Vasily Ilin

Website: vilin97.github.io | Google Scholar: Vasily Ilin | Github: Vilin97 | email: vasilin97@gmail.com

Education

University of Washington | 2020 - 2026

Applied Mathematics, PhD, $15,000 in awards.
Cofounder & Deputy Director of the Math AI Lab, $76,000 in funding.

Boston University | 2015 - 2020

Computer Science, MS.
Mathematics, BA & MA (Magna Cum Laude). Thesis: "Stochastic Simulation Algorithms and Benchmarks", paper.

Skills

Technical Skills: Python, Jax, PyTorch, WandB, SQL, Julia, Java, Lean, cluster computing
CS & Math Skills: AI for Math, deep learning, plasma modeling, data engineering, diffusion, sampling, formalization

Work Experience

AI for Math Researcher at Math.Inc | October 2025-present

Accelerating math research and math formalization with AI agents.

AI Intern at Google Cloud | Summer 2025

Improved a Slides2Video model with a multi-agentic APO algorithm. Achieved F1 score of 0.9. Collaborated with two teams in Google DeepMind. Gave 6 research talks. Was acknowledged in a memo to 4 Vice Presidents and 3 Directors.

ML Intern at Google Cloud (Python, SQL) | Summer 2024

Trained XGBoost to predict defective TPUs from HBM ECC telemetry, achieved 80% accuracy and improved test efficiency 4x.

ML Intern at YouTube (Python, C++, SQL) | Summer 2023

Trained model to achieve a 2x Egress/Ingress improvement in a CDN. Proposed and implemented a load balancing algorithm.

Data Engineering Intern at Android, Google (Flume Java, SQL) | Summer 2022

Built a distributed pipeline for Quick Share, from logs to dashboards. Sped up dashboards by 30x using approximate aggregation.

Google Summer of Code (Julia) | Summer 2021 | code, blog post

Implemented, tested, benchmarked, and optimized algorithms to simulate jump processes. Sped up simulations by 70%.

Leadership

Cofounder & Deputy Director of UW Math AI Lab (Lean, LLMs) | 2022-present | Github org, website

- Ran the Lab independently during the Director's sabbatical; designed and taught UW's first Lean course (MATH 480).
- Built the university's Lean community from near zero to several dozen active users; taught Lean to 200+ students and researchers including 3 faculty members.
- Contributed six lemmas to `mathlib` with Lab students. PRs: 19798, 19896, 19886.
- Created the Math AI Seminar and organized 14 talks; delivered invited talks at JMM 2025, a Lean workshop, and Google.
- Secured $\$76,000$ in Lab funding (grants from UW, Google, Nebius, and individual donations).
- Mentored or administrated 58 projects involving 169 undergraduates, 17 graduate mentors, and 10 faculty mentors.
- Built Lab recruiting and community infrastructure: GitHub organization, LinkedIn organization, Lab logo, and Instagram advertising; generated 150 applications in one quarter.
- Supervised the Lab's most popular research projects, including a theorem-search project with 70 applicants.
- Supported student research dissemination: UW Undergraduate Symposium presentations; coauthored submission to Math4AI Workshop at ICML 2025.

Organizer of plasma and ML seminars | 2024-present

- Organized PlaSMoS (plasma simulation and modeling seminar); delivered three talks.
- Helped organize MLJC (machine learning journal club); delivered one talk.

Publications

From Kernels to Attention: A Transformer Framework for Density and Score Estimation | 2025 | pre-print

A symmetry-equivariant KDE-inspired transformer for density and score estimation, significantly outperforming KDE.

Stability of the homogeneous Landau equation in relative entropy and applications to score-based methods | 2025 | pre-print

A proof of relative entropy stability of the Landau equation with applications to accuracy of score-based particle solvers.

Score-Based Deterministic Density Sampling | ICLR, 2025 | paper

Deterministic sampling of an unnormalized density using on-the-fly score estimation with a neural network.

RealEdit: Reddit Edits As a Large-scale Empirical Dataset for Image Transformations | CVPR, 2025 | project page
A dataset and diffusion model to perform text-guided image edits. The first global edit model trained on real data.

Transport Based Particle Methods for the Fokker-Planck-Landau Equation | CMS, 2025 | paper
An algorithm for simulation of plasma using a neural network, inspired by score-based generative modeling.

Community and Mentorship Through the Experimental Lean Lab | JMM, 2025 | abstract
We share lessons we've learned in building community and mentoring undergraduate research projects in Lean.

Extending `JumpProcesses.jl` for Fast Point Process Simulation with Time-Varying Intensities | JuliaCon, 2024 | paper
An algorithm to efficiently simulate any point process on the real line with a continuous intensity rate.

Catalyst: Fast Biochemical Modeling with Julia | PLOS Comp Bio, 2024 | paper
Julia library for modeling and high-performance simulation of chemical reaction networks.

Selected Projects

University Course Matching (Serper, Gemini API, Streamlit) | Summer 2025 | code
Found and parsed 100,000 courses from 2,000 universities, matched courses to a proprietary database of 11,000 textbooks.

AI DnD Bot (LiteLLM, PostgreSQL, Telegram API) | Summer 2024 | code
Made a Telegram bot for Dungeons & Dragons, including short- and long-term memory, image generation, and persistence.