Yuxuan Ren

Curriculum Vitae

yuxuan.ren@nus.edu.sg — (65) 85851276

EDUCATION

Ph.D. Student in Computer Science
Year-Year: 2026-present
National University of Singapore
School of Computing
Advisor: Yatao Bian

Bachelor of Technology in Chemistry
Year of completion: 2024
University of Science and Technology of China
School of Chemistry and Materials Science

- Overall GPA: 3.45/4.3 Rank: 27/134

EMPLOYMENT

Research Intern
Year-Year: 2023-2024
Microsoft Research AI for Science
Worked on developing the SFM system, a foundational model designed for multitask learning in the chemical domain.
Mentor: Chang Liu

Research Scientist
Year-Year: 2024-2025
Beijing Academy of Artificial Intelligence, Health computing
I worked on OpenCryo and multimodal Protein–protein-language representation and generation.
Mentor: Qiwei Ye

Research Intern
Year-Year: 2026
Zhongguancun Academy & Zhongguancun Institute of Artificial Intelligence
I worked on the post-training of ZGCM-1, a fully open 7B foundation model for math and agentic search.
Mentor: Jiyan He

RESEARCH EXPERIENCE

Scientific Foundation Model & Physical Consistency in Molecular Science Multi-Task Learning
Advisors: Dr. Chang Liu & Dr. Bin Shao Dec. 2023 - Jun. 2024
Microsoft Research AI for Science

- Investigated multi-task learning approaches and current molecular datasets in molecular science.
- Implemented diffusion processes and energy prediction terms for foundational models.
- Enhanced model performance by leveraging scientific laws across tasks, connecting diffusion generation to the energy landscape based on Boltzmann distribution equations and the principle of minimum energy for DFT optimized structures.

OpenCryo
Advisors: Dr. Kai Chen July. 2024 - Dec. 2025
Beijing Academy of Artificial Intelligence, Health computing

- Investigated deep learning-based inverse problem solutions and Cryo-ET data processing pipelines.
- Introduced an energy-based framework for inverse problems, applying it to solve the missing wedge and low signal-to-noise ratio issues in Cryo-ET.

Cross-domain foundation models (SinAE)
Advisors: Dr. Yatao Bian Jan. 2026 - Present
National University of Singapore, Blue Whale Lab

- Developed SinAE, a single-architecture flow-matching autoencoder that serves as a universal, near-lossless geometric tokenizer for molecules, crystals, proteins and metal-organic frameworks (MOFs).
- Shifted the reconstruction burden from the encoder to an iterative flow-matching decoder, reaching near-zero reconstruction RMSD and accurate recovery of chemical identity across domains.
- Relied only on vanilla Transformer backbones, removing domain-specific inductive biases such as explicit graph modeling or equivariant message passing.
- Trained a standard Diffusion Transformer prior in the shared latent space for competitive 3D de novo generation in all three domains, and showed that joint molecule–crystal training improves both, evidence of cross-domain transfer.

Topology-aware molecular language (STAMP)
Advisors: Dr. Yatao Bian 2026 - Present
National University of Singapore, Blue Whale Lab

- Designed STAMP, a fragment-level molecular language that encodes motif identity, attachment sites and tree topology separately, so that fragments are reusable across contexts and subtrees can be edited directly.
- Built the full framework: deterministic encoding and decoding, motif-vocabulary construction, any-order autoregressive generation with fragment infilling, and TreeCrossover optimization.
- Lowered activity-cliff prediction error on 23 of 30 targets against a matched SMILES encoder, and ranked first among six methods on 15 of 23 scaffold-preserving optimization objectives.
- Let frozen language models operate on indexed fragments through a deterministic compiler, improving molecular editing and related tasks by 9 to 56 percentage points over plain SMILES.

ZGCM-1 post-training
Advisors: Dr. Jiyan He 2026
Zhongguancun Academy & Zhongguancun Institute of Artificial Intelligence

- Contributed to the post-training of ZGCM-1, a fully open 7B dense foundation model trained from scratch that pairs internal reasoning with external tool use for math and agentic search.
- Post-training combines supervised fine-tuning on mixed general and agentic data with GRPO-based reinforcement learning using verifiable math and code rewards.
- The released 7B model stays competitive with much larger models on math reasoning and agentic search suites (75.0% on AIME 2026, 97.1% on MATH-500, 63.1% on WebWalkerQA).

PUBLICATIONS

Preprints

SinAE: A Single-Architecture Flow-Matching Autoencoder for Cross-Domain Atomic Systems
Yuxuan Ren, Fan Yang, Jianhua Yao, Yatao Bian, arXiv:2607.12380, 2026.

STAMP as a Topology-Aware Molecular Language for Fragment-Level Learning and Editing
Yuxuan Ren, Jiatong Li, Ziqiao Meng, Yatao Bian, arXiv:2610.18438, 2026.

ZGCM-1: A Fully Open and Extremely Efficient Foundation Model for Math and Agentic Search
ZGCM Team (including **Yuxuan Ren**), arXiv:2609.13356, 2026.

Chemical Chain-of-Thought Functions as a Hallucination-Prone Molecular Scratchpad
Jiatong Li, **Yuxuan Ren**, Weida Wang, Xiao-Yong Wei, Yatao Bian, arXiv:2607.20935, 2026.

Conference Proceedings

EPO: Diverse and Realistic Protein Ensemble Generation via Energy Preference Optimization
Yuancheng Sun, **Yuxuan Ren**, Zhaoming Chen, Xu Han, Kang Liu, Qiwei Ye, The Fortieth AAAI Conference on Artificial Intelligence (AAAI 2026).

CryoGEN: Cryogenic Electron Tomography Reconstruction via Generative Energy Nets
Yunfei Teng*, **Yuxuan Ren***, Kai Chen, Xi Chen, Zhaoming Chen, Qiwei Ye, 13th International Conference on Learning Representations (ICLR 2025).

Physical Consistency Bridges Heterogeneous Data in Molecular Multi-Task Learning
Yuxuan Ren*, Dihan Zheng*, Chang Liu, Peiran Jin, Yu Shi, Lin Huang, Jiyan He, Shengjie Luo, Tao Qin, Tie-Yan Liu, 38th Conference on Neural Information Processing Systems (NeurIPS 2024).

SELECTED COURSES

Stochastic Process (A+)	Probability Theory (A)	Machine Learning Methods (A-)
Operating System (A)	Data Intelligence and Material Science (A)	Algorithm Basics (A-)
Calculus (A+)	Organic Chemistry (A+)	Statistical Mechanics (A)
Quantum Chemistry (A)		