Lucas Mateo Waldburger

Email | GitHub | Google Scholar

SUMMARY

I'm currently building an automated red-teaming system for scalable AI biosecurity. I previously completed PhD research in high-throughput genomics, machine learning for biology, and protein engineering.

EDUCATION

Ph.D., Bioengineering, UC Berkeley and UC San Francisco 2020 - 2026
Advisors: Jay Keasling, Patrick Shih
Selected Coursework: Convex Optimization, ML in Biology & Chemistry, Comp. Immunology, Adv. Immunology

M.S., Electrical Engineering and Computer Science, UC Berkeley 2024 - 2026
Advisors: Jiantao Jiao, Shankar Sastry
Selected Coursework: ML, AI, Deep RL, Agentic AI, ML & Human Behavior, Multi-Agent Systems

B.A., Molecular and Cell Biology, UC Berkeley 2016 - 2018
Honors Degree with Highest Distinction
Selected Coursework: Algorithms, OOP, Comp. Biology, Linear Algebra, Engineering Physics I, II, & III

EXPERIENCE

Research Fellow, Lila Sciences; Machine Alignment, Transparency & Security (MATS) Program 2026 - present
- Built an automated red-teaming system to test model safeguards and elicit dual-use capabilities.
- Implemented evaluation harnesses using inference-time scaling for agricultural pathogen biosecurity assessments.
- Led cross-functional collaborations to develop agentic evaluations on virology and enzymatic DNA synthesis tasks.

Graduate Student Researcher, Berkeley AI Research Lab 2024 - 2026
- Developed a multi-agent LLM framework to simulate human decision-making behavior in dynamic environments.
- Led a team of 3 researchers developing applications of language models for genome assembly.

System Administrator, Lawrence Berkeley National Laboratory 2021 - 2026
- Maintained 2 Arch Linux servers supporting ~10 researchers running genomics tools.

Graduate Student Researcher, Lawrence Berkeley National Laboratory 2021 - 2026
- Modeled gene regulatory network evolution over millions of years using high-throughput assays and ML.
- Built an active learning workflow for protein engineering using language models with calibrated uncertainty, outperforming state-of-the-art methods in biological property prediction and out-of-distribution generalizability.
- Designed a directed evolution platform using deep mutational scanning to maximize genetic transformation.
- Developed an ML-guided screen to mine viral proteomes for gene-regulatory protein discovery and design.
- Built a comparative genomics pipeline to maximize T cell expansion for tuberculosis vaccine development.
- Filed one patent on programmable genetic transformation spanning health and climate applications.

Research Mentorship Program, Google Research 2021
- Developed proficiency in applying ML techniques for DNA and protein annotation using foundation models.

Research Associate, Novome Biotechnologies 2018 - 2020
- Deployed a suite of computational tools to process experimental and genomic datasets supporting cell therapy design.
- Contributed to the first demonstration of programmable microbiome engineering (*Science*).

Student Researcher, UC Berkeley, Dept. of Bioengineering 2016 - 2018
- Modeled and engineered biological feedback control for therapeutic and biomanufacturing applications.
- Contributed to the first demonstration of computer-designed proteins for dynamic control of living cells (*Nature*).

Research Intern, Biohub 2017
- Helped develop a pipeline to process and manage single-cell RNA-seq samples from 12 research labs at Stanford.
- Contributed to one of the first cell atlases, enabling analysis of aging dynamics across the mouse lifespan (*Nature*).

SKILLS

Machine Learning: Fine-tuning and deploying protein and genomic language models; training, optimizing, and evaluating models; uncertainty estimation and calibration; active learning; model interpretability; basic Bayesian optimization
Tooling: Programming Languages (Python, R, Java, JavaScript, MATLAB, C, C++); Data Science (PyTorch, TensorFlow, Keras, XGBoost, NumPy, Pandas, Matplotlib, scikit-learn); Frameworks (Docker, Hugging Face Transformers, DeepSpeed, PyTorch Lightning, Weights & Biases, Inspect AI); Coding Assistants (Claude Code, Codex, OpenCode); LaTeX
Projects: Prediction of operon structures in genomes using an MSA transformer model (2023); Localization of gliomas in MRI scans using a deep Q-network (2022); Autonomous monitoring and cultivation of plants using Raspberry Pis (2022)