

# Aviral Srivastava

• Github: [kebab-mai-haddi](#) • LinkedIn: [/in/sriavi](#) • Website: [aviralsrivastava.com](#)  
• Email: [hi@avisri.com](mailto:hi@avisri.com) • [\[REDACTED\]](#) • Blog

## EDUCATION

**Boston University**, Boston, MA.

Sep 2019 - Dec 2020 (Expected)

- MS, Computer Science. CGPA : 3.77/4
  - **Subjects:** Designing Data Warehouse, Database Security, Advanced Database Management Systems, Graduate Distributed Systems (MIT), ML, AI, Adv. Health Info and Big Data, Advanced Analysis of Algorithms, Information Systems Analysis and Design.

**Vellore Institute of Technology**, Chennai, India.

Jul 2014 - May 2018

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

## SKILLS

- **Languages:** Python, C++, Go, Java, PHP, Rust, SQL, MySQL, PostgreSQL
- **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

**Twitter**, Software Engineer

Apr 2021 - Present

- As part of the Graph Storage team, I am working on real-time storage infrastructure.
- Led the team for building a managed graph database offering for analytics purposes, graph machine learning, and applications. Reduced onboarding time from 1 week to minutes, for teams at Twitter.

**Redhat**, Software Engineering (Data) Intern, Internal Data-Hub team, Boston

May 2020 - 2020-12-17

- Re-partitioned the schema of [SOS reports](#) (**10 Trillion+ rows, 100+ TeraBytes** data) to ensure high data availability and low latency. **Reduced** the count of S3 objects by more than **90%**: from 15M+ objects to < 1M objects.
- Scaling the visualization dashboards for internal analytical teams to prevent outages and provide smooth user journey for our data team members.
- Standardized data workflows for the internal teams using **Argo, Spark, JupyterHub, Kubernetes**

**Boston University**, Research Associate (Software Engineering), Boston

Sep 2019 - May 2020

- Got funded by MIT-ICORPS for commercializing the research for improving software quality.

**Atlan**, Software Engineer(Data)

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**.
- Integrated 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 system for tabular 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 payments in [Collect](#) using **Kafka**.

**Centify**, Co-founder, CTO

Mar 2018 - Jul 2018

- Audience Engagement tools which enable digital media platforms to integrate social media reactions with video or textual content.
- Built the complete system for processing social media content and applying emotional as well as sentimental analysis. Got featured on [IBM's developer blog](#)

**Wingify**, SDE Intern

Dec 2017 - Mar 2018

- Wingify, with clients like **ebay, Disney, AMD, etc.**, has a product VWO which is an experience optimization platform.
- Reduced the query processing time for the product by **60%** by building a micro-service that parses textual inputs leveraging the Database.
- Stack: **PHP, MySQL, CICD, Docker, Vagrant, Blackbox Testing, Unit testing, JMeter**

## OPEN SOURCE

- [Profiling RocksDB-Apache Flink](#)
- [ZCash Service Status Dashboard](#)
- [Other projects](#)