

# Yanzeng Li

Email: liyanzeng@stu.pku.edu.cn

Github: github.com/lsvih

Website: <https://cv.lsvih.com>

Mobile: +86-186-\*\*\*\*-\*\*\*\*

Address: 5<sup>#</sup> Yiheyuan Road, Haidian, Beijing

## RESEARCH INTERESTS

---

Natural Language Processing (NLP), Applications of NLP, Knowledge Graph (KG), and KG incorporated Large Language Models, ML Sys.

## EDUCATION

---

### Peking University

Beijing, China

- Wangxuan Institute of Computer Technology  
PhD candidate, in Computer Application Technology

September 2021 - Present

### Chinese Academy of Sciences

Beijing, China

- Institute of Information Engineering  
M.Eng., in Computer Science and Technology

September 2018 - June 2021

### Central South University

Hunan, China

- Powder Metallurgy Research Institute  
B.Eng., in Powder Materials Science and Engineering

September 2012 - June 2016

## RESEARCH EXPERIENCE

---

### Data Management Lab

WICT, PKU

- *Research Intern & Student, supervised by Lei Zou*  
  - **Knowledge Graph Construction.** Construct KG from unstructured corpus; Information extraction methods; KGC System.
  - **Multimodal Graph Data Management.** Manage multimodal graph data; Graph database optimize; Database query extension.
  - **Machine Learning System.** Multi-model scheduling; Deep learning model accelerate; Distributed system.
  - **Applications of KG.** Knowledge base question and answering; LLM-based methods for KBQA; Visualization and analysis of KG.

June 2021 - Present

### Beijing Institute for General Artificial Intelligence (BIGAI)

Beijing, China

- *Research Intern. Mentored by Zilong Zhen & Wenjuan Han*  
  - **Multimodal Knowledge Graph.** Construct multimodal knowledge graph, and build applications like KBQA, VQA, etc., upon the knowledge graph.
  - **Graph Database OPS.** Establish and maintain infrastructure (gStore graph database) for serving graph data, as a important component in business pipeline.

November 2021 - December 2023

### Knowledge Discovery for Security Group

IIE, CAS

- *Research Intern & Student, supervised by Tingwen Liu*  
  - **Information Extraction.** Extract elements from text, e.g., named entities, aspect.
  - **Pre-training Language Model.** Design and optimize PLM architecture; Accelerate heterogeneous PLMs.
  - **Data Annotation.** Build up data annotation tools; Active learning.

May 2017 - June 2021

### State Key Laboratory for Powder Metallurgy

PMRI, CSU

- *Research Intern, supervised by Huachan Fang*

2015 - April 2016

## SELECTED PUBLICATIONS

---

- Leveraging Large Language Model as Simulated Patients for Clinical Education  
**Yaneng Li**, Cheng Zeng, Jialun Zhong, Ruoyu Zhang, Minhao Zhang, Lei Zou  
*In Preprint*
- gBuilder: A Scalable Knowledge Graph Construction System for Unstructured Corpus  
**Yanzeng Li**, Lei Zou
- LLMaAA: Making Large Language Models as Active Annotators  
Ruoyu Zhang, **Yanzeng Li**, Yongliang Ma, Ming Zhou, Lei Zou.  
*In Findings of the Association for Computational Linguistics: EMNLP 2023*
- CORD: A Three-Stage Coarse-to-Fine Framework for Relation Detection in Knowledge Base Question Answering  
**Yanzeng Li**, Sen Hu, Wenjuan Han, Lei Zou.  
*In Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM), 2023*
- AtTGen: Attribute Tree Generation for Real-World Attribute Joint Extraction  
**Yanzeng Li**, Bingcong Xue, Ruoyu Zhang, Lei Zou.  
*In Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics (ACL), 2023*
- A Novel Table-to-Graph Generation Approach for Document-Level Joint Entity and Relation Extraction  
Ruoyu Zhang, **Yanzeng Li**, Lei Zou.  
*In Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics (ACL), 2023*
- VGStore: A Multimodal Extension to SPARQL for Querying RDF Scene Graph  
**Yanzeng Li**, Zilong Zheng, Wenjuan Han, Lei Zou  
*In Proceedings of the 21st International Semantic Web Conference (ISWC), 2022*
- Introducing Semantic Information for Numerical Attribute Prediction over Knowledge Graphs  
Bingcong Xue, **Yanzeng Li**, Lei Zou  
*In Proceedings of the 21st International Semantic Web Conference (ISWC), 2022*
- Enhancing Chinese Pre-trained Language Model via Heterogeneous Linguistics Graph  
**Yanzeng Li**, Jiangxia Cao, Xin Cong, Zhenyu Zhang, Bowen Yu, Hongsong Zhu, Tingwen Liu.  
*In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL), 2022*
- Relation Extraction based on Data Partition and Representation Integration [*Best Paper Runner up*]  
Jiapeng Zhao, Panpan Zhang, Tingwen Liu, Zhenyu Zhang, **Yanzeng Li**, Jinqiao Shi  
*In Proceedings of 2021 IEEE Sixth International Conference on Data Science in Cyberspace (DSC), 2021*
- Semi-Open Attribute Extraction from Chinese Functional Description Text  
Li Zhang, **Yanzeng Li\***, Ruoyu Zhang, Wenjie Li.  
*In Proceedings of the 13th Asian Conference on Machine Learning (ACML), 2021.*
- FITAnnotator: A Flexible and Intelligent Text Annotation System  
**Yanzeng Li**, Bowen Yu, Li Quangan, Tingwen Liu.  
*In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT) (Demonstrations), 2021*
- Enhancing Pre-trained Chinese Character Representation with Word-aligned Attention  
**Yanzeng Li**, Bowen Yu, Mengge Xue, Tingwen Liu.  
*In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 2020*
- Character-based BiLSTM-CRF Incorporating POS and Dictionaries for Chinese Opinion Target Extraction  
**Yanzeng Li**, Tingwen Liu, Diying Li, Quangan Li, Jinqiao Shi, Yanqiu Wang.  
PMLR 95:518-533, *In Proceedings of the 10th Asian Conference on Machine Learning (ACML), 2018.*

## PROFESSIONAL SERVICE

---

- Journal Reviewer, IEEE Transactions on Knowledge and Data Engineering (TKDE)
- Journal Reviewer, IEEE Open Journal of Signal Processing (OJSP)
- Program Committee Member, the AAAI Conference on Artificial Intelligence (AAAI), 2021/2022/2023/2024
- Program Committee Member, the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2021/2022/2023, & EMNLP Industry Track 2023
- Program Committee Member, the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2023
- Program Committee Member, the Annual Meeting of the Association for Computational Linguistics (ACL), 2020/2021/2022
- Program Committee Member, the Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT), 2021/2022/2024
- Program Committee Member, the ACM International Conference on Multimedia (ACM MM), 2024
- Program Committee Member, the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023/2024
- Program Committee Member, the 1st Conference on Language Modeling (COLM), 2024
- Program Committee Member, The ACM International Conference on Information and Knowledge Management (CIKM), 2023/2024 & Demo Track 2023
- Review Assistant/Sub Reviewer for IJCAI 2018/2021, etc.
- ...

## AWARDS

---

- **The 3rd Prize of PKU Scholarship**  
Peking University 2023
- **Merit Student**  
Peking University 2022, 2023
- **Merit Student**  
Wangxuan Institute of Computer Technology, Peking University 2022, 2023
- **Ubiquant Scholarship**  
Peking University 2022
- **IIE Presidential Special Scholarship (Top 1%)**  
Institute of Information Engineering, Chinese Academy of Sciences 2020
- **Laboratory Excellent Student Scholarship.**  
Institute of Information Engineering, Chinese Academy of Sciences 2020
- **Merit Student**  
University of Chinese Academy of Sciences 2019, 2020
- **Outstanding Student League Cadres**  
Central South University 2015

## (PREVIOUS) INDUSTRY EXPERIENCE

---

- **(Boloni Corp) Jufenqi E-Business Co.,Ltd** Beijing  
*Web Developer, mostly developed the front-end* August 2016 - May 2017
  - Developed B2B&B2C products, ARP&CRM and FMIS, etc.
- **(Boloni Corp) KuaiKuaiZhuang Technology Co., Ltd** Beijing  
*Web Developer & App Developer* June 2016 - September 2016
  - Technology architecture design. Developed official website, iOS app, etc.
- **ShanHe Design Inc. [SHD 834776]** Beijing  
*Full Stack Developer* March 2016 - June 2016
  - Developed official website, online store, recommended system, etc.