

Aviral Srivastava

- Github: [kebab-mai-haddi](#) • LinkedIn: [/in/sriavi](#) • Website: [aviralsrivastava.com](#)
- Email: hi@avisri.com • 617-283-3811

EDUCATION

Boston University, Boston, MA.

Sep 2019 - Dec 2020 (Expected)

- MS, Computer Science. CGPA : 3.77/4

Vellore Institute of Technology, Chennai, India.

Jul 2014 - May 2018

- B.Tech, Computer Science and Engineering. CGPA : 3.56/4

SKILLS

- **Languages:** Python, C++, PHP, Java, SQL
- **Technologies:** Spark, Dask, Kubernetes, Docker, Vagrant, Argo, Airflow, Kafka, Pulsar, Celery, Flink, Snowflake, Cassandra, RocksDB, PrestoDB, KyotoDB, Redis, Django.
- **Platforms:** AWS, GCP, Azure

WORK

EXPERIENCE

Redhat, Data Engineering Intern, Internal Data-Hub team, Boston

May 2020 - Present

- Re-partitioning the schema of [SOS reports](#) to make data accessible.
- Building a data browsing platform to be used by AI and ML teams w/o any hidden technical debts, using **Spark, JupyterHub, Argo, Kubernetes, S3**, and, **Ceph**.
- Built **Argo** workflows for AI and ML engineers to deploy their notebooks directly without any hidden technical debt such as configuration management, Spark OOM issues, etc.
- Reduced S3 object count of all data by more than **90%**, removed the Ceph's limit on object count.

Boston University, Research Associate (Software Engineering), Boston

Sep 2019 - May 2020

- With [Dr Eric](#), I worked on making UML diagrams scalable and reliable for large codebases.
- Got selected in **MIT-ICORPS** Fusion program to commercialize the research.
- Conducted customer discovery interviews with CEOs, CTOs, Engg Managers from various companies to understand the **Customer Segments** and **Value Propositions**.

Atlan, Data Engineer

May 2018 - Jul 2019

- Atlan is a data democratization startup that helps data teams collaborate frictionlessly on data projects.
- Ensured the development and on-premise deployments of dashboards like [United Nations SDGs monitoring](#), [Ujjwala Dashboard](#) using **Python3, Kubernetes, Docker** and **AWS S3**.
- Created serverless JupyterHub: a multi-user hub that spawns, manages, and proxies multiple instances of the single-user Jupyter notebook server. Stack: **papermill, Python3, Docker, Helm, Kubernetes**. Used by companies like **Unilever, Mindshare, McKinsey & Co**.
- Built and deployed Data Repository: a version control for data, including features like 'git diff', rollback, at a scale of over **billion rows in one cycle**. Stack: **Python3, S3, RocksDB, KyotoDB, Hive metastore and PrestoDB**. This scaled the production releases by **300%**: 1 project release to 3 projects release per quarter.
- Built a micro-service with Event Sourcing mechanism for pricing in [Collect](#) using **Kafka, Python3, Docker**.

Wingify, SDE Intern

Dec 2017 - Mar 2018

- Wingify is a B2B software product company with clients like **ebay, Disney, AMD, etc**.
- It's product is VWO which is an experience optimization platform that improves key business metrics by empowering clients to easily discover insights, test ideas, and improve engagement.
- As an intern, I worked with Head of Engg, Ankit Jain, to build a micro-service which parses textual inputs into Domain Specific Language trees and JavaScript code.
- APIs of this microservice scaled query processing of the product by **60%**.
- Stack: **PHP, MySQL, CICD, Docker, Vagrant, Blackbox Testing, Unit testing, JMeter**

Other Internships

- Commutatus, LeagueSX, University of Manchester, University of Liverpool

OPEN SOURCE

- [Profiling RocksDB-Apache Flink](#)
 - Currently working with [Dr Vasia Kalavri](#) (Prime Member Contributor of Apache Flink) and [Dr John Liagouris](#) to determine overhead caused by state-ful operations in Flink streaming engine.
- [ZCash Service Status Dashboard](#)
 - Built a health-check service status dashboard to monitor the core ZCash blockchain, crypto exchanges, blockchain explorers and ZCash communities and forums.
- [Other projects](#)