Qingyi Si

(+86)131-6168-5288 | 1997.12.01 | siqingyi@iie.ac.cn | https://phoebussi.github.io/ |
https://github.com/PhoebusSi | @twit | Beijing, China | My research interests include OOD generalization, VQA, visual & language, and LLMs.

Education

University of Chinese Academy of Sciences (UCAS) *Beijing*

PhD in Cyberspace Security; **MS** in Computer Applied Technology Sep. 2019 - Jun. 2024

Beijing Language and Culture University (BLCU) *Beijing*

BS in Computer Science and Technology Sep. 2015 - Jun. 2019

Publications (Co-*) First author of 10 papers

[1] **Qingyi Si**, Yuchen Mo, Zheng Lin, et al. "Combo of Thinking and Observing for Outside-Knowledge VQA". ***ACL'23***

[2] **Qingyi Si**, Fandong Meng, et al. "Language Prior Is Not the Only Shortcut: A Benchmark for Shortcut Learning in VQA". ***EMNLP'22***
*ps: This paper received **high praise** from **Damien Teney**.*

[3] **Qingyi Si**, Fandong Meng, et al. "Towards Robust VQA: Making the Most of Biased Samples via Contrastive Learning". ***EMNLP'22***

[4] **Qingyi Si**, Zheng Lin, et al. "Check It Again: Progressive Visual Question Answering via Visual Entailment". ***ACL'21***

[5] **Qingyi Si**, Yuanxin Liu, et al. "Learning Class-Transductive Intent Representations for Zero-shot Intent Detection". ***IJCAI'21***

[6] Zhe Wen*, **Qingyi Si***, Zheng Lin, et al. "Schema Item Matters in Knowledge Base Question Answering". *IJCNN'23*

Under Review Papers

[7] **Qingyi Si**, Yuanxin Liu et al. "Compressing And Debiasing Vision-Language Pre-Trained Models for VQA". ***EMNLP'23***

[8] **Qingyi Si**, Tong Wang et al. "An Empirical Study of Instruction-tuning Large Language Models in Chinese". ***EMNLP'23***

[9] Huishan Ji*, **Qingyi Si***, Zheng Lin, et al. "Towards Many-to-one Visual Question Answering". ***NeurIPS'23***

[10] Jie Xu*, **Qingyi Si***, Hanbo Zhang*, et al. "TiO: A Unified Interactive Visual-Language Disambiguation Transformer". ***NeurIPS'23***

Co-author Papers

[11] Duo Zheng, Fandong Meng, **Qingyi Si**, et al. "Visual Dialog for Spotting the Differences between Pairs of Similar Images". ***MM'22***

[12] Ran Li, **Qingyi Si**, et al. "A Multi-channel Neural Network for Imbalanced Emotion Recognition". *ICTAI'19*

[13] Chunhua Liu, Yan Zhao, **Qingyi Si**, et al. "Multi-perspective fusion network for commonsense reading comprehension".*CCL'18*

Research Projects

Alpaca-CoT (***https://github.com/PhoebusSi/alpaca-CoT/***) ***1.7k Star***

An Instruction-tuning Platform with Instruction Data Collection and Unified Large Language Models Interface ***2.4k Downloads last month***

We unified the interfaces of instruction-tuning data, multiple LLMs and parameter-efficient methods (e.g., lora, p-tuning) together for easy use. Besides, we are the first to extend CoT data into LLaMA to improve its reasoning ability. We constructed the largest collection of instruction tuning datasets currently available. I was invited by Machine Heart (jiqizhixin) to give a talk about this project.

A Multimodal Large Language Model ***In Progress***

A small and open-source version of GPT-4, which empowers LLMs with multi-modality capabilities.

Awards and Honors

[1] **National Scholarship for Doctoral students (Top 2%, RMB ¥ 30,000).** *Ministry of Education of P.R. China. 2021.*

[2] **Merit Student, University of Chinese Academy of Sciences** *UCAS. 2020, 2021.*

[3] **Excellent Undergraduate Graduation Project (Thesis).** *Education Commission of Beijing. 2019.*

[4] **Regular Institutions of Higher Education Outstanding Graduate.** *Education Commission of Beijing. 2019.*

[5] **National Scholarship for Undergraduates (Top 2%, RMB ¥ 8,000).** *Ministry of Education of P.R. China. 2017, 2018.*

Working Experience

Microsoft Research Asia (MSRA) *Beijing*

Research Intern in NLC group November.2021 - May.2022

WeChat, Tencent *Beijing*

Research Intern in the Pattern Recognition Center (PRC), WeChat AI March.2021 - November.2021