Zach Siegel

Software Engineering | AI | Data Science

zachary.edmund.siegel@gmail.com
zsiegel92.github.io
linkedin.com/in/zach-edmund-siegel

Passionate about optimization, automation, and statistics-driven decision-making. Opening doors unlocked by AI. Highly proficient with full-stack software development tools that bring analysis to life.

Professional Experience

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

Drove end-to-end product stories. Deployed dozens of ETL pipelines, creating millions of records per day; managed event-handlers for each record to reach users; contributed customer-facing UI to consume data.

Engineered NLP solutions. Rigorously evaluated and deployed generative AI pipelines. Built a robust *named entity resolution* service used in multiple unstructured-to-structured data pipelines. Combined ranking, chunking, and retrieval strategies to identify and generate content about developing trends in real-time.

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

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

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 calculation 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), networked lab instruments, and internal website database. Built a QR-based inventory update system. Fully automated several lab instrument data ingest operations.

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

Teaching Assistant for MBA Program *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.