Hongming Zhang

(+1) 856-883-0473 | hzhangal0330@gmail.com | panda0881.github.io/Hongming_Homepage | panda0881 | Hongming Zhang

Bio

Hongming Zhang is currently a senior researcher at Tencent AI Lab, Bellevue. He leads the research on **Self-evolving LLM Agents** at Tencent AI Lab. In 2024, his team published tens of top-tier conference papers on LLM agents and self-evolving LLMs. Before joining Tencent, Hongming received his Ph.D. degree from HKUST in 2021. He also worked as a research scholar at UPenn with Prof. Dan Roth from 2020 to 2021. By Feb. 2025, Hongming has published hundreds of top-tier conference papers with a total citation of **4,254**. Topics cover LLM agents, Complex Reasoning of LLM, LLM Self-evolving, Advanced RAG systems, Commonsense Reasoning, Information Retrieval, and Information Extraction.

Work Experience

Tencent AI Lab — *Bellevue, WA*

Research Lead — *Aug. 2023 - Present*

- Lead a team of ten to work on front-tier research on self-evolving LLM agents
- Developed and released the first **general-purpose proactive** LLM agent system **"Cognitive Kernel"** that could actively interact with the real world. It achieves SOTA performance on complex daily tasks that require **real-time** info, **private** info, and **long-term memory**. More details can be found at https://github.com/tencent/CogKernel.
- Developed Several leading LLM agent systems for different scenarios, including **Web Agents**, **MLE Agents**, **Agents with Long-term Memory**, and **Code Agents**. We publish these works at top-tier conferences such as ICLR, ACL, and EMNLP.

Tencent AI Lab — *Bellevue, WA*

Senior Researcher — *Nov. 2021 - Present*

- Conduct independent front-tier research. Topics cover **LLM complex reasoning**, **LLM agents**, **Advanced RAG Systems**, **Information Retrieval**, and **Information Extraction**. Publish tens of papers at top-tier conferences, such as ICLR, ICML, Neurips, ACL, EMNLP, and NAACL.

Education

HKUST (Hong Kong University of Science and Technology) — *Hong Kong, China*

Ph.D. in Computer Science — *Sep. 2018 - Nov. 2021*

- Supervisor: Prof. Yangqiu Song
- Research topics cover **Commonsense Reasoning**, **Selectional Preference**, and **Coreference Resolution**

University of Pennsylvania (Hong Kong University of Science and Technology) — *Philadeplphia, PA*

Visiting Scholar — *Jan. 2020 - Nov. 2021*

- Supervisor: Prof. Dan Roth
- Research topics cover **Commonsense Reasoning** and **Information Extraction**

HKUST (Hong Kong University of Science and Technology) — *Hong Kong, China*

M.Phil in TLE (Technology, Leadership, and Entrepreneurship) — *Sep. 2016 - Aug. 2018*

- Supervisor: Prof. Yangqiu Song
- Major in Computer Science. Researching **Graph Neural Networks** and **Information Retrieval**

HKUST (Hong Kong University of Science and Technology) — *Hong Kong, China*

B.E. in Electronic and Computer Engineering — *Sep. 2012 - Aug. 2016*

- Two Minor Degrees in Information Technology and Entrepreneurship
- Cumulative Grade Average: 3.91/4.30, Ranking in major: 3/74
- Academic Achievement Medal (Highest academic award for undergraduates at HKUST)

Selected Publications

- **Hongming Zhang**, Xiaoman Pan, Hongwei Wang, Kaixin Ma, Wenhao Yu, Dong Yu. Cognitive kernel: An open-source agent system towards generalist autopilots.(Technical Report, 2024)
  - Create the first open-source proactive agent system for generalist autopilot tasks.
- **Hongming Zhang**, Xin Liu, Haojie Pan, Haowen Ke, Tianqing Fang, Jiefu Ou, Yangqiu Song. ASER: Towards Large-scale Commonsense Knowledge Acquisition via Higher-order Selectional Preference over Eventualities. (Artificial Intelligence 2022)
  - Propose the largest eventuality knowledge graph for commonsense reasoning.
- **Hongming Zhang**, Liwei Qiu, Lingling Yi, Yangqiu Song. Scalable Multiplex Network Embedding. (IJCAI 2018)
  - Propose multiplex network embedding for better representation learning and recommendation.

Honors & Awards

2024	**Outstanding Paper**, Transactions of Machine Learning Research	*Bellevue, WA, U.S.A*
2023	**Outstanding Paper**, EMNLP	*Singapore*
2022	**Engineering PhD Research Excellence Finalist Award**, HKUST	*Hong Kong, China*
2020	**SENG Academic Award for Continuing PhD students**, HKUST	*Hong Kong, China*
2019	**Microsoft Research Asia Fellowship**, Microsoft Research Asia	*Beijing, China*
2019	**Tencent Rhino-Bird Scholarship**, Tencent	*Shenzhen, China*
2018	**Hong Kong PhD Fellowship**, Hong Kong	*Hong Kong, China*
2018	**Excellent Research Award**, HKUST	*Hong Kong, China*
2016	**Academic Achievement Medal**, HKUST	*Hong Kong, China*
2014	**Academic Excellence Award**, HKUST	*Hong Kong, China*

Program Committees

- **Journal Editorial Board**: Semantic Web Journal special edition on Commonsense Knowledge and Reasoning; Frontiers in Big Data.
- **Journal Reviewer**: Applied Network Science; Computer Science and Language; IEEE Transactions on Neural Networks and Learning Systems; NeuralComputing.
- **Area Chair and Conference Committee**: ICLR'24-25; ICML'25; Neurips'23-24; COLM'24-25; AACL'23; ACL'21-24; EMNLP'21-24; CVPR 22-24; NAACL'21-24; IJCAI'21-23; AAAI'21-23; COLING'20; AACL'20; CIKM'19; AKBC'21.
- **Others**: IJCAI 2019 WebMaster.

All Peer-reviewed Publications

- 2025
  - Liqiang Jing, Zhehui Huang, Xiaoyang Wang, Wenlin Yao, Wenhao Yu, Kaixin Ma, **Hongming Zhang**, Xinya Du, Dong Yu. DSBench: How Far Are Data Science Agents from Becoming Data Science Experts? (ICLR 2025)
  - Siru Ouyang, Wenhao Yu, Kaixin Ma, Zilin Xiao, Zhihan Zhang, Mengzhao Jia, Jiawei Han, **Hongming Zhang**, Dong Yu. RepoGraph: Enhancing AI Software Engineering with Repository-level Code Graph. (ICLR 2025)
- 2024
  - With many other Collaborators. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. (TMLR 2024, Outstanding Paper)
  - Wenhao Yu, **Hongming Zhang**, Xiaoman Pan, peixin cao, Kaixin Ma, Jian Li, Hongwei Wang, Dong Yu. Chain-of-Note: Enhancing Robustness in Retrieval-Augmented Language Models. (EMNLP 2024)
  - Fengyu Cai, Xinran Zhao, Tong Chen, Sihao Chen, **Hongming Zhang**, Iryna Gurevych, Heinz Koeppl. MixGR: Enhancing Retriever Generalization for Scientific Domain through Complementary Granularity. (EMNLP 2024)
  - Tong Chen, Hongwei Wang, Sihao Chen, Wenhao Yu, Kaixin Ma, Xinran Zhao, **Hongming Zhang**, Dong Yu. Dense X Retrieval: What Retrieval Granularity Should We Use? (EMNLP 2024)
  - Ruixin Hong, **Hongming Zhang**, Xiaoman Pan, Dong Yu, Changshui Zhang. Abstraction-of-Thought Makes Language Models Better Reasoners. (Findings of EMNLP 2024)
  - Chunkit Chan, Cheng Jiayang, Yauwai Yim, Zheye Deng, Wei Fan, Haoran Li, Xin Liu, **Hongming Zhang**, Weiqi Wang, Yangqiu Song. NegotiationToM: A Benchmark for Stress-testing Machine Theory of Mind on Negotiation Surrounding. (Findings of EMNLP 2024)
  - Zhaowei Wang, Wei Fan, Qing Zong, **Hongming Zhang**, Sehyun Choi, Tianqing Fang, Xin Liu, Yangqiu Song, Ginny Wong, Simon See. AbsInstruct: Eliciting Abstraction Ability from LLMs through Explanation Tuning with Plausibility Estimation. (ACL 2024)
  - Yinya Huang, Ruixin Hong, **Hongming Zhang**, Wei Shao, Zhicheng YANG, Dong Yu, Changshui Zhang, Xiaodan Liang, Linqi Song. CLOMO: Counterfactual Logical Modification with Large Language Models. (ACL 2024)
  - Hongliang He, Wenlin Yao, Kaixin Ma, Wenhao Yu, Yong Dai, **Hongming Zhang**, Zhenzhong Lan, Dong Yu. WebVoyager: Building an End-to-End Web Agent with Large Multimodal Models. (ACL 2024)
  - Xinran Zhao, **Hongming Zhang**, Xiaoman Pan, Wenlin Yao, Dong Yu, Tongshuang Wu, Jianshu Chen. Fact-and-Reflection (FaR) Improves Confidence Calibration of Large Language Models. (Findings of ACL 2024)
  - Fengyu Cai, Xinran Zhao, **Hongming Zhang**, Iryna Gurevych, Heinz Koeppl. GEOHard: Towards Measuring Class-wise Hardness through Modelling Class Semantics. (Findings of ACL 2024)
  - Ruixin Hong, **Hongming Zhang**, Xinyu Pang, Dong Yu, Changshui Zhang. A Closer Look at the Self-Verification Abilities of Large Language Models in Logical Reasoning. (NAACL 2024)

- Sihao Chen, **Hongming Zhang**, Tong Chen, Ben Zhou, Wenhao Yu, Dian Yu, Baolin Peng, Hongwei Wang, Dan Roth, Dong Yu. Sub-Sentence Encoder: Contrastive Learning of Propositional Semantic Representations. (NAACL 2024)
- Hangfeng He, **Hongming Zhang**, Dan Roth. SocREval: Large Language Models with the Socratic Method for Reference-free Reasoning Evaluation. (Findings of NAACL 2024)
- Tianqing Fang, Zhaowei Wang, Wenxuan Zhou, **Hongming Zhang**, Yangqiu Song, Muhao Chen. Getting Sick After Seeing a Doctor? Diagnosing and Mitigating Knowledge Conflicts in Event Temporal Reasoning. (Findings of NAACL 2024)
- Zhaowei Wang, Haochen Shi, Weiqi Wang, Tianqing Fang, **Hongming Zhang**, Sehyun Choi, Xin Liu, Yangqiu Song. AbsPyramid: Benchmarking the Abstraction Ability of Language Models with a Unified Entailment Graph. (Findings of NAACL 2024)
- Nan Xu, **Hongming Zhang**, and Jianshu Chen. CEO: Corpus-based Open-Domain Event Ontology Induction. (Findings of EACL 2024)
- Tianqing Fang, Wenxuan Zhou, Fangyu Liu, **Hongming Zhang**, Yangqiu Song, Muhao Chen. On-the-fly Denoising for Data Augmentation in Natural Language Understanding. (Findings of EACL 2024)
- Haoyu Wang, **Hongming Zhang**, Kaiqiang Song, Dong Yu, Dan Roth. Event Semantic Classification in Context. (Findings of EACL 2024)

- 2023
  - Haoyu Wang, **Hongming Zhang**, Yueguan Wang, Yuqian Deng, Muhao Chen, Dan Roth. Are All Steps Equally Important? Benchmarking Essentiality Detection of Events. (EMNLP 2023)
  - Hongwei Wang, **Hongming Zhang**, and Dong Yu. On the Dimensionality of Sentence Embeddings. (Findings of EMNLP 2023)
  - Keming Lu, Xiaoman Pan, Kaiqiang Song, **Hongming Zhang**, Dong Yu, and Jianshu Chen. PIVOINE: Instruction Tuning for Open-world Information Extraction. (Findings of EMNLP 2023)
  - Jiayang Cheng, Lin Qiu, Tsz Ho CHAN, Tianqing Fang, Weiqi Wang, Chunkit Chan, Qipeng Guo, **Hongming Zhang**, Yangqiu Song, Yue Zhang, Zheng Zhang. STORYANALOGY: Deriving Story-level Analogies from Large Language Models to Unlock Analogical Understanding. (EMNLP 2023)
  - James Y. Huang, Wenlin Yao, Kaiqiang Song, **Hongming Zhang**, Muhao Chen, Dong Yu. Bridging Continuous and Discrete Spaces: Interpretable Sentence Representation Learning via Compositional Operations. (EMNLP 2023 Outstanding Paper)
  - Xinran Zhao, **Hongming Zhang**, Xiaoman Pan, Wenlin Yao, Dong Yu, and Jianshu Chen. Thrust: Adaptively Propels Large Language Models with External Knowledge. (Neurips 2023)
  - Xiang Li, Jiangwei Yu, Xinran Zhao, **Hongming Zhang**, and Yu-Xiong Wang. Video State-changing Object Segmentation. (ICCV 2023)
  - Ruixin Hong, **Hongming Zhang**, Hong Zhao, Dong Yu, and Changshui Zhang. Faithful Question Answering with Monte-Carlo Planning. (ACL 2023)
  - Zhaowei Wang, Quyet V. Do, **Hongming Zhang**, Jiayao Zhang, Weiqi Wang, Tianqing Fang, Yangqiu Song, Ginny Y. Wong, and Simon See. COLA: Contextualized Commonsense Causality Reasoning from the Causal Inference Perspective. (ACL 2023)
  - Xiaoman Pan, Wenlin Yao, **Hongming Zhang**, Dian Yu, Dong Yu, and Jianshu Chen. Knowledge-in-Context: Towards Knowledgeable Semi-Parametric Language Models. (ICLR 2023)
  - Haoyu Wang, **Hongming Zhang**, Yuqian Deng, Jacob R Gardner, Muhao Chen, and Dan Roth. Extracting or guessing? improving faithfulness of event temporal relation extraction. (EACL 2023)
  - Wenlin Yao, Lifeng Jin, **Hongming Zhang**, Xiaoman Pan, Kaiqiang Song, Dian Yu, Dong Yu, and Jianshu Chen. How do Words Contribute to Sentence Semantics? Revisiting Sentence Embeddings with a Perturbation Method. (Findings of EACL 2023)
  - **Hongming Zhang**, Yintong Huo, Yanai Elazar, Yangqiu Song, Yoav Goldberg, and Dan Roth. CIKQA: Learning Commonsense Inference with a Unified Knowledge-in-the-loop QA Paradigm. (Findings of EACL 2023)
  - Zizheng Lin, **Hongming Zhang**, and Yangqiu Song. Global Constraints with Prompting for Zero-Shot Event Argument Classification. (Findings of EACL 2023)
- 2022
  - **Hongming Zhang**, Xin Liu, Haojie Pan, Haowen Ke, Tianqing Fang, Jiefu Ou, Yangqiu Song. ASER: Towards Large-scale Commonsense Knowledge Acquisition via Higher-order Selectional Preference over Eventualities. (Artificial Intelligence 2022)
  - Xintong Yu, **Hongming Zhang**, Ruixing Hong, Yangqiu Song, and Changshui Zhang. CD-PCR: Improving Visual Dialog with Pronoun Coreference Resolution. (Pattern Recognition 2022)

- Ruixin Hong, **Hongming Zhang**, Xintong Yu, and Changshui Zhang. Learning Event Extraction From a Few Guideline Examples. (IEEE/ACM Transactions on Audio, Speech, and Language Processing 2022)
- Yinya Huang, **Hongming Zhang**, Ruixin Hong, Xiaodan Liang, Changshui Zhang, and Dong Yu MetaLogic: Logical Reasoning Explanations with Fine-Grained Structure. (EMNLP 2022)
- Zhaowei Wang, **Hongming Zhang**, Tianqing Fang, Yangqiu Song, Ginny Y. Wong, and Simon See. SubeventWriter: Iterative Sub-event Sequence Generation with Coherence Controller. (EMNLP 2022)
- Fei Wang, Kaiqiang Song, **Hongming Zhang**, Lifeng Jin, Sangwoo Cho, Wenlin Yao, Xiaoyang Wang, Muhao Chen, and Dong Yu. Salience Allocation as Guidance for Abstractive Summarization. (EMNLP)
- Yue Yang, Wenlin Yao, **Hongming Zhang**, Xiaoyang Wang, Dong Yu, and Jianshu Chen. Z-LaVI: Zero-Shot Language Solver Fueled by Visual Imagination. (EMNLP 2022)
- **Hongming Zhang**, Wenlin Yao, and Dong Yu. Efficient Zero-shot Event Extraction with Context-Definition Alignment. (EMNLP 2022)
- Tianqing Fang, Quyet V. Do, **Hongming Zhang**, Yangqiu Song, Ginny Y. Wong, and Simon See. PseudoReasoner: Leveraging Pseudo Labels for Commonsense Knowledge Base Population. (EMNLP 2022)
- Ying Su, Zihao Wang, Tianqing Fang, **Hongming Zhang**, Yangqiu Song, and Tong Zhang MICO: A Multi-alternative Contrastive Learning Framework for Commonsense Knowledge Representation. (EMNLP 2022)
- Ying Su, **Hongming Zhang**, Yangqiu Song, and Tong Zhang. Multilingual Word Sense Disambiguation with Unified Sense Representation. (COLING 2022)
- Jiayao Zhang, **Hongming Zhang**, Weijie Su, and Dan Roth. Causal Inference Principles for Reasoning about Commonsense Causality. (ICML 2022)
- Hantian Ding, Jinrui Yang, Yuqian Deng, **Hongming Zhang**, and Dan Roth. Towards Open-Domain Topic Classification. (NAACL 2022 Demo Track)
- Ruixin Hong, **Hongming Zhang**, Xintong Yu, and Changshui Zhang. METGEN: A Module-based Entailment Tree Generation Framework for Answer Explanation. (Findings of NAACL 2022)
- Jiaxin Bai, Zihao Wang,**Hongming Zhang**, and Yangqiu Song. Query2Particles: Knowledge Graph Reasoning with Particle Embeddings. (Findings of NAACL 2022)
- Xinran Zhao, **Hongming Zhang**, and Yangqiu Song. PCR4ALL: A Comprehensive Evaluation Benchmark for Pronoun Coreference Resolution in English. (LREC 2022)
- Changlong Yu, **Hongming Zhang**, Yangqiu Song, and Wilfred Ng. CoCoLM: COmplex COmmonsense Enhanced Language Model. (Findings of ACL 2022)
- Ying Su, **Hongming Zhang**, Yangqiu Song, and Tong Zhang. Rare and Zero-shot Word Sense Disambiguation using Z-Reweighting. (ACL 2022)
- Yintong Huo, Yuxin Su, **Hongming Zhang**, and Michael Lyu. ARCLIN: Automated API Mention Resolution for Unformatted Texts. (ICSE 2022)

• 2021
- Haoyu Wang, **Hongming Zhang**, Muhao Chen, and Dan Roth. Learning Constraints and Descriptive Segmentation for Subevent Detection. (EMNLP 2021)
- Yanai Elazar, **Hongming Zhang**, Yoav Goldberg, and Dan Roth. Back to Square One: Bias Detection, Training and Commonsense Disentanglement in the Winograd Schema. (EMNLP 2021)
- Xintong Yu, **Hongming Zhang**, Yangqiu Song, Changshui Zhang, Kun Xu, and Dong Yu . Exophoric Pronoun Resolution in Dialogues with Topic Regularization. (EMNLP 2021)
- Tianqing Fang, Weiqi Wang, Sehyun Choi, Shibo Hao, **Hongming Zhang**, Yangqiu Song, and Bin He. Benchmarking Commonsense Knowledge Base Population with an Effective Evaluation Dataset. (EMNLP 2021)
- Tianqing Fang, Haojie Pan, **Hongming Zhang**, Yangqiu Song, Kun Xu, Dong Yu. Do Boat and Ocean Suggest Beach? Dialogue Summarization with External Knowledge. (AKBC 2021)
- **Hongming Zhang**, Haoyu Wang, and Dan Roth. Zero-shot Label-Aware Event Trigger and Argument Classification. (Findings of ACL 2021)
- Qing Lyu, **Hongming Zhang**, Elior Sulem, and Dan Roth. Zero-shot Event Extraction via Transfer Learning: Challenges and Insights. (ACL 2021)
- Xinran Zhao, Esin Durmus, **Hongming Zhang**, and Claire Cardie. Leveraging Topic Relatedness for Argument Persuasion. (Findings of ACL 2021)
- **Hongming Zhang**, Yintong Huo, Xinran Zhao, Yangqiu Song, and Dan Roth. Learning Contextual Causality from Time-consecutive Images. (CVPR 2021)
- Tianqing Fang, **Hongming Zhang**, Weiqi Wang, Yangqiu Song, and Bin He. DISCOS: Bridging the Gap between Discourse Knowledge and Commonsense Knowledge. (WWW 2021)
- Jiaxin Bai, **Hongming Zhang**, Yangqiu Song, and Kun Xu. Joint Coreference Resolution and Character Link-

ing for Multiparty Conversation. (EACL 2021)

- 2020
  - Muhao Chen, **Hongming Zhang**, Haoyu Wang, and Dan Roth. What Are You Trying to Do? Semantic Typing of Event Processes. (CoNLL 2020)
  - **Hongming Zhang**, Muhao Chen, Haoyu Wang, Yangqiu Song, and Dan Roth. Analogous Process Structure Induction for Sub-event Sequence Prediction. (EMNLP 2020)
  - Haoyu Wang, Muhao Chen, **Hongming Zhang**, and Dan Roth. Joint Constrained Learning for Event-Event Relation Extraction. (EMNLP 2020)
  - Changlong Yu, Jialong Han, Peifeng Wang, Yangqiu Song, **Hongming Zhang**, Wilfred Ng, and Shuming Shi. When Hearst Is not Enough: Improving Hypernymy Detection from Corpus with Distributional Models. (EMNLP 2020)
  - Changlong Yu, **Hongming Zhang**, Yangqiu Song, Wilfred Ng, and Lifeng Shang. Enriching Large-Scale Eventuality Knowledge Graph with Entailment Relations. (AKBC 2020)
  - **Hongming Zhang**, Daniel Khashabi, Yangqiu Song, and Dan Roth. TransOMCS: From Linguistic Graphs to Commonsense Knowledge. (IJCAI 2020)
  - **Hongming Zhang**, Xinran Zhao, and Yangqiu Song. WinoWhy: A Deep Diagnosis of Essential Commonsense Knowledge for Answering Winograd Schema Challenge. (ACL 2020)
  - **Hongming Zhang**, Xin Liu, Haojie Pan, Yangqiu Song, and Cane Wing-Ki Leung. ASER: A Large-scale Eventuality Knowledge Graph. (WWW 2020)
- 2019
  - **Hongming Zhang**, Jiaxin Bai, Yan Song, Kun Xu, Changlong Yu, Yangqiu Song, Wilfred Ng, and Dong Yu. Multiplex Word Embeddings for Selectional Preference Acquisition. (EMNLP 2019)
  - Xintong Yu, **Hongming Zhang**, Yangqiu Song, Yan Song, and Changshui Zhang. What You See is What You Get: Visual Pronoun Coreference Resolution in Conversations. (EMNLP 2019)
  - Nedjma Ousidhoum, Zizheng Lin, **Hongming Zhang**, Yangqiu Song, and Dit-Yan Yeung. Multilingual and Multi-aspect Hate Speech Analysis. (EMNLP 2019)
  - **Hongming Zhang**, Hantian Ding, and Yangqiu Song. SP-10K: A Large-Scale Evaluation Set for Selectional Preference Acquisition. (ACL 2019)
  - **Hongming Zhang**, Yan Song, Yangqiu Song, and Dong Yu. Knowledge-aware Pronoun Coreference Resolution. (ACL 2019)
  - **Hongming Zhang**, Yan Song, and Yangqiu Song. Incorporating Context and External Knowledge for Pronoun Coreference Resolution. (NAACL 2019)
- 2018
  - **Hongming Zhang**, Liwei Qiu, Lingling Yi, and Yangqiu Song. Scalable Multiplex Network Embedding. (IJCAI 2018)