

CURRICULUM VITAE – LIDONG BING

Research Scientist, Alibaba DAMO Academy

CONTACT INFORMATION

Address: AXA Tower, 8 Shenton Way, Singapore 068811,
Email: l.bing@alibaba-inc.com, binglidong@gmail.com
Homepage: <https://lidongbing.github.io/>

POSITIONS HELD

Research Scientist	2018 - present
Alibaba DAMO Academy	
Senior Researcher	2016 - 2018
Tencent AI Lab	
Postdoctoral Fellow	2015 - 2016
Machine Learning Department, Carnegie Mellon University	
Postdoctoral Fellow	2013 - 2015
SE&EM Department, The Chinese University of Hong Kong	

EDUCATION

Ph.D.	2012
Faculty of Engineering, The Chinese University of Hong Kong.	
M.Sc.	2009
School of EECS, Peking University.	
B.Sc.	2006
School of Computer Science, Northeast Normal University.	

RESEARCH INTEREST

General areas: Natural Language Processing, Text Mining, Artificial Intelligence, Machine Learning.

Specific areas: Text Generation, Sentiment Analysis, Low-resource NLP, Information Extraction and Knowledge Base.

CURRENT RESEARCH

Low-resource NLP and MT

Neural methods are effective when there is enough labeled data to train the model. Unfortunately, this is not the case for low-resource languages like Indonesian, Thai, and Vietnamese. These languages often lack enough labeled data and accurate processing tools. Our goal in this project is to study the techniques to effectively process low-resource languages by transferring knowledge from resource-rich languages like English. In particular, we focus on cross-lingual methods for basic language processing tasks (e.g., word segmentation, NER tagging) and applications such as machine translation, and our goal is to solve these problems with no or few labeled examples in the target language.

Aspect-based Sentiment Analysis

Sentiment analysis is important for many applications. In this project, we develop a few Deep Neural Networks frameworks to tackle aspect-based sentiment analysis and aspect term extraction. We also explored some more integrated tasks. For example, we proposed an end-to-end solution for the joint task of aspect term extraction and its sentiment prediction. We also proposed a new task called aspect

sentiment triplet extraction, where one framework needs to extract the aspect term, its sentiment, as well as the cause (i.e. opinion words) of such sentiment. We have generated some promising results, and some papers were accepted at top venues.

Text Generation and Style Editing

Text Generation is a longstanding research topic, and it attracts special attention from researchers and industry recently because the development of seq2seq techniques and the growing need. I am working on a few text generation projects, including description text generation for structured data, style transfer, question generation, etc.

PREVIOUS
RESEARCH

Medical Data Analytics

In this project, we aimed at structuralizing the biomedical data. One type of medical data are from hospitals, namely patient treatment records. To understand the situation of the patient, it is critical to extract specific fields of the treatment records, such as symptom, previous disease history, etc. Another kind of important biomedical corpora are from comprehensive medical websites, such as PubMed and DailyMed. We extract rich medicine and disease knowledge from them and construct a knowledge base with distantly supervised and semi-supervised extraction methods. (Project link: <http://curtis.ml.cmu.edu/gnat/biomed/>)

Abstractive and compressive news summarization

We proposed new approaches to conduct abstractive and compressive document summarization. The sentences are decomposed into basic semantic units by constituency analysis. An optimization model based on integer linear programming was proposed to generate abstractive and compressive summaries with the units. This is a joint project with Noah's Ark Lab of Huawei Technologies Co. Ltd.

Web page structure analysis and data record detection

Two frameworks were proposed to analyze Web page structure with structured output SVM. They can explore the visual features and DOM structure features, and they have been utilized in Web data record extraction and Web page segmentation tasks. Based on the extracted data records, another framework was proposed for extracting entities and their attributes from the extracted data records. A distantly and semi-supervised model with Conditional Random Fields was proposed to learn extractors under the guidance of a proximate record graph.

Probabilistic graphical model based query refinement

A directed probabilistic graphical model was proposed to score candidate queries in the query refinement task. The candidate queries are generated by mining the history query log as well as the online social tagging data.

Ontology enhancement and its application in text mining

A framework was proposed to generate enhanced ontology by merging the entities of online encyclopedias (e.g. Wikipedia) into an existing ontology (e.g. WordNet). An adaptive method was also designed to utilize the enhanced ontology in text mining tasks.

PUBLICATIONS

[2020]

Knowing What, How and Why: A Near Complete Solution for Aspect-based Sentiment Analysis . (Haiyun Peng, Lu Xu, Lidong Bing, Wei Lu, Fei Huang, Luo Si), *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

Open Domain Event Text Generation. (Zihao Fu, Lidong Bing, Wai Lam), *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

Cross-Lingual Low-Resource Set-to-Description Retrieval for Global E-Commerce . (Juntao Li, Chang Liu, Lidong Bing, Xiaozhong Liu, Hong-song Li, Jian Wang, Dongyan Zhao, Rui Yan), *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

GRET: Global Representation Enhanced Transformer. (Rongxiang Weng, Shujian Huang, Hao-Ran Wei, Heng Yu, Weihua Luo, Lidong Bing, Jiajun Chen), *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

[2019]

Improving Question Generation With to the Point Context. (Jingjing Li, Yifan Gao, Lidong Bing, Irwin King, Michael R. Lyu), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

Using Customer Service Dialogues for Satisfaction Analysis with Context-Assisted Multiple Instance Learning. (Kaisong Song, Lidong Bing, Wei Gao, Jun Lin, Lujun Zhao, Jiancheng Wang, Changlong Sun, Xiaozhong Liu, Qiong Zhang), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

Hierarchical Pointer Net Parsing. (Linlin Liu, Xiang Lin, Shafiq Joty, Simeng Han, Lidong Bing), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

Semi-supervised Text Style Transfer: Cross Projection in Latent Space. (Mingyue Shang, Piji Li, Zhenxin Fu, Lidong Bing, Dongyan Zhao, Shuming Shi, Rui Yan), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

Tackling Long-Tailed Relations and Uncommon Entities in Knowledge Graph Completion. (Zihao Wang, Kwunping Lai, Piji Li, Lidong Bing, Wai Lam), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

Transferable End-to-End Aspect-based Sentiment Analysis with Selective Adversarial Learning. (Zheng Li, Xin Li, Ying Wei, Lidong Bing, Yu Zhang, Qiang Yang), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

- A Knowledge Regularized Hierarchical Approach for Emotion Cause Analysis.** (Chuang Fan, Hongyu Yan, Jiachen Du, Lin Gui, Lidong Bing, Min Yang, Ruifeng Xu, Ruibin Mao), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.
- Who Is Speaking to Whom? Learning to Identify Utterance Addressee in Multi-Party Conversations.** (Ran Le, Wenpeng Hu, Mingyue Shang, Zhenjun You, Lidong Bing, Dongyan Zhao, Rui Yan), *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.
- Exploiting BERT for End-to-End Aspect-based Sentiment Analysis .** (Xin Li, Lidong Bing, Wenxuan Zhang, Wai Lam), *EMNLP Workshop W-NUT*, 2019.
- Difficulty Controllable Question Generation for Reading Comprehension.** (Yifan Gao, Lidong Bing, Wang Chen, Jianan Wang, Irwin King, Michael R. Lyu), *The 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 2019.
- An Integrated Approach for Keyphrase Generation via Exploring the Power of Retrieval and Extraction.** (Wang Chen, Hou Pong Chan, Piji Li, Lidong Bing, Irwin King), *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2019.
- Persona-Aware Tips Generation.** (Piji Li, Zihao Wang, Lidong Bing, Wai Lam), *The Web Conference (WWW)*, 2019.
- Generating Distractors for Reading Comprehension Questions from Real Examinations.** (Yifan Gao, Lidong Bing, Piji Li, Irwin King, Michael R. Lyu), *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019.
- A Unified Model for Opinion Target Extraction and Target Sentiment Prediction.** (Xin Li, Lidong Bing, Piji Li, Wai Lam), *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019.
- Learning to Write Creative Stories with Thematic Consistency.** (Juntao Li, Lidong Bing, Lisong Qiu, Min Chen, Dongyan Zhao, Rui Yan), *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019.
- Abstractive Text Summarization by Incorporating Reader Comments.** (Shen Gao, Xiuying Chen, Piji Li, Zhaochun Ren, Lidong Bing, Dongyan Zhao, Rui Yan), *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, 2019.
- [2018]**
- Transformation Networks for Target-Oriented Sentiment Classification.** (Xin Li, Lidong Bing, Wai Lam, and Bei Shi.), *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2018.
- Learning Domain-Sensitive and Sentiment-Aware Word Embeddings.** (Bei Shi, Zihao Fu, Lidong Bing, and Wai Lam), *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2018.

- QuaSE: Sequence Editing (Accurate Text Style Transfer) under Quantifiable Guidance.** (Yi Liao, Lidong Bing, Piji Li, Shuming Shi, Wai Lam, and Tong Zhang), *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018.
- Variational Autoregressive Decoder for Neural Response Generation.** (Jiachen Du, Wenjie Li, Yulan He, Ruifeng Xu, Lidong Bing and Xuan Wang), *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018.
- Hybrid Neural Attention for Agreement/Disagreement Inference in Online Debates.** (Di Chen, Jiachen Du, Lidong Bing and Ruifeng Xu), *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018.
- Estimating Marginal Probabilities of n-grams for Recurrent Neural Language Models.** (Thanapon Noraset, Doug Downey and Lidong Bing), *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018.
- Aspect Term Extraction with History Attention and Selective Transformation.** (Xin Li, Lidong Bing, Piji Li, Wai Lam, and Zhimou Yang), *The 27th International Joint Conference on Artificial Intelligence (IJCAI)*, 2018.
- Semi-Supervised Learning with Declaratively Specified Entropy Constraints.** (Haitian Sun, Lidong Bing, and William W. Cohen), *Conference on Neural Information Processing Systems (NeurIPS)*, 2018.
- Joint Modeling of Participant Influence and Latent Topics for Recommendation in Event-based Social Networks.** (Yi Liao, Wai Lam, Lidong Bing, Xin Shen), *ACM Trans. Inf. Syst. (TOIS)*, 2018.
- Learning a Unified Embedding Space of Web Search from Large-scale Query Log.** (Lidong Bing, Zheng-Yu Niu, Piji Li, Wai lam, Haifeng Wang), *Knowl.-Based Syst. (KBS)*, 2018.
- [2017]**
- Bootstrapping Distantly Supervised IE using Joint Learning and Small Well-structured Corpora.** (Lidong Bing, Bhuwan Dhingra, Kathryn Mazaitis, Jong Hyuk Park, and William W. Cohen), *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, 2017.
- Salience Estimation via Variational Auto-Encoders for Multi-Document Summarization.** (Piji Li, Zihao Wang, Wai Lam, Zhaochun Ren, and Lidong Bing), *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, 2017.
- Towards a Language-independent Solution: Knowledge Base Completion by Searching the Web and Deriving Language Pattern.** (Lidong Bing, Zhiming Zhang, Wai Lam, William W. Cohen), *Knowl.-Based Syst. (KBS)*, volume 115, 80-86, 2017.
- Using Graphs of Classifiers to Impose Declarative Constraints on Semi-supervised Learning.** (Lidong Bing, Bhuwan Dhingra, and William W. Cohen), *The 26th International Joint Conference on Artificial Intelligence (IJCAI)*, 2017.

- Neural Rating Regression with Abstractive Tips Generation for Recommendation.** (Piji Li, Zhaochun Ren, Lidong Bing, and Wai Lam), *The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2017.
- A Deep Recurrent Generative Decoder for Abstractive Text Summarization.** (Piji Li, Lidong Bing, and Wai Lam), *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017.
- Recurrent Attention Networks for Aspect Sentiment Analysis.** (Peng Chen, Zhongqian Sun, Lidong Bing*, and Wei Yang), *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017. (*: corresponding author)
- Abstractive Document Summarization via Hierarchical Memory Networks.** (Piji Li, Wai Lam, Lidong Bing, and Zhengdong Lv), *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017.
- Reader-Aware Multi-Document Summarization: An Enhanced Model and The First Dataset.** (Piji Li, Lidong Bing, and Wai Lam), *Proceedings of the EMNLP 2017 Workshop on New Frontiers in Summarization (EMNLP-NewSum)*, 2017.
- [2016]**
- Distant IE by Bootstrapping Using Lists and Document Structure** (Lidong Bing, Mingyang Ling, Richard C. Wang, and William W. Cohen), *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 2899–2905, 2016.
- Unsupervised Extraction of Popular Product Attributes from E-commerce Web Sites by Considering Customer Reviews** (Lidong Bing, Tak-Lam Wong, and Wai Lam), *ACM Transactions on Internet Technology (TOIT)*, Volume 16 Issue 2, 12:1–12:17, 2016.
- Using Graphs of Classifiers to Impose Constraints on Semi-supervised Relation Extraction** (Lidong Bing, William W. Cohen, Bhuwan Dhingra, and Richard C. Wang), *Proceedings of the 5th Workshop on Automated Knowledge Base Construction (AKBC)*, 2016.
- Detecting Common Discussion Topics Across Culture From News Reader Comments** (Bei Shi, Wai Lam, Lidong Bing, and Yinqing Xu), *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 676–685, 2016.
- CUIS at the NTCIR-12 MobileClick2 Task** (Kwun Ping Lai, Wai Lam, and Lidong Bing), *Proceedings of the 12th NTCIR Conference*, 2016.
- Digesting Multilingual Reader Comments via Latent Discussion Topics with Commonality and Specificity** (Bei Shi, Wai Lam, Lidong Bing, and Yinqing Xu), *Proceedings of The 25th ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 2293–2298, 2016.
- Efficient and Scalable Detection of Overlapping Communities in Big Networks** (Tianshu Lyu, Lidong Bing, Zhao Zhang, and Yan Zhang), *Proceedings of The IEEE International Conference on Data Mining (ICDM)*, pp. 1071–1076, 2016.

Learning a Semantic Space of Web Search via Session Data (Lidong Bing, Zheng-Yu Niu, Wai Lam, and Haifeng Wang), *Proceedings of the 20th Asia Information Retrieval Societies Conference (AIRS)*, pp. 83–97, 2016.

[2015]

Multilingual Viewpoint Detection from news comments (Bei Shi, Wai Lam, Lidong Bing, and Yinqing Xu), *Proceedings of 2015 International Conference on Asian Language Processing (IALP)*, pp. 189–192, 2015.

Improving Distant Supervision for Information Extraction Using Label Propagation Through Lists (Lidong Bing, Sneha Chaudhari, Richard C. Wang, and William W. Cohen), *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 524–529, 2015.

A Unified Posterior Regularized Topic Model with Maximum Margin for Learning-to-Rank (Shoaib Jameel, Wai Lam, Steven Schockaert, and Lidong Bing), *Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 103–112, 2015.

Abstractive Multi-Document Summarization via Phrase Selection and Merging (Lidong Bing, Piji Li, Yi Liao, Wai Lam, Weiwei Guo, Rebecca Passonneau), *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 1587–1597, 2015.

Reader-Aware Multi-Document Summarization via Sparse Coding (Piji Li, Lidong Bing, Wai Lam, Hang Li, Yi Liao), *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1270–1276, 2015.

Supervised Topic Models with Word Order Structure for Document Classification and Retrieval Learning (Shoaib Jameel, Wai Lam, Lidong Bing), *Information Retrieval Journal*, volume 18, 283–330, 2015.

Web Query Reformulation via Joint Modeling of Latent Topic Dependency and Term Context (Lidong Bing, Wai Lam, Tak-Lam Wong, Shoaib Jameel), *ACM Trans. Inf. Syst. (TOIS)*, volume 33, 6:1–6:38, 2015.

Adaptive Concept Resolution for document representation and its applications in text mining (Lidong Bing, Shan Jiang, Wai Lam, Yan Zhang, Shoaib Jameel), *Knowl.-Based Syst. (KBS)*, volume 74, 1–13, 2015.

Nonparametric Topic Modeling Using Chinese Restaurant Franchise with Buddy Customers (Shoaib Jameel, Wai Lam, Lidong Bing), *Proceedings of the 37th European Conference on Information Retrieval (ECIR)*, pp. 648–659, 2015.

[2014]

Web page segmentation with structured prediction and its application in web page classification (Lidong Bing, Rui Guo, Wai Lam, Zheng-Yu Niu, Haifeng Wang), *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 767–776, 2014.

Website Community Mining from Query Logs with Two-Phase Clustering (Lidong Bing, Wai Lam, Shoaib Jameel, Chunliang Lu), *Proceedings of the 15th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing)*, pp. 201–212, 2014.

[2013]

Robust Detection of Semi-structured Web Records Using a DOM Structure-knowledge-driven Model (Lidong Bing, Wai Lam, Tak-Lam Wong), *ACM Trans. Web (TWEB)*, volume 7, 21:1-21:32, 2013.

Web Entity Detection for Semi-structured Text Data Records with Unlabeled Data (Chunliang Lu, Lidong Bing, Wai Lam, Ki Chan, Yuan Gu), *International Journal of Computational Linguistics and Applications*, volume 4, pp. 135-150, 2013.

Wikipedia entity expansion and attribute extraction from the web using semi-supervised learning (Lidong Bing, Wai Lam, Tak-Lam Wong), *Proceedings of the sixth ACM International Conference on Web Search and Data Mining (WSDM)*, pp. 567-576, 2013.

Structured positional entity language model for enterprise entity retrieval (Chunliang Lu, Lidong Bing, Wai Lam), *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 129-138, 2013.

Towards an enhanced and adaptable ontology by distilling and assembling online encyclopedias (Shan Jiang, Lidong Bing, Yan Zhang), *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 1703-1708, 2013.

[PhD study]

Unsupervised Extraction of Popular Product Attributes from Web Sites (Lidong Bing, Tak-Lam Wong, Wai Lam), *Proceedings of the 8th Asia Information Retrieval Societies Conference on Information Retrieval Technology (AIRS)*, pp. 437-446, 2012.

Normalizing web product attributes and discovering domain ontology with minimal effort (Tak-Lam Wong, Lidong Bing, Wai Lam), *Proceedings of the Forth International Conference on Web Search and Web Data Mining (WSDM)*, pp. 805-814, 2011.

Ontology enhancement and concept granularity learning: keeping yourself current and adaptive (Shan Jiang, Lidong Bing, Bai Sun, Yan Zhang, Wai Lam), *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1244-1252, 2011.

Using query log and social tagging to refine queries based on latent topics (Lidong Bing, Wai Lam, Tak-Lam Wong), *Proceedings of the 20th ACM Conference on Information and Knowledge Management (CIKM)*, pp. 583-592, 2011.

Towards a unified solution: data record region detection and segmentation (Lidong Bing, Wai Lam, Yuan Gu), *Proceedings of the 20th ACM Conference on Information and Knowledge Management (CIKM)*, pp. 1265-1274, 2011.

Investigation of Web Query Refinement via Topic Analysis and Learning with Personalization (Lidong Bing, and Lam Wai), *Proceedings of the 2nd SIGIR Workshop on Query Representation and Understanding (SIGIR-QRU)*, 2011.

Learning ontology resolution for document representation and its applications in text mining (Lidong Bing, Bai Sun, Shan Jiang, Yan Zhang, Wai Lam), *Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM)*, pp. 1713-1716, 2010.

[Master study]

Weighting Links Using Lexical and Positional Analysis in Web Ranking (Yi Zhang, Yexin Wang, Lidong Bing, Yan Zhang), *Proceedings of the ninth International Conference on Web-Age Information Management (WAIM)*, pp. 9-16, 2008.

Primary content extraction with Mountain Model (Lidong Bing, Yexin Wang, Yan Zhang, Hui Wang), *Proceedings of 8th IEEE International Conference on Computer and Information Technology (CIT)*, pp. 479-484, 2008.

LET: Towards More Precise Clustering of Search Results (Yi Zhang, Lidong Bing, Yexin Wang, Yan Zhang), *Proceedings of the fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 385-389, 2007.

PRESENTATION AND
TALK

Knowledge Extraction from Semi-structured Web Records and Entity-centric Documents. *Noah's Ark Lab of Huawei Inc.* 2016.

Knowledge Extraction from Semi-structured Web Records and Entity-centric Documents. *Department of Computer Science, The University of Iowa.* 2016.

Distant IE by Bootstrapping Using Lists and Document Structure. *AAAI, Phoenix, Arizona USA.* 2016.

Knowledge Extraction from Semi-structured Web Records and Entity-centric Documents. *Department of Computer Science, University of Illinois at Chicago.* 2016.

Knowledge Collection from Unstructured Web Data. *NELL Group, Machine Learning Department, Carnegie Mellon University.* 2015.

Reader Interest-Aware News Summarization for Smart Phone Users. *Noah's Ark Lab of Huawei Inc.* 2014.

Web Page Segmentation with Structured Prediction and its Application in Web Page Classification. *SIGIR, Gold Coast, Australia.* 2014.

Information Discovery from Semi-structured Record Sets on the Web - From Unstructured Web to Knowledge. *Web Data Mining Group of Baidu Inc.* 2014.

Website Community Mining from Query Logs with Two-phase Clustering. *CICLing, Kathmandu, Nepal.* 2014.

Web Knowledge Collection and Web Query Analysis. *Shanghai Jiao Tong University.* 2014.

Abstractive Multi-Document Summarization via Phrase Selection and Merging. *Noah's Ark Lab of Huawei Inc.* 2014.

Information Discovery from Semi-structured Record Sets on the Web - From Unstructured Web to Knowledge. *Fudan University.* 2013.

Wikipedia Entity Expansion and Attribute Extraction from the Web Using Semi-supervised Learning. *WSDM, Rome, Italy.* 2013.

Using Query Log and Social Tagging to Refine Queries Based on Latent Topics. *CIKM, Glasgow, UK.* 2011.

Towards a Unified Solution: Data Record Region Detection and Segmentation. *CIKM, Glasgow, UK.* 2011.

PROFESSIONAL ACTIVITIES

Associate Editor and Reviewer of journals

Associate Editor of Neurocomputing, since 2019.

Transactions of the Association for Computational Linguistics, since 2019

ACM Transactions on Information Systems (TOIS), since 2014

IEEE Transactions on Knowledge and Data Engineering (TKDE), since 2014

ACM Transactions on the Web (TWEB), since 2014

ACM Transactions on Intelligent Systems and Technology (TIST), since 2014

Computational Linguistics, since 2014

Journal of the Association for Information Science and Technology (JASIST), since 2014

Expert Systems With Applications, since 2016

Information Processing and Management, since 2016

Artificial Intelligence Review, since 2015

Knowledge-Based Systems (KBS), since 2014

Neural Networks, since 2014

Transactions on Asian and Low-Resource Language Information Processing (TALLIP), since 2013

Neural Computing and Applications Journal (NCA), 2012

Regular AC, SPC and PC of conferences

The Annual Meeting of the Association for Computational Linguistics (ACL)

The Conference on Empirical Methods in Natural Language Processing (EMNLP)

The AAAI Conference on Artificial Intelligence (AAAI)

International World Wide Web Conference (WWW)

The International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)

The International Joint Conference on Artificial Intelligence (IJCAI)

The ACM International Conference on Information and Knowledge Management (CIKM)

The International Joint Conference on Natural Language Processing (IJCNLP)

The Asia Information Retrieval Societies Conference (AIRS)

JOINT PHD STUDENTS

Qingyu Tan (National University of Singapore, 2020 -)

David Low (Nanyang Technological University, 2019 -)

Ran Zhou (Nanyang Technological University, 2019 -)

Linlin Liu (Nanyang Technological University, 2019 -)

Liyang Cheng (Singapore University of Technology and Design, 2019 -)

Lu Xu (Singapore University of Technology and Design, 2019 -)

OTHER STUDENT
MENTORING

Zhanming Jie (Singapore University of Technology and Design, 2019 -)
Yan Zhang (Singapore University of Technology and Design, 2019 -)
Qian Yu (The Chinese University of Hong Kong, 2018 -)
Yifan Gao (The Chinese University of Hong Kong, 2018 -)
Jingjing Li (The Chinese University of Hong Kong, 2018 -)
Zihao Fu (The Chinese University of Hong Kong, 2018 -)
Juntao Li (Peking University, 2018 -)
Mingyue Shang (Peking University, 2018 - 2018)
Xin Li (The Chinese University of Hong Kong, 2017 -)
Yi Liao (The Chinese University of Hong Kong, 2013 - 2014, 2017 - 2018)
Mingyang Ling (Carnegie Mellon University, 2015)
Sneha Chaudhari (Carnegie Mellon University, 2015)
Piji Li (The Chinese University of Hong Kong, 2014 - 2015)
Bei Shi (The Chinese University of Hong Kong, 2014 - 2015)
Shan Jiang (Peking University, 2011 - 2013)

HONOR AND
FELLOWSHIP

Staff Travel Grant, The Chinese University of Hong Kong, 2013 and 2014
Postdoctoral Fellowship of Faculty of Engineering, The Chinese University of Hong Kong, 2013
Postgraduate Student Travel Grant, The Chinese University of Hong Kong, 2011
Postgraduate Scholarship, The Chinese University of Hong Kong, 2009 - 2012
Outstanding Scholarship of CNPC, Peking University, 2008
Second Grade Scholarship, Northeast Normal University, 2006
Second Grade National Scholarship, 2005
First Grade Scholarship, Northeast Normal University, 2003, 2004 and 2005
Outstanding Graduate, Northeast Normal University, 2006
Outstanding University Student, Changchun city, 2005

LANGUAGE

Mandarin (Native), English (Fluent)