

## Education

- **Johns Hopkins University, Baltimore, MD, USA** Aug. 2014 – present  
**Ph.D. in Computer Science**  
Center for Language and Speech Processing  
Advisor: Prof. Benjamin Van Durme
- **Peking University, Beijing, China** Sep. 2010 – Jul. 2014  
**B.Sc. in Computer Science**  
Department of Computer Science  
Advisor: Prof. Junfeng Hu  
Thesis: *Large-Scale Unsupervised Word Segmentation for Classical Chinese: Research & System.*

## Research Interests

I design new methods to tackle various problems arisen from knowledge discovery in NLP: question answering, information extraction, entity search etc., with an emphasis on scalability to large corpora, while drawing inspirations from related fields such as information retrieval, machine learning, and computational semantics.

## Publications

- [5] **Tongfei Chen**, Benjamin Van Durme (2017): Discriminative information retrieval for question answering sentence selection. In *Proceedings of the European Chapter of the Association for Computational Linguistics: Volume 2 (EACL)*, pp. 719-725.
- [4] Junhao Zhang, **Tongfei Chen**, Junfeng Hu (2015): On the relationship between Gaussian stochastic blockmodels and label propagation algorithms. *Journal of Statistical Mechanics: Theory and Experiment (J. Stat. Mech.)*. 2015(3), P03009.
- [3] Ni Sun, **Tongfei Chen**, Liumingjing Xiao, Junfeng Hu (2014): Diachronic deviation features in continuous space word representations. In *Proceedings of the 13th China National Conference on Computational Linguistics (CCL; LNAI 8801)*, pp. 23-33.
- [2] **Tongfei Chen**, Xiaojun Zou, Weimeng Zhu, Junfeng Hu (2013): Human-computer interactive Chinese word segmentation: An adaptive Dirichlet process mixture model approach. In *Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP)*, pp. 1278-1284.
- [1] **Tongfei Chen**, Weimeng Zhu, Xueqiang Lv, Junfeng Hu (2013): A Kalman filter based human-computer interactive segmentation system for ancient Chinese texts. In *Proceedings of the 12th China National Conference on Computational Linguistics (CCL; LNAI 8202)*, pp. 25-35.

## Research & Work Experience

- **Research Intern** May. 2017 – Aug. 2017  
IBM Thomas J. Watson Research Center, Yorktown Heights, NY  
Conducted research on question answering and model calibration under the supervision of Dr. Jiri Navratil and Dr. Bing Xiang.
- **Research Assistant** Aug. 2014 – present

*Center for Language and Speech Processing, Johns Hopkins University*

Conducted research on information extraction and question answering under the supervision of Prof. Benjamin Van Durme.

- **Research Assistant** *Jun. 2012 – Jul. 2014*  
*Institute of Computational Linguistics, Peking University*  
Conducted research on Chinese word segmentation, named-entity recognition and social network analysis under the supervision of Prof. Junfeng Hu.

## Selected Projects

- **SCALE 2016: Computer-Aided Discovery, Extraction and Translation** *Jun. 2016 – Aug. 2016*  
Participated in the 2016 Summer Camp for Applied Language Exploration (SCALE) workshop at the Human Language Technology Center of Excellence (HLTCOE) at Johns Hopkins University. Completed a system for user-customizable trainable cross-lingual information retrieval.
- **SCALE 2015: Chinese Entity Discovery and Linking** *Jun. 2015 – Jul. 2015*  
Participated in the 2015 Summer Camp for Applied Language Exploration (SCALE) workshop at the Human Language Technology Center of Excellence (HLTCOE) at Johns Hopkins University. Worked on entity linking and coreference resolution on Chinese data.
- **Reviewer Assignment System for Funding Applications** *Sep. 2013 – Oct. 2013*  
Collaborated in the development of an intelligent reviewer assignment system for the National Science Foundation of China (NSFC). Utilized techniques such as recommendation systems, graph-based key-phrase extraction and unsupervised ontology construction.
- **Human-Computer Interactive Segmenter for Classical Chinese** *Nov. 2012 – May. 2013*  
Conducted research on various methods, including Kalman filters and Dirichlet processes, for Chinese word segmentation. This project has resulted in two published papers.

## Teaching Experience

- **Teaching Assistant** *Sept. 2016 – Dec. 2016*  
*EN.600.365: Knowledge Discovery from Text, Johns Hopkins University*  
Graded assignments; gave guest lectures; and aided students in their course projects.
- **Guest Lecturer** *Nov. 2013*  
*048-30530: Introduction to Computing (Honor Track), Peking University*  
Gave lectures on introductory Bayesian statistics and machine learning.

## Honors and Awards

- Outstanding Undergraduate Thesis, Peking University. *Jun. 2014*
- Founder Inc. Scholarship, Peking University. Top 4 in 73. *Sep. 2013*

## Academic Service

- Program committee member:
  - KG4IR 2017: The First Workshop on Knowledge Graphs and Semantics for Text Retrieval and Analysis.
- Peer reviewing:
  - Secondary reviewer for ACL 2014, EMNLP 2014, WWW 2015, NAACL 2015, ACL 2015, TACL 2015, EACL 2017, TACL 2017, ACL 2017, ACL 2017 Demo Track, EMNLP 2017.

## Skills

- Programming languages:
  - Scala, Java, Python, C/C++, C#, Haskell
- Natural languages:
  - Mandarin Chinese (*native*), English
- Libraries and tools:
  - Deep learning: TensorFlow, PyTorch
  - Information retrieval: Lucene
  - Data serialization: Thrift
  - Distributed computing: Spark
  - Scala ecosystem: Cats, Shapeless