

Salayhin Md Sirajus

Tokyo, Japan

+81-080-6918-4753

linkedin.com/in/salayhin | github.com/salayhin

https://salayhin.github.io/me/ | salayhin.lab@gmail.com

SUMMARY

Tech Lead / Staff Data Engineer with 14+ years of experience owning and leading **large-scale, cloud-native data platforms**, including **near-real-time telemetry and analytics pipelines**. Acted as the **primary technical decision-maker** for data architecture, driving design reviews, tooling choices, and scalable solutions that balance **performance, cost, and developer productivity**.

Specialized in **lakehouse architectures, dbt-driven analytics, data governance, and FinOps** on GCP with AWS expertise, enabling **high-volume event and telemetry data processing** for AI/ML workloads while delivering major performance gains and cost reductions.

SKILLS

Languages & Scripting: Python, Scala, SQL, Bash

Cloud Platforms: GCP (BigQuery, Cloud Storage, Dataflow, Dataproc, Pub/Sub, Composer), AWS (transferable expertise): S3, Athena, Redshift, EMR

Table Formats: Apache Iceberg

Orchestration & ELT: dbt Core, Apache Airflow

Streaming & Batch: Pub/Sub, Dataflow, Dataproc, Apache Spark

Architecture & Modeling: Data Lakehouse, Data Warehouse, Dimensional Modeling

FinOps & Monitoring: GCP audit logs, Billing Export, Asset Inventory, Cloud Monitoring

DevOps & CI/CD: Docker, Kubernetes, Terraform, GitHub Actions

BI & Analytics: Tableau, Redash, Looker

Governance & Security: GDPR Compliance, Data Lineage, Automated Testing

EXPERIENCE

Staff Data Engineer (Permanent)

November 2023 - Present

[Zeals Co., Ltd](https://zeals.co.jp) | Tokyo, Japan

- Acted as primary technical decision-maker for the data platform, evaluating architectural trade-offs, tooling choices, and scalability constraints to balance performance, cost, and developer productivity.
- Owned the technical vision and led the design and implementation of a modern data lakehouse and analytics platform, aligning data architecture with business and ML requirements.
- Led architecture and design reviews for data pipelines and analytics workloads, ensuring consistency with platform standards, reliability requirements, and long-term scalability.
- Led organization-wide adoption of dbt Core and dbt Cloud across multiple teams, defining standards and CI/CD workflows; improved data quality by ~30% and reduced pipeline maintenance effort by ~40% through automated testing and documentation.
- Drove platform-wide BigQuery performance and cost optimization using system tables and billing exports, reducing query runtime by up to 90%, scan volume 80%, and monthly costs by ~40%.

- Led the design of the gold layer using dimensional modeling (star schema), transforming raw events into optimized dim_ and fact_ tables, improving query performance by 70% and reducing scan volume by 80%.
- Owned and drove the platform-wide data governance strategy, defining data contracts, quality standards, and SLAs across teams, resulting in a 25% improvement in data integrity and accuracy.
- Owned the design and implementation of a GCP-wide utilization and cost governance framework using Asset Inventory, introducing automated alerts and rightsizing policies to improve resource efficiency and cost visibility across the platform.
- Mentored junior engineers, driving a 20% uplift in technical proficiency and fostering a culture of continuous learning.

Senior Big Data Engineer (Permanent)
[Zeals Co., Ltd](#) | Tokyo, Japan

November 2022 - October 2023

- Implemented data governance using Google Dataplex, improving data management efficiency by 20% and ensuring compliance with internal and external standards.
- Deployed CI/CD pipelines with Terraform, automating data workflows, reducing manual intervention by 30%, and increasing infrastructure reliability.
- Designed and rolled out a FinOps pipeline using GCP billing exports, reducing infrastructure costs by 15% through proactive cost monitoring.
- Built a semantic analytics layer for a customer-facing dashboard, improving data access speed by 20% and enhancing overall user experience.
- Integrated BI tools (e.g., Tableau) with the Data Lakehouse Gold layer, enabling real-time insights and improving decision-making for business stakeholders by 25%.

Senior Big Data Engineer (Remote)
[Zeals Co., Ltd](#) | Tokyo, Japan

May 2021 - October 2022

- Developed a robust data lakehouse architecture on GCP using BigQuery and Cloud Storage, improving data scalability and accessibility for internal stakeholders.
- Partnered with R&D and analytics teams to support ML model development and BI dashboarding, driving a 20% improvement in data-driven decision-making.
- Audited and remediated data storage systems in collaboration with the security team, strengthening data security protocols.

Senior Data Engineer Level 2
[Pathao Ltd](#) | Dhaka, Bangladesh

June 2018 - May 2021

- Designed and operated large-scale near-real-time telemetry and event data pipelines (trip lifecycle, location updates, driver/rider actions), supporting time-series analytics, operational monitoring, and performance optimization across city-scale mobility operations.

- Established a scalable analytics ecosystem using Google Dataproc, Google Dataflow, BigQuery, and Data Studio, increasing analytics capabilities by 40%.
- Spearheaded the development of a data lake and warehouse in Google Cloud Storage and BigQuery, reducing data retrieval times by 50%.
- Enhanced the data pipeline framework to handle over 1,000 attributes, boosting data processing efficiency by 35%.
- Designed and implemented a scheduled reporting framework on the data warehouse, generating 10+ financial and operational reports, improving team productivity by 70%.

Senior Data Engineer

June 2015 - March 2018

[Augmedix Bangladesh Ltd.](#) | Dhaka, Bangladesh

- Designed and built a data warehouse using SQL Server 2016 as a core contributor, improving business decision-making speed by 25%.
- Integrated a data pipeline framework to ingest data from 3rd-party APIs, Google Sheets, and MySQL, processing 100+ attributes and improving data efficiency by 30%.
- Developed a reporting framework for the data warehouse, generating 20+ operational reports and increasing stakeholder access to decision-ready data by 35%.
- Built backend services to serve millions of records from the warehouse, reducing data retrieval latency by 20%.

Software Engineer

June 2013 - May 2015

[Nascenia Ltd.](#) | Dhaka, Bangladesh

- Deployed a B2B marketplace platform for farmers using Ruby on Rails and MySQL, increasing seller participation by 30% and boosting transactions by 25%.
- Built a web-based tool for social media analytics and post scheduling with Ruby on Rails and PostgreSQL, improving engagement by 40%.
- Enhanced a school management system by extending core features using Ruby on Rails and PostgreSQL, improving system efficiency and reducing administrative workload by 35%.

Software Engineer

November 2011 - May 2013

[Right Brain Solution Ltd.](#) | Dhaka, Bangladesh

- Developed and deployed an API for social media mobile apps in collaboration with the dev team, increasing user engagement by 20%.
- Co-authored a high-traffic automotive portal for the Toronto Star, boosting monthly user visits by 35% and improving customer engagement.

- Built a multilingual B2C ski rental shop, increasing rentals by 20% and expanding the international customer base by 15%.
- Optimized an OMR reader for university online exams, reducing manual data entry time by 95% through improved input handling.

Assistant Programmer

May 2011 - October 2011

[Nanosoft](#) | Dhaka, Bangladesh

- Simplified a web-based appointment scheduling system using PHP and MySQL, improving user experience and reducing booking time by 20%.

Software Developer

June 2010 - April 2011

[Sitebloom](#) | Dhaka, Bangladesh

- Developed a web-based school management system using CodeIgniter and MySQL, improving administrative efficiency and reducing manual workload by 30%.

EDUCATION

Jagannath University | Dhaka, Bangladesh | B.Sc in Physics
IICT, BUET | Dhaka, Bangladesh | Post Graduate Diploma in ICT

2008
2019

CERTIFICATION

- [Data Engineering with Google Cloud Specialization](#) (Google Cloud Training) (Coursera 2023)
- [Data Engineering, Big Data, and Machine Learning on GCP Specialization](#) (Google Cloud Training) (Coursera 2023)