

# Ilan Goodman

Data Engineering, ML Systems, and Computer Science Education  
ilangoodman.dev • igoodman@alumni.stanford.edu • (805) 705-8872  
ilangoodman.dev • linkedin.com/in/ilangoodman

## Summary

---

Senior Data & ML Engineer turned Computer Science Lecturer with 8+ years of experience building production data systems at Meta, Robinhood, and the Chan Zuckerberg Initiative. Deep expertise across cloud data warehouses, scalable ETL/ELT pipelines, telemetry, and machine learning infrastructure. Currently designing and teaching large-scale data engineering and algorithm courses at Washington University in St. Louis, bringing rigorous production-system judgment into CS education.

## Technical Skills

---

**Languages:** Python, SQL, Java, C++, C, PySpark, Ruby on Rails, LaTeX

**Data, ML, & Cloud Systems:** Snowflake, Airflow, Spark, Redshift, BigQuery, PostgreSQL, AWS, Matillion, Kafka, Flink, ETL/ELT architecture, telemetry pipelines, data modeling

**AI & Rapid Prototyping:** Expert in leveraging LLMs (Gemini, Codex) for rapid prototyping and AI-assisted full-stack development (including TypeScript, Cloudflare Workers/D1, and GitHub Actions)

**Foundations:** Data structures, algorithms, probability/statistics, machine learning, theoretical computer science

## Professional Experience

---

### Washington University in St. Louis

Aug. 2024 – Present

*Lecturer, Computer Science and Engineering*

*St. Louis, MO*

- Designed and launched CSE 3104 (Data Manipulation and Management), a high-yield introductory data engineering course, and CSE 5114, a graduate course on large-scale and streaming data systems.
- Built comprehensive curricula around Snowflake, Airflow, PySpark, Kafka, Flink, and NoSQL. Developed novel AI-focused assignments, including hybrid streaming anomaly detection with LLMs.
- Co-instruct CSE 247/2407 (Data Structures and Algorithms), modernizing the curriculum for the AI era while preserving rigorous mastery of algorithm design, proofs, and programming abstractions.
- Instruct classes of up to 266 students, manage TA teams of up to 86 staff members, and advise 19 undergraduate students as well as master's thesis and capstone projects.

### Meta Platforms, Inc.

Nov. 2022 – Feb. 2024

*Machine Learning Engineer, Reality Labs XR Tech Ground Truth Data*

*Remote*

- Architected and implemented telemetry for ingestion and delivery pipelines used by XR ground-truth data workflows.
- Improved pipeline automation and built a robust interpolation framework for hand and body annotation data.
- Operated at the boundary between data infrastructure, ML data quality, and annotation systems for multimodal XR datasets.

### Robinhood Markets, Inc.

Apr. 2021 – Oct. 2022

*Senior Data Engineer*

*Remote*

- Built data tooling and established team standards for production analytics and data engineering workflows.
- Developed core business metrics and authored predictive models to assess the firm's liquidity risk.
- Designed reusable frameworks for scalable Spark and Airflow pipelines, significantly improving consistency and reducing deployment times across enterprise data projects.

### Chan Zuckerberg Initiative

Jul. 2016 – Feb. 2021

*Senior Software Engineer, Data*

*Redwood City, CA*

- Built the foundational data infrastructure for the Summit Learning Platform, including the initial Redshift warehouse and the subsequent large-scale migration to Snowflake.
- Designed warehouse architecture, data models, ETL pipelines, and complex data deletion workflows using Matillion and Airflow.
- Led cross-functional projects with up to five engineers and conducted hundreds of technical interviews for software engineering, data engineering, and data science roles.

### Early Career Internships & Research

2012 – 2015

*Facebook AML, Khan Academy, Snapchat, SLAC*

*CA*

- **Facebook:** Built a language model for OCR shape-model parsing on the Applied Machine Learning Solutions team.
- **Khan Academy & Snapchat:** Designed A/B tests, wrote BigQuery utilities, and built Snapchat's first spam-detection

framework.

- **Stanford / SLAC:** Conducted original weak gravitational lensing research via large-scale data simulation; served as (Head) TA for 19 class iterations across 12 quarters.

## Selected Projects & Technical Leadership

---

- **IG Games & Personal Site (AI-Assisted):** Leveraged Gemini and Codex to rapidly build and deploy `ilangoodman.dev` and a serverless games platform (TypeScript, Cloudflare Workers/D1).
- **Baseball Analytics:** Built a public ABS Challenge Leaderboard and MLB analysis tooling for interpreting challenge-value data.
- **Applied AI & Probabilistic Modeling:** Developed diverse projects including a music-composition neural network, BuzzFeed emotional-reaction prediction via NLP, MLB hit prediction, and a novel probabilistic search algorithm for game playing.

## Education & Awards

---

<b>Stanford University</b> – M.S. Computer Science (Theory and Artificial Intelligence); GPA 4.0	2016
<b>Stanford University</b> – B.S. Physics; GPA 3.6	2015
<b>UC Santa Barbara, College of Creative Studies</b> – Physics and Mathematics while in high school; GPA 4.0	2011

*Awards:* Centennial Teaching Assistant Award (Stanford, 2016) • National Merit Scholar • National AP Scholar