

Harrison (Harry) deBoer

San Francisco, CA 94117 • (415) 717-8072 • harrisondeboer@berkeley.edu

EDUCATION

University of California, Berkeley | May 2021

B.A. in Data Science with an emphasis in Economics

Coursework: Data Science, Software Development, A.I., Machine Learning, Statistics, Time Series, Linear Algebra

PROJECTS

Merlin - AI Dungeon Master Chat Bot

Fall 2025

- Built an async FastAPI backend that powers an AI-driven D&D experience, managing characters, sessions, and story progression through structured LLM interactions
- Designed a prompt orchestration and turn service layer to manage dialogue flow, enforce game rules, and generate context-aware LLM responses
- Integrated OpenAI models via a modular adapter supporting JSON-mode outputs, circuit breakers, and token budgeting for stable, cost-controlled generation
- Implemented character creation and chat APIs enabling dynamic storytelling, skill checks, and in-character decision-making across persistent sessions.

Seen Them – Concert Tracking Social Media App

Summer 2025

- Built a full-stack social app for users to track, rate, and share concerts with friends
- Developed modular Go backend with Chi routing, pgx-based Postgres queries, and clean architecture
- Implemented Google OAuth 2.0 login with JWT session management, refresh handling, and cookie-based auth
- Developed a responsive React frontend using React Query for client-side caching and Bootstrap for polished UI

WORK EXPERIENCE

Moveworks - (acquired by ServiceNow)

Mountain View, CA

Data Scientist

January 2025 - present

- Engineered large-scale data processing pipelines in PySpark on AWS EMR, transforming millions of daily log events from S3/Athena into Snowflake models powering internal products and customer analytics
- Designed scalable schemas and data models supporting new product features and behavioral metrics, optimizing for query performance and long-term maintainability
- Enhanced Moveworks' internal pipeline framework, building reusable abstractions (transformers, loaders, and utility modules) that accelerated new pipeline development across teams
- Built and deployed an end-to-end LLM-based classification pipeline, leveraging the company's internal LLM Gateway to measure conversational helpfulness and extract behavioral insights by user role
- Developed a feature-access computation system integrating user protos and org-level configuration from the internal Recon service, mapping millions of users to hundreds of features to enable adoption/engagement analysis
- Collaborated with product engineers and data stakeholders to deliver data APIs that improved product visibility and decision-making across multiple internal platforms

TikTok

Mountain View, CA

Data Scientist

October 2022 - December 2024

- Developed an internal reporting tool using Python and Google Drive API, automating the creation of user research survey reports, reducing manual report generation time by 95% and improving efficiency across multiple teams
- Designed and maintained interactive dashboards that tracked TikTok Shop campaign performance, providing real-time insights that led increased campaign-driven purchases and improved decision-making speed
- Automated text response analysis for user research surveys using a custom TF-IDF model developed in Python, reducing manual analysis time by 90% and enhancing insights from user text with quantifiable metrics
- Developed user survey data pipeline from Kafka logs to Hive tables utilizing PySpark, which was then used as the core data source for user research teams to derive analyses from user research surveys

SKILLS, CERTIFICATIONS, & INTERESTS

Technical Skills: Python, Java, Javascript, Typescript, Go, React, CSS, HTML, NodeJS, SQL, AWS, Spark, Airflow

Certifications: AWS Certified Cloud Practitioner

Interests: Recreational tennis player • NBA fanatic • Avid cook and burger enthusiast • Board Game Connoisseur