

Patrick (Pengcheng) Jiang

Ph.D. Candidate in Computer Science

University of Illinois at Urbana-Champaign

☎ +1-(217)-298-6575 | ✉ pj20@illinois.edu | 🌐 pat-jj.github.io | 🐙 github.com/pat-jj

Education

University of Illinois at Urbana-Champaign

Ph.D. in Computer Science

Champaign, USA

Aug. 2024 – Present

- Thesis advisor: Prof. Jiawei Han

M.S. in Computer Science (with thesis)

Aug. 2022 – May 2024

- Advisors: Prof. Jiawei Han and Prof. Jimeng Sun

Waseda University

B.E. in Computer Science and Engineering

Tokyo, Japan

Sep. 2017 – Sep. 2021

- Advisors: Prof. Shinichi Honiden and Prof. Kenji Tei

Industrial Experience

Anthropic

Research Fellow

San Francisco, USA

May 2026 – Sep. 2026

- Research in retargeting the search.

Google Research

Student Researcher

Mountain View, USA

Jun. 2025 – Feb. 2026

- Research in training agent-distilled diffusion model with RL for information retrieval. Output: [\[ICML 2026\]](#)

GEHC Research

Research Scientist Intern

Seattle, USA

May 2024 – Aug. 2024

- Developed KARE, which constructs multi-sourced clinical concept knowledge graphs and integrates them with RAG for LLM-based healthcare reasoning & predictions. Output: [\[ICLR 2025\]](#)

Ant Group

Software Engineer

Hangzhou, China

Nov. 2021 – Jul. 2022

- Developed a platform for safe ML model deployment.

Selected Lead-Author Research

2025–now [Agentic Search, RL & Alignment](#)

[Preprint](#) **P. Jiang**, Zhiyi Shi, Kelly Hong, Xueqiang Xu, Jiashuo Sun, Jimeng Sun, Hammad Bashir, and Jiawei Han, “[Harness-1: Reinforcement Learning for Search Agents with State-Externalizing Harnesses](#)”. [\[pdf\]](#) [\[code\]](#)

[Preprint](#) **P. Jiang***, Luxi He*, Jifan Zhang, Jiawei Han, and Danqi Chen, “[Measuring and Strengthening Behavioral Suppression in Language Models](#)”. [\[pdf\]](#) [\[code\]](#)

[ICML’26](#) **P. Jiang**, Judith Li, Moonkyung Ryu, R. Lily Hu, Kun Su, Zhong Yi Wan, Liam Hebert, Hao Peng, Jiawei Han, Dima Kuzmin, and Craig Boutilier, “[Efficient, Property-Aligned Fan-Out Retrieval via RL-Compiled Diffusion](#)”. [\[pdf\]](#)

[ICLR’26](#) **P. Jiang**, Lang Cao, Ruike Zhu, Minhao Jiang, Yunyi Zhang, Jiaming Shen, Jimeng Sun, and Jiawei Han, “[RAS: Retrieval-And-Structuring for Knowledge-Intensive LLM Generation](#)”. [\[pdf\]](#)[\[code\]](#)

[EMNLP’25](#) **P. Jiang**, Xueqiang Xu, Jiacheng Lin, Jinfeng Xiao, Zifeng Wang, Jimeng Sun, and Jiawei Han, “[s3: You Don’t Need That Much Data to Train a Search Agent via RL](#)”. [\[pdf\]](#)[\[code\]](#)

COLM'25 **P. Jiang**, Jiacheng Lin, Lang Cao, Runchu Tian, SeongKu Kang, Zifeng Wang, Jimeng Sun, and Jiawei Han, “[DeepRetrieval: Hacking Real Search Engines and Retrievers with Large Language Models via Reinforcement Learning](#)”. [\[pdf\]](#)[\[code\]](#)

2023–2024 Knowledge Graphs, Retrieval & LLMs

ICLR'25 **P. Jiang**, Cao Xiao, Minhao Jiang, Parminder Bhatia, Taha Kass-Hout, Jimeng Sun, and Jiawei Han, “[Reasoning-Enhanced Healthcare Predictions with Knowledge Graph Community Retrieval](#)”. [\[pdf\]](#)[\[code\]](#)

AAAI'25 **P. Jiang**, Cao Xiao, Tianfan Fu, Parminder Bhatia, Taha Kass-Hout, Jimeng Sun, and Jiawei Han, “[Bi-level Contrastive Learning for Knowledge-Enhanced Molecule Representations](#)”. [\[pdf\]](#)[\[code\]](#)

NeurIPS'24 **P. Jiang**, Lang Cao, Cao Xiao, Parminder Bhatia, Jimeng Sun, and Jiawei Han, “[KG-FIT: Knowledge Graph Fine-Tuning Upon \(LLM's\) Open-World Knowledge](#)”. [\[pdf\]](#)[\[code\]](#)

ICLR'24 **P. Jiang**, Cao Xiao, Adam Cross, and Jimeng Sun, “[GraphCare: Enhancing Healthcare Predictions with Personalized Knowledge Graphs](#)”. [\[pdf\]](#)[\[code\]](#)

NAACL'24 **P. Jiang**, Cao Xiao, Zifeng Wang, Parminder Bhatia, Jimeng Sun, and Jiawei Han, “[TriSum: Learning Summarization Ability from Large Language Models](#)”. [\[pdf\]](#)[\[code\]](#)

NAACL'24 **P. Jiang**, Jiacheng Lin, Zifeng Wang, Jimeng Sun, and Jiawei Han, “[GenRES: Rethinking Evaluation for Generative Relation Extraction in the Era of LLMs](#)”. [\[pdf\]](#)[\[code\]](#)

ACL'23 **P. Jiang**, Shivam Agarwal, Bowen Jin, Xuan Wang, Jimeng Sun, and Jiawei Han, “[Text-Augmented Open Knowledge Graph Completion via Pre-Trained Language Models](#)”. [\[pdf\]](#)[\[code\]](#)

Survey & Tutorial Papers

Preprint **P. Jiang***, Jiacheng Lin*, Zhiyi Shi*, Zifeng Wang, Luxi He, Yichen Wu, Ming Zhong, Qizheng Zhang, Peiyang Song, Heng Wang, Xueqiang Xu, Hanwen Xu, Pengrui Han, Dylan Zhang, Jiashuo Sun, Chaoqi Yang, Kun Qian, Tian Wang, Changran Hu, Manling Li, Quanzheng Li, Sheng Wang, Hao Peng, Jiaxuan You, Liyuan Liu, Pan Lu, Yu Zhang, Heng Ji, Yejin Choi, Dawn Song, Jimeng Sun, and Jiawei Han[†], “[Adaptation of Agentic AI](#)”. [\[pdf\]](#)[\[repo\]](#) (*survey*)

KDD'26 **P. Jiang**, Jiashuo Sun, Wonbin Kweon and Jiawei Han, “[Structure Shapes the Future of DataxLLM Systems: Retrieval, Structuring, and Reasoning](#)”. [\[pdf\]](#) (*tutorial*)

KDD'25 **P. Jiang***, Siru Ouyang*, Yizhu Jiao, Ming Zhong, Runchu Tian, and Jiawei Han, “[Retrieval And Structuring Augmented Generation with Large Language Models](#)”. [\[pdf\]](#) (*tutorial & survey*)

KDD'25 Ran Xu, **P. Jiang**, Linhao Luo, Cao Xiao, Adam Cross, Shirui Pan, Jimeng Sun, and Carl Yang, “[A Survey on Unifying Large Language Models and Knowledge Graphs for Biomedicine and Healthcare](#)”. [\[pdf\]](#) (*tutorial & survey*)

KDD'23 C. Yang*, Z. Wu*, **P. Jiang**, Zhen Lin, Junyi Gao, Benjamin Danek, and Jimeng Sun, “[PyHealth: A Deep Learning Toolkit for Healthcare Predictive Modeling](#)”. [\[pdf\]](#)[\[doc\]](#)[\[code\]](#) (*tutorial*)

Technical Reports & Others (full list in [Google Scholar](#))

Tech Report Hammad Bashir, Kelly Hong, **P. Jiang**, and Zhiyi Shi “[Context-1: Training a Self-Editing Search Agent](#)”, 2026. [\[report\]](#)[\[code\]](#)

User Study **P. Jiang**, Megan Lim, Adam Cross, and Jimeng Sun, “[MedKG: Empowering Medical Education with Interactive Construction and Visualization of KGs via LLMs](#)”, 2023. [\[pdf\]](#)[\[code\]](#)[\[software\]](#)

SSCI'21 **P. Jiang**, Kenji Tei, “[OACAL: Finding Module-consistent Specifications to Secure Systems from Weakened User Obligations](#)”. [\[pdf\]](#)

Invited Talks & Tutorials

- 2026.07 Search Agent Training with a State-externalizing Harness, AlphaSense
- 2025.10 Poster Talk (s3) invited by Meta PhD Forum, Meta
- 2025.10 “RL for Retrieval” Guest Talk invited by CS598JH, UIUC
- 2025.08 Google Research Internal Talk, Google Research
- 2025.08 “MedKLM: Unifying Large Language Models and Knowledge Graphs for Medicine and Healthcare”, KDD’25 Tutorial
- 2025.08 “Retrieval And Structuring Augmented Generation with Large Language Models”, KDD’25 Tutorial
- 2025.06 General Research Talk, invited by Hong Kong University of Science and Technology (HKUST)
- 2025.05 “DeepRetrieval: Pushing the Boundaries of Information Retrieval with LLMs”, invited by National Institutes of Health (NIH)
- 2024.12 “Reasoning-Enhanced Healthcare Predictions with Knowledge Graphs”, invited by BioNLP Group at UMass Amherst
- 2023.08 PyHealth Hands-on Tutorial, KDD’23 Tutorial

Professional Services

Paper Reviewer: NeurIPS, ICLR, ICML, EMNLP, NAACL, ACL, AAAI, AISTATS, KDD, TMLR