SHI QIU

(+1) 919-923-9983 Email: stephenshiqiu@gmail.com ◇ Homepage ◇ Google Scholar

EDUCATION

University of North Carolina at Chapel Hill 2026.09 - now
Ph.D. in Computer Science (incoming); Advisor: Prof. Huaxiu Yao

Peking University 2022.09 - 2026.07
B.S. in Astronomy, School of Physics

RESEARCH INTEREST

My research focuses on Agentic AI and AI4Physics, with an emphasis on self-evolving autoresearch systems for scientific discovery — agents that iteratively generate hypotheses, design and run experiments, and write papers with minimal human intervention. I explore two directions: (1) building robust, self-evolving agentic frameworks that automate the full research loop; and (2) grounding agents in domain-specific physics knowledge for rigorous scientific reasoning.

My open-source work has collectively earned over **22,000 GitHub stars**. I hold a **U.S. permanent resident green card** and am actively seeking for internship opportunities.

SELECTED PUBLICATIONS

*: equal contributions. Full publication list is in Google Scholar.

[1] **Shi Qiu**, Shaoyang Guo, Zhuo-Yang Song, Yunbo Sun, Zeyu Cai, Jiashen Wei, ..., Qing-Hong Cao, Hua Xing Zhu, PHYBench: Holistic Evaluation of Physical Perception and Reasoning in Large Language Models, in *NeurIPS 2025*. [2025]

[2] Jiaqi Liu, **Shi Qiu**, Mairui Li, Bingzhou Li, James Zou, Caiming Xiong, Huaxiu Yao et al, AutoResearchClaw: Self-Reinforcing Autonomous Research with Human-AI Collaboration, *arXiv preprint*. **Github 12.5k+ stars.**[project] [2026]

[3] **Shi Qiu**, Zeyu Cai, Jiashen Wei, Zhuo-Yang Song, Xing-Bo Yuan, Hua Xing Zhu, Qing-Hong Cao, Ming-Xing Luo, ColliderAgent: An LLM-Powered System for End-to-End Collider Phenomenology, under review. [2026]

[4] Haonian Ji*, **Shi Qiu***, Siyang Xin, Siwei Han, Zhaorun Chen, Dake Zhang, Hongyi Wang, Huaxiu Yao, From EduVisBench to EduVisAgent: A Benchmark and Multi-Agent Framework for Reasoning-Driven Pedagogical Visualization, in *ICLR 2026*. [2026]

[5] **Shi Qiu**, Zhun Wang, Tianneng Shi, Zhaorun Chen, Wenbo Guo, Dawn Song, AgentXploit: End-to-End Red-Teaming for AI Agents Powered by Multi-Agent Systems, *under review*. [2025]

[6] Peng Xia*, Siwei Han*, **Shi Qiu***, Yiyang Zhou, Zhaoyang Wang, Wenhao Zheng, Zhaorun Chen, Chenhang Cui, Mingyu Ding, Linjie Li, Lijuan Wang, Huaxiu Yao, MMIE: Massive Multimodal Interleaved Comprehension Benchmark for Large Vision-Language Models, in *ICLR 2025 (Oral)*. [Paper] [Code] [2024]

[7] Yiyang Zhou*, Linjie Li*, **Shi Qiu***, Haibo Tong, Lijuan Wang, Huaxiu Yao, GLIMPSE: Do Large Vision-Language Models Truly Think With Videos or Just Glimpse at Them?, in *EMNLP 2025 Main Conference (Oral)*. [2025]

[8] Haibo Tong, Zhaoyang Wang, Zhaorun Chen, Haonian Ji, **Shi Qiu**, Siwei Han, Zhongkai Xue, Yiyang Zhou, Peng Xia, Kexin Geng, Mingyu Ding, Rafael Rafailov, Chelsea Finn, Huaxiu Yao MJ-VIDEO: Fine-Grained Benchmarking and Rewarding Video Preferences in Video Generation, in *NeurIPS 2025 (Spotlight)*. [Paper] [2025]

[9] Wangchunshu Zhou, Yuchen Eleanor Jiang, Long Li, Jialong Wu, **Shi Qiu**, Shuai Wang, Jiamin Chen, Tiannan Wang, Jintian Zhang, Jing Chen, Wentao Zhang, Jiyu Chen, Ruipu Wu, Shiding Zhu, Xiangru Tang, Peng Cui, Huajun Chen, Ningyu Zhang, Mrinmaya Sachan, Agents: An Open-source Framework for Autonomous Language Agents, in *ICLR 2024 Workshop on Large Language Models for Agents.*[Paper], **Github 5.4k stars.**[project] [2023]

HONORS AND AWARDS

Beijing Municipal Outstanding Graduate **2026**
Peking University Outstanding Graduate **2026**
Merit Student Pacesetter, Peking University **2026**
National Scholarship **2025**
Peking University Excellence in Research Award **2024**
Peking University Third Class Scholarship **2024**
Shu Qi Scholarship **2023**

RESEARCH EXPERIENCES

Sunblaze Group, UC Berkeley — Research Intern 2025 - Now
Mentor: Prof. Dawn Song. Lead researcher of AgentXploit[5], an open-source, end-to-end agentic framework for automatic injection-point exploitation and validation.

AI4Physics Group, Peking University — Research Intern 2025 - Now
Mentors: Prof. Hua Xing Zhu, Prof. Qing-Hong Cao, Prof. Ming-Xing Luo. Initiator of PHYBench[1], coordinating 178 students to build the first high-difficulty physics-reasoning benchmark; core contributor to IdeaSearch for LLM-assisted optimization in HEP, astronomy, and fluid dynamics.

UNC-Chapel Hill — Research Intern 2024 - Now
Mentor: Prof. Huaxiu Yao. Multimodal alignment and evaluation: co-led MMIE[6] (20,103-sample interleaved benchmark with an MLLM-as-judge metric), proposed GLIMPSE[7], and released the MJ-Video reward model[8].

AIWaves — Research Intern 2023
Mentor: Wangchunshu Zhou. Core developer of the open-source Agents[9] framework for self-evolving autonomous language agents.

Advisor, AlphaXiv 2025 - Now
alphaxiv.org/about

EXTRACURRICULAR ACTIVITIES

President of the Peking University Jump-Rope Team 2023-2024

- Beijing College Shuttlecock and Rope Skipping League: Men's 30-Second Double Under: First Place

Class Monitor of Class 8 in the 22nd Undergraduate Cohort 2022-Now

ACADEMIC SKILLS

English Skills GRE **326** | TOEFL **112** | CET-6 **696**