CHENGYU WANG

chywang2013@gmail.com ⋄ https://chywang.github.io/

Room 105 (West), Mathematics Building, 3663 North Zhongshan Road,

East China Normal University, Shanghai, 200062, China

EDUCATION

Ph.D. in Software Engineering   *Sept. 2015 – June 2020 (Expected)*
(Data Science and Engineering)
School of Computer Science and Software Engineering, **East China Normal University** (ECNU)
Supervisors: Prof. **Xiaofeng He** and Prof. **Aoying Zhou**
Research interests: relation extraction, taxonomy learning, knowledge graphs, lexical semantics, natural language understanding, etc.
Overall GPA: 90.29/100, Major GPA: 90.44/100
Won **National Scholarship for PhD Students** twice (2016, 2017)

B.Eng. in Software Engineering   *Sept. 2011 – June 2015*
Software Engineering Institute, **East China Normal University** (ECNU)
Overall GPA: 3.77/4.0, Major GPA: 3.84/4.0
Thesis: Regional Traffic Pattern Mining Based on Latent Semantics (Supervisor: Prof. **Xiaofeng He**)
Won **National Scholarship** three times (2012, 2013, 2014) and the honor of **Shanghai Outstanding Graduate** (2015)

PH.D. THESIS RESEARCH PROJECTS

Learning hypernymy relations for taxonomy construction   *Sept. 2015 – Present*

- Design word embedding based iterative and transductive learning algorithms to predict Chinese hypernymy ("is-a") relations among term pairs. Construct a large-scale Chinese taxonomy from user generated categories. (**KAIS**, **COLING** 2016, **ACL** 2017)

- Propose adversarial and transfer learning frameworks for cross-domain and cross-lingual hypernymy prediction with limited training data. (**AAAI** 2019, **WWW** 2019)

- Survey 100+ papers on state-of-the-art techniques for hypernymy learning. (**EMNLP** 2017)

Knowledge extraction and understanding for Chinese short texts   *Jan. 2017 – Present*

- Propose an unsupervised graph-based approach to harvest relational textual patterns and extract relations from Chinese short texts. (**TKDE**, **EMNLP** 2017)

- Design a three-layer data-driven system to extract noun-based relations from Chinese noun phrases, under the framework of Open Information Extraction. (**TKDE** (submitted))

Entity linking and link analysis of large-scale semantic networks   *Mar. 2015 – June 2016*

- Present a two-stage ranking and classification method to detect linking errors in homogeneous information networks, with case studies of English and Chinese Wikipedia. (**CIKM** 2016)

- Propose an exploratory entity linking algorithm to performance entity linking and to detect new entities for heterogeneous information networks, with a case study of DBLP. (**KAIS** (submitted))

SELECTED NON-THESIS PROJECTS

Data-driven analysis of online educational public opinions   *Dec. 2014 – Dec. 2016*
Joint work with the National Institute of Educational Policy Research, ECNU

- Participate in developing an educational public opinion analysis platform based on massive Web-scale educational data. Hotspots in educational public opinions are detected and analyzed. Popular Web opinions towards these hotspots are summarized.

- Participate in writing and publishing two books in Chinese (i.e., the "Annual Report of Chinese Online Educational Public Opinion Analysis" series in 2014 and 2015) based on the results of educational public opinion analysis.

Taxi pick-up points recommendation based on large trajectory data *July 2013 – Dec. 2013*

- Participate in developing a distributed spatio-temporal data analysis toolkit to process large-scale traffic trajectory data on Hadoop clusters.

- Design a location-based taxi pick-up point recommender based on the prediction of the waiting time for vacant taxis. Regression models and the non-homogeneous Poisson process are employed.

- Win the Second Prize in the First National Big Data Technology Innovation Contest (A 30K RMB prize, top-4 among 660 teams).

PUBLICATIONS

Manuscripts in Submission:

1. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *Commonsense Relation Extraction from Chinese Noun Phrases.* submitted to IEEE Transactions on Knowledge and Data Engineering (**TKDE**)

2. **Chengyu Wang**, Xiaofeng He, Aoying Zhou. *HEEL: Exploratory Entity Linking for Heterogeneous Information Networks.* submitted to Knowledge and Information Systems (**KAIS**)

Refereed Journal Articles:

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**) (in press)

2. **Chengyu Wang**, Yan Fan, Xiaofeng He, Aoying Zhou. *Predicting Hypernym-Hyponym Relations for Chinese Taxonomy Learning.* Knowledge and Information Systems (**KAIS**) (in press)

3. **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)

4. **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)

5. 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)

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)

Refereed Conference Papers:

1. **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), accepted

2. **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), accepted

3. **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

4. **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

5. **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

6. **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

7. **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

8. **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)

9. **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

10. **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

11. **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)

12. 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), accepted

13. 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

14. 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

15. 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

16. 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

17. 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

18. 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

Refereed Workshop Papers:

1. **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

2. 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

3. 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

4. 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

Refereed Software Demonstrations:

1. 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

2. 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

3. Chi Zhang, Yanhua Wang, **Chengyu Wang**, Wenliang Cheng, Xiaofeng He. *SNExtractor: 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

Papers in Chinese with English Abstract:

1. 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)

SKILLS AND ABILITIES

Technical Skills:

Programming	Java, Python, C++, SQL, LaTeX
Platforms & Tools	Hadoop, Keras, TensorFlow, Mahout, Lucene, Weka
Qualifications	EMC Academic Associate, Cloud Infrastructure and Services
	EMC Academic Associate, Information Storage and Management

Language Proficiency:

Chinese	Mandarin (Native Speaker), Wu (Dialect)
English	Good Command of Reading, Writing, Translation and Interpretation
	Shanghai Advanced English Interpretation (Dec. 2013)
	CET-6: 636/710 CET-4: 663/710
Japanese	Intermediate Level with Basic Communication Skills
	New Japanese Language Proficiency Test (JLPT): N3 Level (July 2013)

SELECTED AWARDS AND SCHOLARSHIPS

Scholarships:

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**

Honors and Grants:

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**

Awards:

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
	(Ranked 2 out of 12 submissions)
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

Conference Presentation:

2019	AAAI 2019 (Honolulu, USA)
2016	WWW 2016 (Montréal, Canada), CIKM 2016 (Indianapolis, USA), COLING 2016 (Osaka, Japan), APWeb 2016 (Suzhou, China), WISE 2016 (Shanghai, China)
2015	DESWeb@ICDE 2015 (Seoul, South Korea)
2013	WISE 2013 (Nanjing, China)

Academic Services:

Conference PC Member	ACL 2019 (Word-level Semantics Track)
Conference Review (Secondary)	NDBC 2015, APWeb 2016, KDD 2017, CIKM 2017, ICDM 2017, COLING 2018, COLING 2018 Mentoring Program, KDD 2018, ICDM 2018, ICBK 2018
Journal Review	IEEE Access
Conference Volunteer	AAAI 2019 (Honolulu, USA)

Teaching Experience:

2016	TA, Database Systems and Applications (Spring Semester)
2015	TA, Web System Development (Autumn Semester)

REFERENCES

- Xiaofeng He, Professor
  E-mail: xfhe@sei.ecnu.edu.cn
  School of Computer Science and Software Engineering, East China Normal University

- Aoying Zhou, Professor
  E-mail: ayzhou@dase.ecnu.edu.cn
  Vice President, East China Normal University

(Last updated: February 2019)