Ziqi Wang

ziqiw@umich.edu · (617) 309-9097 · linkedin.com/in/ziqi-wang · arxiv.org/abs/2507.14267

Trustworthy scientific LLM agent researcher. First author of DREAMS.

EDUCATION

Ph.D., Mechanical Engineering and Scientific Computing 2021–2026 (expected)
Carnegie Mellon University & University of Michigan · Advisor: Prof. Venkatasubramanian Viswanathan

B.S., Mechanical Engineering & B.S., Computer Science 2016–2021
University of Michigan & Florida Institute of Technology

RESEARCH EXPERIENCE

DREAMS First author · arXiv 2025

Fully autonomous LLM agent that runs expert-level materials simulations

- Reached human-expert accuracy on multi-step autonomous workflows: reproduced a long-debated surface-chemistry benchmark (CO/Pt(111)) to within 0.2% and matched expert structures on all 27 systems of a crystal benchmark, where a single-agent baseline failed every run and a multi-agent baseline missed by 389%.
- Canvas: a shared communication-and-memory scheme for long-horizon, complex runs, with report-based history compression suited to scientific workflows, giving finer control over agent context than general-purpose agent frameworks.
- Deterministic and LLM-based safety guards that vouch for each result's credibility, validating every parameter setting against customizable rule levels (R1, R2, …) and user-defined sensitive parameters.
- Full traceability and white-box transparency: every result is registered with its inputs and outputs, verified by reference ID (ref_id), and linked to its sources through a provenance DAG, with each claim backed by a context-aware reasoning chain annotated with those IDs.
- Enables recursive, deep debugging of long, multi-step runs.
- Benchmarked against representative single-agent and multi-agent frameworks, reporting accuracy, success rate, and token usage (cost), using up to ~3x fewer tokens than the baseline framework on long-horizon tasks.

DREAMS-OER Manuscript in preparation

Autonomous agent for open-ended catalyst discovery over a ~380,000-material space

- Scaled the agent from a single fixed task to open-ended search over ~380,000 candidate materials (Google DeepMind's GNoME set), choosing what to study next (material, surface, site, three intermediates) from accumulated evidence rather than a fixed plan. (Demonstrated on behavioral runs; production screening in progress.)
- Built a multi-level relational experiment log, enabling efficient information retrieval and progress tracking across hundreds of partial, interdependent studies.
- Gave the agent live time- and compute-budget awareness to submit work opportunistically, keep the cluster busy while reasoning, and re-plan under pressure, completing a bounded 7-hour study without overrunning.
- Hardened it against scale-only failure modes: a disposition gate enforcing genuine engagement with each result (removing a queue-occupancy gaming pattern), a runaway-loop guard, retrieval-augmented literature grounding, and required hypothesis/limitation logging, including a “too credulous” failure now being addressed.

Non-linear Material-Property Prediction with Machine-Learning Interatomic Potentials Manuscript in preparation

- Built a committee-based active-learning loop to fine-tune several universal ML interatomic potentials (MACE, NequIP, GRACE, MatterSim) into accurate, low-cost surrogates for the Li-Mg alloy system, cutting prediction error on target physical properties by roughly 5x versus off-the-shelf models.
- Showed that weighting training toward higher-order (stress) information, not just energy, most improves accuracy on hard second-derivative properties, with the gain transferring across multiple properties.
- Mapped how alloying produces non-linear excess effects in stiffness and atomic-transport barriers across composition, quantifying behavior that cannot be inferred from the pure end-members alone.

PUBLICATIONS & PREPRINTS

- **Z. Wang**, H. Huang, H. Zhao, C. Xu, S. Zhu, J. Janssen, V. Viswanathan. “DREAMS: Density Functional Theory Based Research Engine for Agentic Materials Simulation.” arXiv:2507.14267 (2025).

TECHNICAL SKILLS

Agentic AI / LLM: multi-agent system design, harness design, SKILL design, guardrail and provenance systems, agent evaluation and benchmarking, LangGraph, LangChain.
Machine Learning: PyTorch, machine learning interatomic potentials (MACE, NequIP, MatterSim, GRACE, SchNet), graph neural networks, active learning
Atomistic Simulation: DFT (Quantum ESPRESSO), ASE, molecular dynamics (LAMMPS), Monte Carlo, nudged elastic band, Bayesian error estimation (BEEF)
Computing: Python, C++, HPC / SLURM, Git

HONORS & AWARDS

- Selected for Anthropic AI for Science Program (2026)
- University Honors, University of Michigan (2019, 2020, 2021)