Zach Siegel

Full-Stack Engineering | Data Science

zachary.edmund.siegel@gmail.com

zsiegel92.github.io

linkedin.com/in/zach-edmund-siegel

I'm a proficient and enthusiastic application builder passionate about optimization, automation, and statistics-driven decision-making.

Professional Experience

Full-Stack Data/AI Engineer *Current AI | Walleye Capital* *March 2024 - Present*

Drove product end-to-end, from ETL to UI. Deployed bulk ingestion pipelines for millions of records per day; synced data across transactional and analytical storage; designed customer-facing features.

Grounded generative AI in software best practices. Rigorously evaluated and optimized generated content to align to SME preferences with a custom prompt optimization workflow. Built a robust *named entity resolution* service.

Innovated on tooling. Used a custom agentic loop to generate thousands of unique configurations to scrape high-quality structured data (without AI) from different websites each day.

Centered software rigor. Maintained high test coverage for backend and frontend Python and Typescript code, and extended testing paradigms to evaluate non-deterministic AI features.

Won company-wide hackathon. Placed first of 48 teams with a resume-screening platform that aligns to manager preferences via annotations of synthetic data that balances exploration and exploitation and prompt optimization.

Software Engineer & Data Scientist *Capsida Biotherapeutics, Inc.* *January 2022 - March 2024*

Led widely-used internal software platform (AWS+React+Django) providing: customized genomic design tools; long-running bioinformatics pipelines; scheduling automation and visualization; executive BI analyses; automated handoffs across platforms; integrations with lab robots; IP-related data mining. Onboarded cross-functional contributors, maintained testing suite, documented, and socialized tools.

Automated lab data plumbing by connecting platforms including: lab information database (Benchling), project management software (Smartsheet), and networked lab instruments. Built a QR-based inventory update system. Fully automated several lab instrument data ingest operations.

Developed forecasting models for primate immunogenicity, with a decision-support tool that anticipates spontaneous interruptions to animal availability (a widespread challenge in biotech). Built automated analysis pipeline that delivers increasingly precise recommendations to decision-makers as new assay results are recorded in lab information database. C suite relied on forecasts for (expensive!) contracts.

Teaching Assistant for Graduate Students *Anderson School of Management, UCLA* *2020 - 2021*

Taught "Data and Analytics" to students in the MBA programs as part of PhD track teaching requirements.

Created (great) statistics course materials with a JupyterHub/git-based deployment pipeline. Students clicked a "magic link" to access cloud-provisioned, SSO-enabled compute environments.

Freelance Software Development

Automated Scheduling *Sinai Temple* *2018, 2019*
Automated and optimized scheduling Bar/Bar Mitzvah dates for ~130 students via mixed-integer linear programming.

Carpool Assignment Optimization *GroupThere - grouptherenow.com (offline)* *2017 - 2020*
Launched a carpool optimization tool for organizations. Minimizes total drive-time sum across groups of 2-100.

Community Safety Intervention Modeling *LA Community Action Network* *2017*
Re-implemented LAPD's "hotspot" generation algorithm. Compared hotspots to historical arrest/citation/crime report data from City of Los Angeles. Contributed results to community-generated report "Predictive Policing in Los Angeles".

Supply Chain Forecasting, Automation, and Optimization *FactoryOfEverything* *2016 - 2017*
Developed a model for purchasing, production, shipping, and holding over a factory-warehouse-retail system. Forecasting using classical signal processing, regression, and machine learning. Implemented MVP in MATLAB.

Education

University of California, Los Angeles *MS - Operations Research. GPA 3.94* *2019-2021*

Pomona College *BA - Mathematics, Computer Science minor. GPA 3.63* *2010-2014*
Awarded **"Llewellyn Bixby Mathematics Prize"** 2014: to the student with highest achievement in the Mathematics dept.

Research

Pandemic Mitigation Optimization *Anderson School of Management, UCLA* *2021*
Optimizes decisions that affect compartment flow parameters in discrete-time SIRD disease progression model.

Fairness, Efficiency, and Feature-Awareness *Anderson School of Management, UCLA* *2020*
Extends strategies for algorithmic fairness from machine learning community to resource-allocation optimization setting.

Generative Models and Sparse Coding *Department of Mathematics, Pomona College* *2014*
Formalizes connections between *Boltzmann Machine* Distribution and unsupervised learning via sparse coding.

Anomaly Detection Using Dictionary Learning *University of Minnesota, Minneapolis* *2013*
Unsupervised anomaly detection in video data using dictionary learning and sparse coding. An NSF-funded REU. Awarded **"Outstanding Presentation Award"** at Joint Mathematics Meeting, 2014: top 15% undergraduate groups.