

# Xiusi Chen

201 N Goodwin Ave, Urbana, IL 61801

☎ (+1) (310)-913-7838 • ✉ xiusic@illinois.edu

## RESEARCH INTERESTS

---

My ultimate research goal is to build automated language agents that can assist human in a variety of domains. To achieve this goal, I worked on:

- **Large Language Models**, especially low-resource LLM reasoning (*Gotta*, *MinPrompt*, *TPD*) and alignment (*IterAlign*);
- **Language and Other Modality Applications**, especially in retrieval and recommendation (*CP-LSH*, *BLaIR*);
- **Complex Dynamics Modeling and Decision Making**, especially the planning tasks in sports domain (*ReLiable*, *PlayBest*).

## EDUCATION

---

### University of Illinois Urbana-Champaign

Postdoctoral Research Fellow

**Advisor:** Prof. Heng Ji.

### University of California, Los Angeles

Ph.D. in Computer Science

**Thesis Committee:** Prof. Wei Wang (Chair), Prof. Yizhou Sun, Prof. Kai-Wei Chang, Prof. Jeffrey Brantingham

### Peking University

B.S & M.S. in Computer Science

**Thesis Committee:** Prof. Jun Gao (Chair), Prof. Bin Cui, Prof. Ming Zhang, and Prof. Shuai Ma

## INDUSTRIAL EXPERIENCES

---

### Amazon Search

Query Understanding team, Applied Scientist Intern

6/2023 - 12/2023

### Amazon Search

PECOS team, Applied Scientist Intern

6/2022 - 9/2022

### NEC Laboratories America

Data Science & System Security, Research Intern

6/2019 - 9/2019

### Alibaba

DAMO Academy, Research Intern

7/2016 - 12/2016

## SELECTED PUBLICATIONS

---

(\*: Equal contribution)

**Xiusi Chen**, Jyun-Yu Jiang, Wei-Cheng Chang, Cho-Jui Hsieh, Hsiang-Fu Yu, Wei Wang, *MinPrompt: Graph-based Minimal Prompt Data Augmentation for Few-shot Question Answering*. In Proceedings of The 62nd Annual Meeting of the Association for Computational Linguistics (**ACL**), main conference, 2024.

**Xiusi Chen**, Hongzhi Wen, Sreyashi Nag, Chen Luo, Qingyu Yin, Ruirui Li, Zheng Li, Wei Wang, *ITERALIGN: Iterative Constitutional Alignment of Large Language Models*. In Proceedings of The 2024 Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), main conference, 2024.

Yu Zhang\*, **Xiusi Chen**\*, Bowen Jin\*, Sheng Wang, Shuiwang Ji, Wei Wang, Jiawei Han, *A Comprehensive Survey of Scientific Large Language Models and Their Applications in Scientific Discovery*. In Proceedings of The 2024 Conference on Empirical Methods in Natural Language Processing (**EMNLP**), main conference, 2024. **Totaling 400+ stars as of August 2024.**

**Xiusi Chen**, Wei-Yao Wang, Ziniu Hu, David Reynoso, Kun Jin, Mingyan Liu, P. Jeffrey Brantingham, Wei Wang, *Professional Basketball Player Behavior Synthesis via Planning with Diffusion*. In Proceedings of The 33<sup>rd</sup> ACM International Conference on Information and Knowledge Management (**CIKM**), 2024.

**Xiusi Chen**, Yu Zhang, Jinliang Deng, Jyun-Yu Jiang, Wei Wang, *GOTTA: Generative Few-shot Question Answering by Prompt-based Cloze Data Augmentation*. In Proceedings of the 2023 SIAM International Conference on Data Mining (**SDM**), 2023. **Long paper, Best Poster Award (Honorable Mention)**

**Xiusi Chen**, Jyun-Yu Jiang, Kun Jin, Yichao Zhou, Mingyan Liu, Jefferey Brantingham, Wei Wang, *RELIABLE: Offline Reinforcement Learning for Tactical Strategies in Professional Basketball Games*. In Proceedings of The 31<sup>st</sup> ACM International Conference on Information and Knowledge Management (**CIKM**), 2022.

**Xiushi Chen**, Jyun-Yu Jiang, Wei Wang, *Scalable Graph Representation Learning via Locality-Sensitive Hashing*. In Proceedings of The 31<sup>st</sup> ACM International Conference on Information and Knowledge Management (**CIKM**), 2022.

Yijun Tian\*, Yikun Han\*, **Xiushi Chen**\*, Wei Wang, Nitech V. Chawla, *TinyLLM: Learning a Small Student from Multiple Large Language Models*. on ArXiv, 2024.

Yu Wang, Yifan Gao, **Xiushi Chen**, Haoming Jiang, Shiyang Li, Jingfeng Yang, Qingyu Yin, Zheng Li, Xian Li, Bing Yin, Jingbo Shang, Julian McAuley, *MemoryLLM: Toward Self-Updating Large Language Models*. In Proceedings of The Forty-first International Conference on Machine Learning (**ICML**), 2024.

Bowen Jin, Hansi Zeng, Guoyin Wang, **Xiushi Chen**, Tianxin Wei, Ruirui Li, Zhengyang Wang, Zheng Li, Yang Li, Hanqing Lu, Suhang Wang, Jiawei Han, Xianfeng Tang, *Language Models as Semantic Indexers*. In Proceedings of The Forty-first International Conference on Machine Learning (**ICML**), 2024.

Yupeng Hou, Jiacheng Li, Zhankui He, An Yan, **Xiushi Chen**, Julian McAuley, *Bridging Language and Items for Retrieval and Recommendation*. Also known as **Amazon Reviews'23**, on ArXiv, 2024. **Totaling 300k+ downloads as of September 2024.**

Yu Zhang, Bowen Jin, **Xiushi Chen**, Yanzhen Shen, Yunyi Zhang, Yu Meng, Jiawei Han, *Weakly Supervised Multi-Label Classification of Full-Text Scientific Papers*. In Proceedings of 29<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2023.

Jinliang Deng, **Xiushi Chen**, Renhe Jiang, Xuan Song, Ivor W. Tsang, *ST-Norm: Spatial and Temporal Normalization for Multi-variate Time Series Forecasting*. In Proceedings of 27<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2021.

Chao Zhang, Fangbo Tao, **Xiushi Chen**, Jiaming Shen, Meng Jiang, Brian Sadler, Michelle Vanni and Jiawei Han, *TaxonGen: Constructing Topical Concept Taxonomy by Adaptive Term Embedding and Clustering*. In Proceedings of 24<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2018. Oral.

Haorui Wang, Rongzhi Zhang, Yinghao Li, Ling kai Kong, Yuchen Zhuang, **Xiushi Chen**, Chao Zhang, *TPD: Enhancing Student Language Model Reasoning via Principle Discovery and Guidance*. In Proceedings of the 1<sup>st</sup> Conference on Language Modeling (**COLM**), 2024.

Yu Zhang, Yanzhen Shen, SeongKu Kang, **Xiushi Chen**, Bowen Jin, Jiawei Han, *Chain-of-Factors Paper-Reviewer Matching*. on ArXiv, 2023.

Yu Wang, Ruihan Wu, Zexue He, **Xiushi Chen**, Julian McAuley, *Large Scale Knowledge Washing*. on ArXiv, 2024.

Yu Wang, Chi Han, Tongtong Wu, Xiaoxin He, Wangchunshu Zhou, Nafis Sadeq, **Xiushi Chen**, Zexue He, Wei Wang, Gholamreza Haffari, Heng Ji, Julian McAuley, *Towards LifeSpan Cognitive Systems*. on ArXiv, 2024.

Jonathan Light, Min Cai, Weiqin Chen, Guanzhi Wang, **Xiushi Chen**, Wei Cheng, Yisong Yue, Ziniu Hu, *Strategist: Learning Strategic Skills by LLMs via Bi-Level Tree Search*. on ArXiv, 2024.

## OTHER PUBLICATIONS

---

Yu Wang, Xinshuang Liu, **Xiushi Chen**, Sean O'Brien, Junda Wu, Julian McAuley, *Self-Updatable Large Language Models with Parameter Integration*. on ArXiv, 2024.

Junyu Luo, Xiao Luo, **Xiushi Chen**, Zhiping Xiao, Wei Ju, Ming Zhang, *SEMIEVOL: Semi-supervised Fine-tuning for LLM Adaptation*. on ArXiv, 2024.

Jinliang Deng, **Xiushi Chen**, Renhe Jiang, Du Yin, Yi Yang, Xuan Song, Ivor W. Tsang, *Disentangling Structured Components: Towards Adaptive, Interpretable and Scalable Time Series Forecasting*. In IEEE Transactions on Knowledge and Data Engineering (**TDKE**), 2024.

Jessie W. Y. Ko, Shengquan Ni, Alexander Taylor, **Xiushi Chen**, Yicong Huang, Avinash Kumar, Sadeem Alsudais, Zuozhi Wang, Xiaozhen Liu, Wei Wang, Chen Li, Suellen Hopfer, *How the Experience of California Wildfires Shapes Twitter Climate Change Framings*. **Climatic Change**, 2024.

Alexander Taylor, Yicong Huang, Junheng Hao, Xinyuan Lin, **Xiushi Chen**, Wei Wang, Chen Li, *Data Science Tasks Using Scripts versus GUI-based Workflows: The Good, the Bad, and the Ugly*. In Proceedings of the 3rd International Workshop on Data Platform Design, Management, and Optimization. (**DataPlat@ICDE**), 2024

Shengquan Ni, Yicong Huang, Jessie W. Y. Ko, Alexander Taylor, **Xiushi Chen**, Avinash Kumar, Sadeem Alsudais, Zuozhi Wang, Xiaozhen Liu, Wei Wang, Suellen Hopfer, Chen Li, *Using Texera to Characterize Climate Change Discussions on Twitter During Wildfires*. (**KDD**), 2023. (poster paper)

Jinliang Deng, **Xiushi Chen**, Renhe Jiang, Xuan Song, Ivor W. Tsang, *A Multi-view Multi-task Learning Framework for Multi-variate Time Series Forecasting*. In IEEE Transactions on Knowledge and Data Engineering (**TDKE**), 2022.

Yu Zhang\*, Shweta Garg\*, Yu Meng, **Xiusi Chen**, Jiawei Han, *MotifClass: Weakly Supervised Text Classification with Higher-order Metadata Information*. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining (**WSDM**), 2022.

Yichao Zhou, Jyun-Yu Jiang, **Xiusi Chen**, Wei Wang. *#StayHome or #Marathon? Social Media Enhanced Pandemic Surveillance on Spatial-temporal Dynamic Graphs*. In Proceedings of The 30<sup>th</sup> ACM International Conference on Information and Knowledge Management (**CIKM**), 2021.

Jyun-Yu Jiang, Yichao Zhou, **Xiusi Chen**, Yan-Ru Jhou, Liqi Zhao, Sabrina Liu, Po-Chun Yang, Jule Ahmar, Wei Wang. *COVID-19 Surveiller: Toward a Robust and Effective Pandemic Surveillance System based on Social Media Mining*. In Philosophical Transactions of the Royal Society A (**Philosophical Transactions A**), 2022.

Seungbae Kim, **Xiusi Chen**, Jyun-Yu Jiang, Jinyoung Han, Wei Wang. *Evaluating audience loyalty and authenticity in influencer marketing via multi-task multi-relational learning*. In The 15th International AAAI Conference on Web and Social Media (**ICWSM**), 2021.

Jinliang Deng, **Xiusi Chen**, Zipei Fan, Renhe Jiang, Xuan Song, Ivor W. Tsang. 2021. *The Pulse of Urban Transport: Exploring the Co-evolving Pattern for Spatio-temporal Forecasting*. In ACM Transactions on Knowledge Discovery in Data (**TKDD**), 2021.

Yu Zhang, **Xiusi Chen**, Yu Meng, Jiawei Han, *Hierarchical Metadata-Aware Document Categorization under Weak Supervision*. In Proceedings of the Fourteenth ACM International Conference on Web Search and Data Mining (**WSDM**), 2021.

Ruirui Li, Xian Wu, **Xiusi Chen**, Wei Wang. *Self-Attentive Few-Shot Learning for New User Recommendation in Location-based Social Networks*. In Proceedings of The Web Conference 2020 (**WWW**), 2020. Oral.

Fangbo Tao\*, Chao Zhang\*, **Xiusi Chen**, Meng Jiang, Tim Hanratty, Lance Kaplan and Jiawei Han, *Doc2Cube: Automated Document Allocation to Text Cube via Dimension-Aware Joint Embedding*. In Proceedings of the 17<sup>th</sup> IEEE International Conference on Data Mining (**ICDM**), 2018. Oral.

Chang Zhou, Jinze Bai, Junshuai Song, Xiaofei Liu, Zhengchao Zhao, **Xiusi Chen**, and Jun Gao, *ATRank: An Attention-Based User Behavior Modeling Framework for Recommendation*. In Proceedings of the 32<sup>nd</sup> AAAI Conference on Artificial Intelligence(**AAAI**), 2018. Oral.

## US PATENT

---

Wenchao Yu, Bo Zong, Wei Cheng, Haifeng Chen, **Xiusi Chen**, *Knowledge Graph Alignment with Uncertainty Embedding*.

## ACADEMIC SERVICE

---

### Conference Program Committee

KDD'20-24; WSDM'22-24; WWW'24; SSL@WWW'21;  
ICML'23-24; NeurIPS'22-24; ICLR'22-24; AAAI'21-24; IJCAI-ECAI'22-24;  
ACL Rolling Review.

### Conference Subreviewer

KDD'18, CIKM'18, WISE'18, IJCAI'19

### Journal Reviewer

IEEE Transactions on Big Data (TBD), IEEE Transactions on Knowledge and Data Engineering (TKDE).

### Student Volunteer

KDD'20, KDD'23.

## HONORS AND AWARDS

---

### Best Poster Award (Honorable Mention)

*SIAM International Conference on Data Mining (SDM)*

Apr. 2023

### SIAM Student Travel Award

*SIAM*

Feb. 2023

### SIGIR Student Travel Grant

*ACM SIGIR*

Sept. 2022

### SIGIR Student Travel Grant

*ACM SIGIR*

Sept. 2021

### Leo KoGuan Fellowship

*top 2%, Peking University*

Sept. 2016

<b>Merit Student</b> <i>top 5%, Peking University</i>	Sept. 2016
<b>Graduate Student Academic Fellowship</b> <i>Peking University</i>	Sept. 2015, 2016, 2017
<b>3<sup>rd</sup> Place of 17<sup>th</sup> Chinese University Basketball Association (CUBA) Championships, Beijing Division</b> <i>On behalf of Peking University, Federation of University Sports of China</i>	Oct. 2014
<b>6<sup>th</sup> Place of the High School Students Swimming Contest of Hunan Province, 100M Breaststroke</b> <i>The Education Department of Hunan Province</i>	Sept. 2010

## TEACHING EXPERIENCE

---

<b>CS 245: Big Data Analysis</b> <i>Teaching Associate, University of California, Los Angeles</i>	Fall 2020
<b>CS 143: Database Systems</b> <i>Teaching Assistant, University of California, Los Angeles</i>	Spring 2020
<b>CS 31: Introduction to Computer Science</b> <i>Teaching Assistant, University of California, Los Angeles</i>	Fall 2019, Winter 2020
<b>Introduction to Database Systems (Honor Track)</b> <i>Teaching Assistant, Peking University</i>	Spring 2016
<b>Database Systems Principles and Technologies</b> <i>Teaching Assistant, Peking University</i>	Fall 2016

## REFERENCES

---

**Heng Ji**, Professor, Amazon Scholar  
 Department of Computer Science  
 University of Illinois Urbana-Champaign  
 ✉ [hengji@illinois.edu](mailto:hengji@illinois.edu)

**Wei Wang**, Leonard Kleinrock Chair Professor, ACM Fellow, IEEE Fellow  
 Department of Computer Science  
 University of California, Los Angeles  
 ✉ [weiwang@cs.ucla.edu](mailto:weiwang@cs.ucla.edu)

**ChengXiang Zhai**, Donald Biggar Willett Professor, ACM Fellow  
 Department of Computer Science  
 University of Illinois Urbana-Champaign  
 ✉ [czhai@illinois.edu](mailto:czhai@illinois.edu)

**Yizhou Sun**, Associate Professor  
 Department of Computer Science  
 University of California, Los Angeles  
 ✉ [yzsun@cs.ucla.edu](mailto:yzsun@cs.ucla.edu)

**P. Jeffrey Brantingham**, Professor  
 Department of Anthropology  
 University of California, Los Angeles  
 ✉ [branting@ucla.edu](mailto:branting@ucla.edu)