Aarash Heydari

http://linkedin.com/in/aarashy

Email : aheyd@berkeley.edu
Mobile : +1-540-282-8104

Statement of Purpose

I am an engineer who takes the time to understand things deeply but prefers to execute quickly. My experience working at early- to mid-stage startups gave me the opportunity to wear many hats and perform functions that stretched beyond the role of an architect. I love writing code and explaining complicated concepts in computing. Above all, I am versatile.

In my next role, I hope to lead business-critical functions and be part of a team that challenges me to deepen my technical abilities and soft skills.

Education

University of California, Berkeley — Berkeley, CA

Bachelor of Arts: Computer Science; GPA: 3.73 — *Aug. 2015 – May 2019*

- **TA / Student Instructor**: Taught four classes over my last five semesters:
  Data Structures | Artificial Intelligence | Algorithms (two semesters) | Machine Learning

Experience

Dataland (Y Combinator, S20) — New York, NY

Founding Engineer — *August 2021 - Present*

Dataland is a private startup that helps internal teams use one UI to work across all their systems of record. As the first hire, I built their sophisticated product from the ground up, gained expertise in high-performance systems programming in Rust, and developed very strong engineering fundamentals.

- **SQL time machine**: Built a SQL CLI application with extended time travel syntax, i.e. "time travel monday 4pm". Followup queries are executed from the perspective of that point in time. This is made possible by replicating a PostgreSQL transaction log over Kafka to construct a history database that includes every version of every row, and utilizing the open-source **PostgreSQL query parser** to rewrite queries to incorporate time travel CTEs.
- **Kubernetes resource controllers**: Developed an infrastructure-as-code deployment experience for our customers to submit table schemas and JavaScript trigger functions via the Dataland CLI. Used Kubernetes operators to reconcile their intended configuration and run their code in sandboxed V8 isolates (Deno).
- **DevOps and infrastructure**: Used kpt and GKE for operational management of backend microservices. Used Cloud DNS to configure static IP masquerading. Used GCP logging and alerting to monitor system health.
- **Open-source**: Utilized and deeply studied a variety of open source projects to understand their architecture, illustrating my ability to quickly acclimate to new projects and frameworks. To call out four in particular: In the C programming language, I have read thousands of lines of **PostgreSQL** and **the Kafka client library**. In Go, I studied and utilized **Ory Kratos** to build a simple and secure Authn/Authz service. In C++, I have contributed bug reports, issues and pull requests to **DuckDB**, an OLAP SQL engine.

Clumio — Santa Clara, CA

Software Engineer / Tech Lead — *July 2019 - July 2021*

Clumio is a private startup that provides cloud data protection at scale. As a member and lead of multiple teams, I architected common platform components, delivered features that enhanced customer UX, and was responsible for the reliability of critical infrastructure.

- **Alerts and progress logs**: As tech lead, spearheaded the overhaul of the internal and external alerting framework. Identified a lack of transparency into the step-by-step progression of backend workflows. Designed and delivered alerting and log solutions that delight customers. Educated fellow developers on how to integrate with this. Mentored a Customer Success engineer, empowering him to build his first microservice which proactively opens Zendesk tickets in response to failures in the system, driving high customer confidence in the platform.

- **Public REST APIs/SDKs**: As tech lead, managed the platform, design, deployment, and versioning of 150+ APIs built on Swagger/OpenAPI and API Gateway. Met weekly with developers and PMs, working across time zones to review APIs and documentation of all engineering teams.
  Wrote a blog about Clumio's API design principles and architecture.
- **Scheduler team**: Scheduled 100k backup jobs for EC2, RDS, and VM disks per day using SNS work queues. Managed the health of key platform services, including a configurable job scheduler and resource manager. Participated in on-call rotation to handle escalations.
- **BI dashboards**: Queried against a 40 million row datalake by integrating Snowflake with embedded Sigma dashboards, powering unprecedented visibility into the history of changes in customer data environments.

Microsoft, Azure Identity Redmond, WA
Software Engineer Intern *May 2018 - Aug. 2018*

- **Feature flag auto-rollout**: Used Azure Functions to automate the configuration management infrastructure for a global LDAP service to minimize human error in deployment, saving ~3 hours of engineer time per enabled feature.

Datalogue Montreal, Quebec
Backend Software Engineer Intern *May 2017 - Aug. 2017*

- **Data pipeline**: Contributed to stateless, purely functional Scala API that accepts petabyte-scale datasets to deliver massive processing power. Built transcoders for JSON, Excel, XML, and MongoDB.

Research Projects

Big Data in Radiology (BDRad) UCSF Radiology and Biomedical Imaging Department
Advised by Dr. Jae Ho Sohn *Sept. 2018 - December 2019*

BDRad is an interdisciplinary research group which brings together data scientists, computer scientists, and radiologists in order to understand how big data computational resources and computer vision advances can provide innovative solutions in imaging sciences.

- **Annotation/visualization system for 3D cross-sectional images**: Integrated a Python plugin for a medical image viewer app to translate between segmentation maps and tabular point-and-radius CSV data. Used this to visualize ML predictions on CT scans and provide radiologists a highly efficient annotation system for producing training data.
- **Presentation at Radiological Society of North America, 2019**: "A DICOM-Embedded Annotation System for 3D Cross-Sectional Imaging Data" (**Poster** | **Source Code**).
- **Peer-reviewed publication**: "High precision localization of pulmonary nodules on chest CT utilizing axial slice number labels" (**Publication**)
- **Weekly reading group**: Participated in weekly discussions of the group's projects and machine learning theory. Presented lectures on topics such as ROC curves and confidence intervals.

Fake News and Misinformation UC Berkeley
Advised by Professor Gireeja Ranade *Feb. 2018 - July 2019*

Professor Ranade assembled a research group to perform data-driven studies on the sociological consequences of social media on news and politics.

- **Data collection**: Used YouTube APIs to collect and analyze data on comments in political media coverage.
- **ML**: Used scikit-learn to apply Random Forest, SVM, etc. to predict the bias category of YouTube videos based on comment statistics.
- **Publication**: "YouTube Chatter: Understanding Online Comments Discourse on Misinformative and Political YouTube Videos", First Author (**Publication** | **Source Code**)

Skills

- **Languages**: Rust, Python, Go, Scala, Java, TypeScript, C, C++
- **Technologies**: AWS, GCP, Kafka, Snowflake, PostgreSQL, Unix, Docker, Kubernetes (deep mastery)
- **Humanities**: Eagle Scout, Farsi, French, Classical Piano, Jazz Drums