KAIYUAN GAO

(086) 152 7187 6956 ◇ im_kai@hust.edu.cn ◇ https://kygao.github.io

SHORT BIO

I am a final-year Ph.D. candidate in a joint program between Huazhong University of Science and Technology and Microsoft Research, focusing on **multimodal generative models**. During my time at Microsoft Research, I worked on **3D molecular structure modeling** and molecule generation using diffusion and language models, with in-depth research on molecular docking. Currently, I am a core contributor to the **Qwen-Image** project, where I actively contribute to research on image editing and the development of visual generation within the Qwen team.

EDUCATION

Huazhong University of Science and Technology September 2021 - Present
Ph.D. in Computer Science *Wuhan, China*

· Joint Ph.D. Program with **Microsoft Research AI4Science**
· Ph.D Supervisor: Tie-yan Liu and Kun He
· Overall GPA: 90.84/100.00 (rank 1/14)

Huazhong University of Science and Technology September 2017 - June 2021
B.Eng. in Artificial Intelligence and Automation *Wuhan, China*

· Overall GPA: 3.95/4.00 (90.3/100.00)

RESEARCH EXPERIENCE

Renmin University of China December 2024 - Present
collaboration with Min Zhao and Chongxuan Li *Beijing, China*

· Work on diffusion distillation on video domain.

Microsoft Research AI4Science April 2022 - Present
Joint-PhD in MSR AI4Sci Group, mentored by Lijun Wu *Beijing, China*

· Work on 3D structure modeling of protein-protein and protein-ligand interactions.

Microsoft Research Asia November 2020 - July 2021
Research Intern in Machine Learning Group *Beijing, China*

· Work on language encoding approaches in multilingual Transformer.

PROJECT EXPERIENCE

Qwen-Image Project April 2025 - Present
https://github.com/QwenLM/Qwen-Image *Beijing, Qwen VGen Team*

· The most powerful open-sourced text-to-image model in terms of text generation fidelity.
· Precise and ID-preserving image editing capabilities.

Physics Structure Model July 2023 - December 2023, June 2024 - November 2024
[Key words] Diffusion, DiT, AlphaFold 3 *Beijing, MSR AI4Sci*

· A Transformer-based structure foundation model trained using diffusion algorithms on molecular, protein, material, and complex structures and force fields.
· Developing DDPM training and inference code on pocket-given and blind docking task.

Nature Language Model July 2023 - June 2024
https://naturelm.github.io *Beijing, MSR AI4Sci*

· A generative foundation model continue-trained and instruction-tuned from powerful open-source LLMs by converting all scientific modalities into wrapped sequences.
· Processing protein and molecule data and developing instruction tuning code for LLaMA3 pipeline.

SELECTED PUBLICATIONS

[**Tech Report**] Chenfei Wu, Jiahao Li, Jingren Zhou, Junyang Lin, **Kaiyuan Gao**, Kun Yan, Shengming Yin, Shuai Bai, Xiao Xu, Yilei Chen, Yuxiang Chen, Zecheng Tang, Zekai Zhang, Zhengyi Wang, An Yang, Bowen Yu, Chen Cheng, Dayiheng Liu, Deqing Li, Hang Zhang, Hao Meng, Hu Wei, Jingyuan Ni, Kai Chen, Kuan Cao, Liang Peng, Lin Qu, Minggang Wu, Peng Wang, Shuting Yu, Tingkun Wen, Wensen Feng, Xiaoxiao Xu, Yi Wang, Yichang Zhang, Yongqiang Zhu, Yujia Wu, Yuxuan Cai, Zenan Liu. "Qwen-Image Technical Report." Arxiv abs/2508.02324 (2025)

[**Preprint**] **Kaiyuan Gao**, Yusong Wang, Haoxiang Guan, Zun Wang, Qizhi Pei, John E. Hopcroft, Kun He, Lijun Wu. "Tokenizing 3D Molecule Structure with Quantized Spherical Coordinates." Arxiv abs/2412.01564 (2024).

[**KDD**] **Kaiyuan Gao**, Qizhi Pei, Gongbo Zhang, Jinhua Zhu, Tao Qin, Kun He, Tie-Yan Liu and Lijun Wu. "FABind+: Enhancing Molecular Docking through Improved Pocket Prediction and Pose Generation." Proceedings of the 31th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (2025).

[**KDD**] Yangzhe Peng*, **Kaiyuan Gao***, Liang He, Yuheng Cong, Haiguang Liu, Kun He, Lijun Wu. "CovDocker: Benchmarking Covalent Drug Design with Tasks, Datasets, and Solutions" Proceedings of the 31th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (2025).

[**KDD**] **Kaiyuan Gao**, Lijun Wu, Jinhua Zhu, Tianbo Peng, Yingce Xia, Liang He, Shufang Xie, Tao Qin, Haiguang Liu, Kun He and Tie-Yan Liu. "Pre-training Antibody Language Models for Antigen-Specific Computational Antibody Design." Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (2023).

[**NeurIPS**] Qizhi Pei* , **Kaiyuan Gao*** , Lijun Wu , Jinhua Zhu, Yingce Xia, Shufang Xie, Tao Qin, Kun He, Tie-Yan Liu, Rui Yan. "FABind: Fast and Accurate Protein-Ligand Binding." Advances in Neural Information Processing Systems (2023).

[**ICLR**] Jinsong Chen*, **Kaiyuan Gao***, Gaichao Li and Kun He. "NAGphormer: A Tokenized Graph Transformer for Node Classification in Large Graphs." International Conference on Learning Representations (2023).

[**Preprint**] **Kaiyuan Gao*** , Sunan He* , Zhenyu He* , Jiacheng Lin* , Qizhi Pei* , Jie Shao* , Wei Zhang*. "Examining User-Friendly and Open-Sourced Large GPT Models: A Survey on Language, Multimodal, and Scientific GPT Models." Arxiv abs/2308.14149 (2023).

RELEVANT PUBLICATIONS

[**ICLR**] Zizhuo Zhang, Lijun Wu, **Kaiyuan Gao**, Jiangchao Yao, Tao Qin, Bo Han. "Fast and Accurate Blind Flexible Docking." International Conference on Learning Representations (2025).

[**ICLR**] Qizhi Pei, Lijun Wu, **Kaiyuan Gao**, Jinhua Zhu, Rui Yan. "3D-MolT5: Towards Unified 3D Molecule-Text Modeling with 3D Molecular Tokenization." International Conference on Learning Representations (2025).

[**ACL Findings**] Qizhi Pei, Lijun Wu, **Kaiyuan Gao**, Xiaozhuan Liang, Yin Fang, Jinhua Zhu, Shufang Xie, Tao Qin, Rui Yan. "BioT5+: Towards Generalized Biological Understanding with IUPAC Integration and Multi-task Tuning." Findings of the Association for Computational Linguistics ACL (2024).

[**EMNLP**] Qizhi Pei, Wei Zhang, Jinhua Zhu, Kehan Wu, **Kaiyuan Gao**, Lijun Wu, Yingce Xia, Rui Yan. "BioT5: Enriching Cross-modal Integration in Biology with Chemical Knowledge and Natural Language Associations." Empirical Methods in Natural Language Processing (2023).

[**Preprint**] Shengjie Luo*, **Kaiyuan Gao***, Shuxin Zheng, Guolin Ke, Di He, Liwei Wang, Tie-Yan Liu. "Revisiting language encoding in learning multilingual representations." ArXiv abs/2102.08357 (2021).