

CHENGYU WANG

Alibaba Group, Hangzhou, 311121, China

chywang2013@gmail.com, chengyu.wcy@alibaba-inc.com ◊ <https://chywang.github.io/>

WORK EXPERIENCE

Algorithm Expert

July 2020 – Current

Machine Learning Platform of AI (PAI), **Alibaba Cloud**

Major Responsibility I: AI Platforms

- Major developer of two open-source NLP toolkits named EasyTransfer (in Tensorflow) and EasyNLP (in PyTorch), supporting over 20 BUs within Alibaba Group and various clients/partners of Alibaba Cloud
 - Pre-trained language models enhanced by relational knowledge graphs and linguistic knowledge
 - Prompt-based parameter-efficient learning and in-context learning for large language models
 - Cross-modality content understanding, such as CLIP-style models for vision and language modeling and ARoBERT for speech and language modeling
 - AIGC models, such as diffusion-based models for text-to-image generation
 - A variety of transfer learning/meta-learning algorithms for NLP tasks across different domains and tasks
- Major developer of a speech recognition toolkit named EasyASR (in Tensorflow)
- Integration of AI toolkits to PAI products including PAI-DSW for Notebook-based model development, PAI-DLC for cloud-native training, PAI-EAS for online serving, and PAI-Designer for zero-code model training

Major Responsibility II: AI Applications

Major developer of several popular AI solutions built upon PAI products, including but not limited to the following:

- Fine-grained text tagging for e-commerce UGC analysis and understanding
- Parameter-efficient fine-tuning and inference pipelines for AIGC applications such as text-to-image generation and ChatGPT-style large language models
- Cross-modality representation learning and feature extraction for search and recommendations

EDUCATION

Ph.D. in Software Engineering

Sept. 2015 – June 2020

School of Software Engineering, **East China Normal University** (ECNU)

Supervisors: Prof. **Xiaofeng He** and Prof. **Aoying Zhou** Overall GPA: 90.29/100, Major GPA: 90.44/100

Ph.D. Thesis: Algorithmic Studies on Relation Extraction from Chinese Short Texts

Won **National Scholarship for PhD Students** three times (2016, 2017, 2019), **ECNU President Scholarship** (2019) and **Shanghai Outstanding Graduate** (2020)

B.Eng. in Software Engineering

Sept. 2011 – June 2015

School of Software Engineering, **East China Normal University** (ECNU)

Supervisor: Prof. **Xiaofeng He** Overall GPA: 3.77/4.0, Major GPA: 3.84/4.0

Won **National Scholarship** three times (2012, 2013, 2014) and **Shanghai Outstanding Graduate** (2015)

PUBLICATIONS

(* denotes equal contribution; # denotes corresponding author.)

[2023]

1. Ziyun Xu*, **Chengyu Wang***, Minghui Qiu, Fuli Luo, Runxin Xu, Songfang Huang, Jun Huang. *Making Pre-trained Language Models End-to-end Few-shot Learners with Contrastive Prompt Tuning*. Proceedings of the 16th ACM Conference on Web Search and Data Mining (**WSDM** 2023), pages 438–446

2. Jianing Wang, **Chengyu Wang**, Jun Huang, Ming Gao, Aoying Zhou. *Uncertainty-aware Self-training for Low-resource Neural Sequence Labeling*. Proceedings of the 37th AAAI Conference on Artificial Intelligence (**AAAI** 2023) (accepted)
3. Zhongjie Duan, **Chengyu Wang**, Cen Chen, Wenmeng Zhou, Jun Huang, Weining Qian. *Match4Match: Enhancing Text-Video Retrieval by Maximum Flow with Minimum Cost*. Proceedings of the Web Conference 2022 (**WWW** 2022), pages 3257–3267
4. Xiaoqing Chen*, **Chengyu Wang***, Junwei Dong, Minghui Qiu, Liang Feng, Jun Huang. *Boosting Prompt-based Few-shot Learners through Out-of-domain Knowledge Distillation*. Proceedings of the 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP** 2022) (accepted)
5. Boyu Hou*, **Chengyu Wang***, Xiaoqing Chen, Minghui Qiu, Liang Feng, Jun Huang. *Prompt-Distiller: Few-shot Knowledge Distillation for Prompt-based Language Learners with Dual Contrastive Learning*. Proceedings of the 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP** 2022) (accepted)
6. Moming Tang, **Chengyu Wang**, Jianing Wang, Chuanqi Tan, Songfang Huang, Cen Chen, Weining Qian. *XtremeCLIP: Extremely Parameter-efficient Tuning for Low-resource Vision Language Understanding*. Findings of the Association for Computational Linguistics: ACL 2023 (**ACL** (Findings) 2023) (accepted)
7. Lei Li, **Chengyu Wang**, Minghui Qiu, Cen Chen, Ming Gao, Aoying Zhou. *Accelerating BERT Inference with GPU-Efficient Exit Prediction*. Frontiers of Computer Science (**FCS**) (accepted)
8. Jiapeng Wang, **Chengyu Wang**, Xiaodan Wang, Jun Huang, Lianwen Jin. *ConaCLIP: Exploring Distillation of Fully-Connected Knowledge Interaction Graph for Lightweight Text-Image Retrieval*. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (**ACL** 2023) (Industry Track) (accepted)
9. Xiaodan Wang, **Chengyu Wang**, Lei Li, Zhixu Li, Ben Chen, Linbo Jin, Jun Huang, Yanghua Xiao, Ming Gao. *FashionKLIP: Enhancing E-Commerce Image-Text Retrieval with Fashion Multi-Modal Conceptual Knowledge Graph*. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (**ACL** 2023) (Industry Track) (accepted)
10. Moming Tang, **Chengyu Wang**, Jianing Wang, Cen Chen, Ming Gao, Weining Qian. *ParaSum: A Contrastive Paraphrasing Framework for Low-resource Extractive Summarization*. Proceedings of the 16th International Conference on Knowledge Science, Engineering and Management (**KSEM** 2023) (accepted)
11. Lei Li, Tingting Liu, **Chengyu Wang**, Minghui Qiu, Cen Chen, Ming Gao, Aoying Zhou. *Resizing Codebook of Vector Quantization Without Retraining*. Multimedia Systems (**MS**) (accepted)
12. Bingyan Liu, Weifeng Lin, Zhongjie Duan, **Chengyu Wang**, Ziheng Wu, Zipeng Zhang, Kui Jia, Lianwen Jin, Cen Chen, Jun Huang. *Rapid Diffusion: Building Domain-Specific Text-to-Image Synthesizers with Fast Inference Speed*. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (**ACL** 2023) (Industry Track) (accepted)
13. Zhen-Ru Zhang, Chuanqi Tan, Haiyang Xu, **Chengyu Wang**, Jun Huang, Songfang Huang. *Towards Adaptive Prefix Tuning for Parameter-Efficient Language Model Fine-tuning*. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (**ACL** 2023) (accepted)
14. Xiaodan Wang, Lei Li, Zhixu Li, Xuwu Wang, Xiangru Zhu, **Chengyu Wang**, Jun Huang, Yanghua Xiao. *AGREE: Aligning Cross-Modal Entities for Image-Text Retrieval Upon Vision-Language Pre-trained Models*. Proceedings of the 16th ACM Conference on Web Search and Data Mining (**WSDM** 2023), pages 456–464
15. Huimin Chen, Fengran Mo, Yanhao Wang, Cen Chen, Jian-Yun Nie, **Chengyu Wang**, Jamie Cui. *A Customized Text Sanitization Mechanism with Differential Privacy*. Findings of the Association for Computational Linguistics: ACL 2023 (**ACL** (Findings) 2023) (accepted)

[2022]

1. **Chengyu Wang**, Suyang Dai, Yipeng Wang, Fei Yang, Minghui Qiu, Kehan Chen, Wei Zhou, Jun Huang. *ARoBERT: An ASR Robust Pre-trained Language Model for Spoken Language Understanding*. IEEE/ACM

2. **Chengyu Wang**, Minghui Qiu, Jun Huang. *Building Natural Language Processing Applications with EasyNLP*. Proceedings of the 31st ACM International Conference on Information and Knowledge Management (**CIKM** 2022), pages 5100–5101 (Industry Day Speaker)
3. **Chengyu Wang**, Minghui Qiu, Taolin Zhang, Tingting Liu, Lei Li, Jianing Wang, Ming Wang, Jun Huang, Wei Lin. *EasyNLP: A Comprehensive and Easy-to-use Toolkit for Natural Language Processing*. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP** 2022), pages 22–29 (Demo Paper)
4. Jianing Wang*, **Chengyu Wang***, Fuli Luo, Chuanqi Tan, Minghui Qiu, Fei Yang, Qiuhui Shi, Songfang Huang, Ming Gao. *Towards Unified Prompt Tuning for Few-shot Text Classification*. Findings of the Association for Computational Linguistics: EMNLP 2022 (**EMNLP** (Findings) 2022), pages 524–536
5. Tingting Liu*, **Chengyu Wang***, Xiangru Zhu, Lei Li, Minghui Qiu, Ming Gao, Yanghua Xiao, Jun Huang. *ARTIST: A Transformer-based Chinese Text-to-Image Synthesizer Digesting Linguistic and World Knowledge*. Findings of the Association for Computational Linguistics: EMNLP 2022 (**EMNLP** (Findings) 2022), pages 881–888
6. Jianing Wang*, **Chengyu Wang***, Minghui Qiu, Qiuhui Shi, Hongbin Wang, Jun Huang, Ming Gao. *KECP: Knowledge-Enhanced Contrastive Prompting for Few-shot Extractive Question Answering*. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP** 2022), pages 3152–3163
7. Jianing Wang, **Chengyu Wang**, Chuanqi Tan, Minghui Qiu, Songfang Huang, Jun Huang, Ming Gao. *Span-Proto: A Two-stage Span-based Prototypical Network For Few-shot Named Entity Recognition*. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP** 2022), pages 3466–3476
8. Tingting Liu, **Chengyu Wang**, Cen Chen, Ming Gao, Aoying Zhou. *Understanding Long Programming Languages with Semantics-Aware Sparse Attention*. Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR** 2022), pages 2093–2098
9. Yiwu Yao*, **Chengyu Wang***, Jun Huang. *Quantized Winograd Acceleration for Conv1D Equipped ASR Models on Mobile Devices*. Proceedings of the 2022 IEEE International Conference on Acoustics, Speech, and Signal Processing (**ICASSP** 2022), pages 4804–4807
10. Taolin Zhang*, **Chengyu Wang***, Nan Hu, Minghui Qiu, Chengguang Tang, Xiaofeng He, Jun Huang. *DKPLM: Decomposable Knowledge-enhanced Pre-trained Language Model for Natural Language Understanding*. Proceedings of the 36th AAAI Conference on Artificial Intelligence (**AAAI** 2022), pages 11703–11711
11. Yan Fan, **Chengyu Wang**, Peng He, Yunhua Hu. *Building Multi-turn Query Interpreters for E-commercial Chatbots with Sparse-to-dense Attentive Modeling*. Proceedings of the 15th ACM Conference on Web Search and Data Mining (**WSDM** 2022), pages 1577–1580 (Demo Paper)
12. Runxin Xu, Fuli Luo, **Chengyu Wang**, Baobao Chang, Jun Huang, Songfang Huang, Fei Huang. *From Dense to Sparse: Contrastive Pruning for Better Pre-trained Language Model Compression*. Proceedings of the 36th AAAI Conference on Artificial Intelligence (**AAAI** 2022), pages 11547–11555
13. Lei Li, Dongxiang Zhang, **Chengyu Wang**, Cheqing Jin, Ming Gao, Aoying Zhou. *Enhancing Seq2seq Math Word Problem Solver with Entity Information and Math Knowledge*. Proceedings of the 23rd International Conference on Web Information Systems Engineering (**WISE** 2022), pages 370–385
14. Taolin Zhang, Junwei Dong, Jianing Wang, **Chengyu Wang#**, Ang Wang, Yinghui Liu, Jun Huang, Yong Li, Xiaofeng He. *Revisiting and Advancing Chinese Natural Language Understanding with Accelerated Heterogeneous Knowledge Pre-training*. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP** 2022) (Industry Track), pages 570–580
15. Dongyang Li, Taolin Zhang, Nan Hu, **Chengyu Wang**, Xiaofeng He. *HiCLRE: A Hierarchical Contrastive Learning Framework for Distantly Supervised Relation Extraction*. Findings of the Association for Computational Linguistics: ACL 2022 (**ACL** (Findings) 2022), pages 2567–2578

[2021]

1. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *Open Relation Extraction for Chinese Noun Phrases*. IEEE Transactions on Knowledge and Data Engineering (**TKDE**) 33(6): 2693–2708 (2021)
2. **Chengyu Wang**, Haojie Pan, Minghui Qiu, Jun Huang, Fei Yang, Yin Zhang. *Meta Distant Transfer Learning for Pre-trained Language Models*. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2021**), pages 9742–9752
3. **Chengyu Wang***, Jianing Wang*, Minghui Qiu, Jun Huang, Ming Gao. *TransPrompt: Towards an Automatic Transferable Prompting Framework for Few-shot Text Classification*. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2021**), pages 2792–2802
4. **Chengyu Wang***, Haojie Pan*, Yuan Liu, Kehan Chen, Minghui Qiu, Wei Zhou, Jun Huang, Haiqing Chen, Wei Lin, Deng Cai. *MeLL: Large-scale Extensible User Intent Classification for Dialogue Systems with Meta Lifelong Learning*. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD 2021**), pages 3649–3659
5. **Chengyu Wang**, Minghui Qiu, Jun Huang, Xiaofeng He. *KEML: A Knowledge-Enriched Meta-Learning Framework for Lexical Relation Classification*. Proceedings of the 35th AAAI Conference on Artificial Intelligence (**AAAI 2021**), pages 13924–13932
6. **Chengyu Wang**, Mengli Cheng, Xu Hu, Jun Huang. *EasyASR: A Distributed Machine Learning Platform for End-to-end Automatic Speech Recognition*. Proceedings of the 35th AAAI Conference on Artificial Intelligence (**AAAI 2021**), pages 16111–16113 (Demo Paper)
7. Mengli Cheng*, **Chengyu Wang***, Jun Huang, Xiaobo Wang. *Weakly Supervised Construction of ASR Systems from Massive Video Data*. Proceedings of the 22nd Annual Conference of the International Speech Communication Association (**Interspeech 2021**), pages 4533–4537
8. Haojie Pan*, **Chengyu Wang***, Minghui Qiu, Yichang Zhang, Yaliang Li, Jun Huang. *Meta-KD: A Meta Knowledge Distillation Framework for Language Model Compression across Domains*. Proceedings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (**ACL-IJCNLP 2021**), pages 3026–3036
9. Taolin Zhang*, **Chengyu Wang***, Minghui Qiu, Bite Yang, Zerui Cai, Xiaofeng He, Jun Huang. *Knowledge-Empowered Representation Learning for Chinese Medical Reading Comprehension: Task, Model and Resources*. Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021 (**ACL-IJCNLP (Findings) 2021**), pages 2237–2249
10. Ziyun Xu, **Chengyu Wang**, Peng Li, Yang Li, Ming Wang, Boyu Hou, Minghui Qiu, Chengguang Tang, Jun Huang. *When Few-shot Learning Meets Large-scale Knowledge-enhanced Pre-training: Alibaba at Few-CLUE*. Proceedings of the 10th CCF International Conference on Natural Language Processing and Chinese Computing (**NLPCC 2021**), pages 422–433, (First Place in NLPCC 2021 Shared Task 2: FewCLUE)
11. Cen Chen, **Chengyu Wang**, Minghui Qiu, Dehong Gao, Linbo Jin, Li Wang. *Cross-domain Knowledge Distillation for Retrieval-based Question Answering Systems*. Proceedings of the Web Conference 2021 (**WWW 2021**), pages 2613–2623
12. Zhi Li, Yuchen Zhai, **Chengyu Wang**, Minghui Qiu, Kailiang Li, Yin Zhang. *UnCLE: Leveraging Semantic Similarity to Compress Word Embeddings*. Findings of the Association for Computational Linguistics: EMNLP 2021 (**EMNLP (Findings) 2021**), pages 1815–1828
13. Minghui Qiu, Peng Li*, **Chengyu Wang***, Haojie Pan*, Ang Wang, Xianyan Jia, Le Jiang, Yaliang Li, Jun Huang, Jun Yang, Lin Wang, Deng Cai, Wei Lin. *EasyTransfer: A Simple and Scalable Deep Transfer Learning Platform for NLP Applications*. Proceedings of the 30th ACM International Conference on Information and Knowledge Management (**CIKM 2021**), pages 4075–4084
14. Taolin Zhang, Zerui Cai, **Chengyu Wang**, Peng Li, Yang Li, Minghui Qiu, Chengguang Tang, Xiaofeng He, Jun Huang. *HORNET: Enriching Pre-trained Language Representations with Heterogeneous Knowledge*

Sources. Proceedings of the 30th ACM International Conference on Information and Knowledge Management (**CIKM 2021**), pages 2608–2617

15. Haojie Pan, Cen Chen, **Chengyu Wang**, Minghui Qiu, Liu Yang, Feng Ji, Jun Huang, Haiqing Chen. *Learning to Expand: Reinforced Response Expansion for Information-seeking Conversations*. Proceedings of the 30th ACM International Conference on Information and Knowledge Management (**CIKM 2021**), pages 4055–4064
16. Taolin Zhang, Zerui Cai, **Chengyu Wang**, Minghui Qiu, Bite Yang, Xiaofeng He. *SMedBERT: A Knowledge-Enhanced Pre-trained Language Model with Structured Semantics for Medical Text Mining*. Proceedings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (**ACL-IJCNLP 2021**), pages 5882–5893
17. Zerui Cai, Taolin Zhang, **Chengyu Wang**, Xiaofeng He. *EMBERT: A Pre-trained Language Model for Chinese Medical Text Mining*. Proceedings of the 5th APWeb-WAIM Joint International Conference on Web and Big Data (**APWeb-WAIM 2021**), pages 242–257
18. Cen Chen, Xinjing Huang, Feng Ji, **Chengyu Wang**, Minghui Qiu, Jun Huang, Yin Zhang. *CAT-BERT: A Context-Aware Transferable BERT Model for Multi-Turn Machine Reading Comprehension*. Proceedings of the 26th International Conference on Database Systems for Advanced Applications (**DASFAA 2021**), pages 152–167

[2020]

1. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *HEEL: Exploratory Entity Linking for Heterogeneous Information Networks*. Knowledge and Information Systems (**KAIS**) 62(2): 485–506 (2020)
2. **Chengyu Wang**, Minghui Qiu, Jun Huang, Xiaofeng He. *Meta Fine-tuning Neural Language Models for Multi-domain Text Mining*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2020**), pages 3094–3104
3. **Chengyu Wang**, Xiaofeng He. *BiRRE: Learning Bidirectional Residual Relation Embeddings for Supervised Hypernymy Detection*. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (**ACL 2020**), pages 3630–3640
4. **Chengyu Wang**, Xiaofeng He, Xuegong Gong, Aoying Zhou. *Word Embedding Projection Models for Hypernymy Relation Prediction*. Chinese Journal of Computers (**CJC**) 43(5):868–883 (2020) (in Chinese with English abstracts)
5. Hui Cai, **Chengyu Wang**, Xiaofeng He. *Debiasing Learning to Rank Models with Generative Adversarial Networks*. Proceedings of the 4th APWeb-WAIM Joint Conference on Web and Big Data (**APWeb-WAIM 2020**), pages 45–60

[2019]

1. **Chengyu Wang**, Yan Fan, Xiaofeng He, Aoying Zhou. *Decoding Chinese User Generated Categories for Fine-grained Knowledge Harvesting*. IEEE Transactions on Knowledge and Data Engineering (**TKDE**) 31(8): 1491–1505 (2019)
2. **Chengyu Wang**, Yan Fan, Xiaofeng He, Aoying Zhou. *Predicting Hypernym-hyponym Relations for Chinese Taxonomy Learning*. Knowledge and Information Systems (**KAIS**) 58(3): 585–610 (2019)
3. **Chengyu Wang**, Yan Fan, Xiaofeng He, Hongyuan Zha, Aoying Zhou. *Idiomaticity Prediction of Chinese Noun Compounds and Its Applications*. IEEE Access (**IEEE ACCESS**) 7: 142866–142878 (2019)
4. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *SphereRE: Distinguishing Lexical Relations with Hyperspherical Relation Embeddings*. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (**ACL 2019**), pages 1727–1737
5. **Chengyu Wang**, Yan Fan, Xiaofeng He, Aoying Zhou. *A Family of Fuzzy Orthogonal Projection Models for Monolingual and Cross-lingual Hypernymy Prediction*. Proceedings of the 2019 World Wide Web Conference (**WWW 2019**), pages 1965–1976

6. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *Improving Hypernymy Prediction via Taxonomy Enhanced Adversarial Learning*. Proceedings of the 33rd AAAI Conference on Artificial Intelligence (**AAAI** 2019), pages 7128–7135
7. Yan Fan, **Chengyu Wang**, Boxing Chen, Zhongkai Hu, Xiaofeng He. *SPMM: A Soft Piecewise Mapping Model for Bilingual Lexicon Induction*. Proceedings of the 19th SIAM International Conference on Data Mining (**SDM** 2019), pages 244–252

[2018]

1. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *Event Phase Oriented News Summarization*. World Wide Web - Internet and Web Information Systems (**WWWJ**) 21(4): 1069–1092 (2018)
2. **Chengyu Wang**, Guomin Zhou, Xiaofeng He, Aoying Zhou. *NERank+: A Graph-based Approach for Entity Ranking in Document Collections*. Frontiers of Computer Science (**FCS**) 12(3): 504–517 (2018)
3. Jihong Yan, **Chengyu Wang**, Wenliang Cheng, Ming Gao, Aoying Zhou. *A Retrospective of Knowledge Graphs*. Frontiers of Computer Science (**FCS**) 12(1): 55–74 (2018)
4. Yan Fan, **Chengyu Wang**, Xiaofeng He. *Exploratory Neural Relation Classification for Domain Knowledge Acquisition*. Proceedings of the 27th International Conference on Computational Linguistics (**COLING** 2018), pages 265–2276
5. Guohai Xu, **Chengyu Wang**, Xiaofeng He. *Improving Clinical Named Entity Recognition with Global Neural Attention*. Proceedings of the 2nd APWeb-WAIM Joint Conference on Web and Big Data (**APWeb-WAIM** 2018), pages 264–279

[2017]

1. **Chengyu Wang**, Yan Fan, Xiaofeng He, Aoying Zhou. *Learning Fine-grained Relations from Chinese User Generated Categories*. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (**EMNLP** 2017), pages 2577–2587
2. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *A Short Survey on Taxonomy Learning from Text Corpora: Issues, Resources and Recent Advances*. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (**EMNLP** 2017), pages 1190–1203
3. Yan Fan, **Chengyu Wang**, Guomin Zhou, Xiaofeng He. *DKGBuilder: An Architecture for Building a Domain Knowledge Graph from Scratch*. Proceedings of the 22nd International Conference on Database Systems for Advanced Applications (**DASFAA** 2017), pages 663–667 (Demo Paper)
4. **Chengyu Wang**, Junchi Yan, Aoying Zhou, Xiaofeng He. *Transductive Non-linear Learning for Chinese Hypernym Prediction*. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (**ACL** 2017), pages 1394–1404
5. Lu Wang, **Chengyu Wang**, Keqiang Wang, Xiaofeng He. *BiUCB: A Contextual Bandit Algorithm for Cold-start and Diversified Recommendation*. Proceedings of the 8th IEEE International Conference on Big Knowledge (**ICBK** 2017), pages 248–253
6. Ruolan Yong, **Chengyu Wang**, Xiaofeng He. *A Transfer Learning based Boosting Model for Emotion Analysis*. Proceedings of the 8th IEEE International Conference on Big Knowledge (**ICBK** 2017), pages 264–269
7. Kun Wang, Guohai Xu, **Chengyu Wang**, Xiaofeng He. *A Hybrid Abnormal Advertising Traffic Detection Method*. Proceedings of the 8th IEEE International Conference on Big Knowledge (**ICBK** 2017), pages 236–241

[2016]

1. **Chengyu Wang**, Xiaofeng He. *Chinese Hypernym-hyponym Extraction from User Generated Categories*. Proceedings of the 26th International Conference on Computational Linguistics (**COLING** 2016), pages 1350–1361

2. **Chengyu Wang**, Rong Zhang, Xiaofeng He, Aoying Zhou. *Error Link Detection and Correction in Wikipedia*. Proceedings of the 25th ACM International Conference on Information and Knowledge Management (**CIKM** 2016), pages 307–316
3. **Chengyu Wang**, Rong Zhang, Xiaofeng He, Guomin Zhou, Aoying Zhou. *Event Phase Extraction and Summarization*. Proceedings of the 17th International Conference on Web Information Systems Engineering (**WISE** 2016), pages 473–488 (Best Paper Nominee)
4. **Chengyu Wang**, Rong Zhang, Xiaofeng He, Guomin Zhou, Aoying Zhou. *NERank: Bringing Order to Named Entities from Texts*. Proceedings of the 18th Asia Pacific Web Conference (**APWeb** 2016), pages 15–27
5. **Chengyu Wang**, Rong Zhang, Xiaofeng He, Aoying Zhou. *NERank: Ranking Named Entities in Document Collections*. Proceedings of the 25th World Wide Web Conference Companion (**WWW** 2016), pages 123–124 (Poster Paper)
6. Jihong Yan, Wenliang Cheng, **Chengyu Wang**, Jun Liu, Ming Gao, Aoying Zhou. *Optimizing Word Set Coverage for Multi-event Summarization*. Journal of Combinatorial Optimization (**JCO**) 30(4):996–1015 (2016)
7. Chi Zhang, Yanhua Wang, **Chengyu Wang**, Wenliang Cheng, Xiaofeng He. *SNE extractor: A Prototype for Extracting Semantic Networks from Web Documents*. Proceedings of the 17th International Conference on Web-Age Information Management (**WAIM** 2016), pages 527–530 (Demo Paper)
8. Jinyang Li, Yanhua Wang, Yan Fan, **Chengyu Wang**, Rong Zhang, Xiaofeng He. *Chinese Taxonomy Construction and Search System*. Journal of Computer Applications (**JCA**) 36(S1):207–209 (2016) (in Chinese with English abstracts)

[2015 and Before]

1. **Chengyu Wang**, Bing Xiao, Xiang Li, Jiawen Zhu, Xiaofeng He, Rong Zhang. *Predicting Users' Age Range in Micro-blog Network*. Proceedings of the 14th International Conference on Web Information Systems Engineering (**WISE** 2013), pages 516–521 (WISE 2013 Challenge Runner-up Award)
2. **Chengyu Wang**, Ming Gao, Xiaofeng He and Rong Zhang. *Challenges in Chinese Knowledge Graph Construction*. Proceedings of the 6th International Workshop on Data Engineering meets the Semantic Web (**DESWeb** 2015), pages 59–61, in conjunction with **ICDE** 2015
3. Jinyang Li, **Chengyu Wang**, Xiaofeng He, Rong Zhang, Ming Gao. *User Generated Content Oriented Chinese Taxonomy Construction*. Proceedings of the 17th Asia Pacific Web Conference (**APWeb** 2015), pages 623–634
4. Wenliang Cheng, **Chengyu Wang**, Bing Xiao, Weining Qian, Aoying Zhou. *On Statistical Characteristics of Real-life Knowledge Graphs*. Proceedings of the 6th International Workshop on Big data benchmarks, Performance, Optimization and Emerging hardware (**BPOE** 2015), pages 37–49, in conjunction with **VLDB** 2015
5. Leyi Song, **Chengyu Wang**, Xiaoyi Duan, Bing Xiao, Xiao Liu, Rong Zhang, Xiaofeng He, Xueqing Gong. *TaxiHailer: A Situation-specific Taxi Pick-up Points Recommendation System*. Proceedings of the 19th International Conference on Database Systems for Advanced Applications (**DASFAA** 2014), pages 523–526 (Demo Paper)
6. Fan Xia, Qunyan Zhang, **Chengyu Wang**, Weining Qian, Aoying Zhou. *On the Rise and Fall of Sina Weibo: Analysis Based on a Fixed User Group*. Proceedings of the 1st International Workshop on Scalable Social Event Processing and Management (**SSEPM** 2015), pages 224–231, in conjunction with **ICDE** 2015
7. Yonghao Su, Chi Zhang, Jinyang Li, **Chengyu Wang**, Weining Qian, Aoying Zhou. *Cross-lingual Entity Query from Large-scale Knowledge Graphs*. Proceedings of the 2nd International Workshop on Web Data Mining and Applications (**WDMA** 2015), pages 139–150, in conjunction with **APWeb** 2015

Honors and Grants:

- Nov. 2022 **Second Prize of CCF Science and Technology Achievement Award**
April 2022 Innovation Pioneer, Alibaba Group
Sept. 2020 Excellent Dissertation Award of the Faculty of Information, ECNU
May 2020 **Shanghai Outstanding Graduate**
Aug. 2019 ACL 2019 Student Volunteer Program
Jan. 2019 AAAI 2019 Student Volunteer and Scholarship Program
Oct. 2018 **Microsoft Research Asia Fellowship Nominee**
(Top 20 out of all Ph.D. candidate applicants in Asia-Pacific regions)
Dec. 2016 Recipient, ECNU Outstanding Doctoral Dissertation Cultivation Plan of Action
(Only 2 among all Ph.D. candidates in School of Computer Science and Software Engineering, ECNU)
Dec. 2016 COLING 2016 Student Support Program
Nov. 2016 WISE 2016 Best Paper Nominee
Oct. 2016 SIGIR Student Travel Grant (CIKM 2016)
April 2016 Google Student Travel Grant (WWW 2016)
May 2015 **Shanghai Outstanding Graduate**

Scholarships:

- Dec. 2019 **ECNU President Scholarship** (Top-10 graduate students in ECNU)
Dec. 2019 **National Scholarship for PhD Students**
Dec. 2018 Student Science and Technology Excellence Scholarship
(Top-1 in School of Computer Science and Software Engineering, ECNU)
Dec. 2017 **National Scholarship for PhD Students**
Dec. 2016 **National Scholarship for PhD Students**
Dec. 2015 ECNU Graduate Freshman Scholarship
Dec. 2014 **National Scholarship**
Dec. 2013 **National Scholarship**
Dec. 2012 **National Scholarship**

Competition Awards:

- July 2021 First Place in NLPCC 2021 Shared Task 2: FewCLUE
May 2018 Third Prize in National English Contest for College Students (Level A)
June 2014 Special Prize in Mathematical Modeling Contest for College Students
May 2014 **Special Prize in National English Contest for College Students (Level C)**
(Top 0.1% among all participants)
Dec. 2013 **Second Prize in First National Big Data Technology Innovation Contest**
(Among 3 teams that won the Second Prize out of 660 teams worldwide)
Oct. 2013 WISE 2013 Challenge Runner-up Award
May 2013 Second Prize in National English Contest for College Students (Level C)
May 2012 Third Prize in National English Contest for College Students (Level C)

ACADEMIC ACTIVITIES

Regular Conference Program Committee (PC) Membership:

- Annual Meeting of the Association for Computational Linguistics (ACL)
- Conference on Empirical Methods in Natural Language Processing (EMNLP)
- Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)
- International Conference on Computational Linguistics (COLING)
- Conference of the European Chapter of the Association for Computational Linguistics (EACL)
- AAAI Conference on Artificial Intelligence (AAAI)

- International Joint Conference on Artificial Intelligence (IJCAI)
- International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
- ACM Multimedia Conference (ACM MM)
- ACL Rolling Review (for *ACL conferences)

Invited Reviewer for Journals:

- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)
- ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE Transactions on Computational Social Systems (TCSS)
- IEEE Transactions on Artificial Intelligence (TAI)
- Expert Systems With Applications (ESWA)
- Natural Language Engineering (NLE)
- Knowledge-Based Systems (KBS)
- Scientific Reports
- Expert Systems
- Neurocomputing
- AI Open

SKILLS AND LANGUAGES

Programming	Python, SQL, \LaTeX
Platforms & Tools	TensorFlow, PyTorch, Keras, Hadoop, Linux
English	Shanghai Advanced English Interpretation (Dec. 2013) CET-6: 636/710 CET-4: 663/710
Japanese	New Japanese Language Proficiency Test (JLPT): N3 Level (July 2013)

(Last updated: June 2023)