Information Propagation. In *Proceedings of The Ninth edition of International Conference on Language Resources and Evaluation (LREC'14)*.

20. Jun Araki, **Zhengzhong Liu**, Eduard Hovy & Teruko Mitamura. 2014. Detecting Subevent Structure for Event Coreference Resolution. In *Proceedings of The Ninth edition of International Conference on Language Resources and Evaluation (LREC'14)*.
21. Xu, J., Lu, Q., & **Zhengzhong Liu**. 2012. Aggregating skip bigrams into key phrase-based vector space model for web person disambiguation. In *Proceedings of KONVENS'12*.
22. Xu, J., Lu, Q., **Zhengzhong Liu**, & Chai, Junyi. 2012. Topic Sequence Kernel. In H. Yuexian, N. Jian-Yun, S. Le, W. Bo, & Z. Peng (Eds.), *Lecture Notes in Computer Science*.
23. Xu, J., Lu, Q., & **Zhengzhong Liu**. 2012. PolyUCOMP: Combining semantic vectors with skip bigrams for semantic textual similarity. In *SemEval'12 Proceedings of the First Joint Conference on Lexical and Computational Semantics*.
24. Xu, J., Lu, Q., & **Zhengzhong Liu**. 2012. Combining Classification with Clustering for Web Person Disambiguation. In *Proceedings of the 21st international conference companion on World Wide Web (WWW'12)*. Lyon.
25. Xu Jian, **Zhengzhong Liu**, Qin Lu, Yu-lan Liu, & Chenchen Wang. 2011. PolyUCOMP in TAC 2011 Entity Linking and Slot-Filling. In *Proceedings of the Fourth Text Analysis Conference (TAC'11)*.
26. **Zhengzhong Liu**, Lu, Qin, & Xu, Jian. 2011. High Performance Clustering for Web Person Name Disambiguation Using Topic Capturing. In *The first International Workshop on Entity Orientation Search, SIGIR'11*.