

孙宇石

手机: +86 18907174483

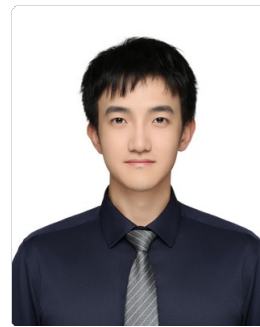
微信: sys18907174483

年龄: 27

邮箱: ysunbp@gmail.com

性别: 男

个人链接: <https://github.com/ysunbp/>



教育及工作经历

腾讯科技有限公司

互动娱乐事业群 光子工作室 | 高级研究员

2025.08 – 至今

研究方向: 强化学习 (AI Bot), Data Agent, Agent Memory

香港科技大学

计算机科学与工程 | 博士

2021.09 – 2025.08

研究方向: 数据管护 (Data Curation, 包含数据标注, 整合和组织), 大语言模型, 检索增强生成 (RAG)

导师: 陈雷教授 (IEEE Fellow, ACM Fellow)

获香港政府奖学金 (HKPFS) 资助

香港科技大学

计算机科学与应用数学 | 本科

2017.09 - 2021.06

绩点: 4.03/4.30, top 1% in HKUST

一等荣誉学位毕业

赴新加坡南洋理工大学 (NTU) 和瑞士苏黎世联邦理工学院 (ETH) 交换

发表论文 (节选) (*共同一作, +通讯作者)

MedKGI: Iterative Differential Diagnosis with Medical Knowledge Graphs and Information-Guided Inquiring

Qipeng Wang, Rui Sheng, Yafei Li, Huamin Qu, Yushi Sun+, Min Zhu+

Submitted to ACL 2026

[KG]

LakeHopper: Cross Data Lakes Column Type Annotation through Model Adaptation

Yushi Sun, Nan Tang, Lei Chen

Submitted to KDD 2026

[Data Preparation]

CacheRAG: A Novel Approach to Enhance KG-based RAG Through Caching Mechanisms

Yushi Sun, Lei Chen

Under Revision VLDB 2026

[LLM, RAG, KG]

KERAG: Knowledge-Enhanced Retrieval-Augmented Generation for Advanced Question Answering

Yushi Sun, Kai Sun, Yifan Ethan Xu, Xiao Yang, Xin Luna Dong, Nan Tang, Lei Chen

EMNLP 2025

[LLM, RAG, KG]

CRAG - Comprehensive RAG Benchmark

Xiao Yang*, Kai Sun*, Hao Xin*, Yushi Sun*, Nikita Bhalla, Xiangsen Chen, Sajal Choudhary, Rongze Daniel Gui, Ziran Will Jiang, Ziyu Jiang, Lingkun Kong, Brian Moran, Jiaqi Wang, Yifan Ethan Xu, An Yan, Chenyu Yang, Eting Yuan, Hanwen Zha, Nan Tang, Lei Chen, Nicolas Scheffer, Yue Liu, Nirav Shah, Rakesh Wanga, Anuj Kumar, Wen-tau Yih,

Xin Luna Dong

(* indicates equal contribution)

NeurIPS 2024 (ccf-A), collaboration with Meta Reality Labs [LLM, RAG, Data Organization]

Are Large Language Models a Good Replacement of Taxonomies?

Yushi Sun, Hao Xin, Kai Sun, Yifan Ethan Xu, Xiao Yang, Xin Luna Dong, Nan Tang, Lei Chen

VLDB 2024 (ccf-A), collaboration with Meta Reality Labs [LLM, Data Organization]

Cross-domain-aware Worker Selection with Training for Crowdsourced Annotation

Yushi Sun, Jiachuan Wang, Peng Cheng, Libin Zheng, Lei Chen, Jian Yin

ICDE 2024 (ccf-A)

[Data Annotation]

RECA: Related Tables Enhanced Column Semantic Type Annotation Framework

Yushi Sun, Hao Xin, Lei Chen

VLDB 2023 (ccf-A)

[Data Preparation]

在校经历+实习经历

蚂蚁研究院实习

2025.02 - 2025.05

与蚂蚁研究院的团队探索了多知识图谱的检索增强生成，提交了两份专利提案，计划以第一作者的身份投稿论文。

参与共同主办KDD Cup 2024

2024.01 - 2024.08

与Meta Reality Labs的团队共同主办了KDD Cup 2024的竞赛。主要负责比赛数据准备的工作，之后以共同一作的身份在NeurIPS 2024上发表了题为CRAG-Comprehensive RAG Benchmark的论文。

荣誉奖项

红鸟奖学金 (2024/06)

红鸟奖学金 (2023/06)

香港政府奖学金 (2021/09)

香港政府奖学金 - Reaching Out Award (2020/07) Lee Hysan

Foundation Scholarship (2020/06)

University's Scholarship Scheme for Continuing Undergraduate Students (2019/12) The Cheng

Foundation Scholarships for Mainland Students (2019/02)

Pilot Mainland Experience Scheme for Post-Secondary Students (2019/01) University's Scholarship

Scheme for Continuing Undergraduate Students (2018/12) Admission Scholarship for student with an

outstanding curriculum (2017/06) Scholarship for outstanding students in Hubei by Tsinghua

University (2017/03)