

# Saswata Dutta

saswat.dutta@gmail.com | 8178349520 | Bengaluru | [linkedin.com/in/saswata-dutta](https://www.linkedin.com/in/saswata-dutta)

## EXPERIENCE

### DOMINO DATA LAB | TECHNICAL LEAD

Bengaluru | Nov 2019 – Present

- Working in Common Services, Frameworks and Enterprise team. ( Scala, Play, Python )
- Implemented public REST apis for model monitoring.
- Refactoring and upgrading persistence layers to support latest MongoDB.

### RIVIGO SERVICES PVT. LTD. | PRINCIPAL ENGINEER

Gurugram | Jun 2018 – Nov 2019

- Created platform for ingesting/querying relevant data in/out of our Data Lake. Added a Feature Store to version and share data for ML models.  
( Scala, Spark, Airflow, Python, Pandas, PySpark, AWS : Lambda, S3, Glue, Athena, EMR )
- Refactored "Payments/Commerce" services into separate micro-services corresponding to Payments (pay-in, pay-out), Wallet, Incentives, Fuel, and Fastag.  
( Java, Scala, Spring, Hibernate, MySql )
- Implemented realtime pay-out to suppliers via IMPS/NEFT, with automated retries and refunds, and best effort payment mode selection based on banks' recent success rates.
- Realtime Wallet and Transactions for demand and supply users
- Integration with all major vendors : IOCL, HPCL, Reliance, BPCL, to generate fuel vouchers.
- Rule based dynamic Incentive engine to distribute coupons to users in realtime.
- KYC service by performing penny drop on beneficiary account to verify name and account details.
- Implemented framework for integration and stress testing of entire payment/commerce ecosystem.
- Implemented queryable application log repository in AWS Glue/Athena, for ELK cost savings.

### ADOBE INDIA PVT. LTD. | COMPUTER SCIENTIST

Noida/Bengaluru | 2013 – 2018

- Feature development and performance optimisation in multiple releases of AEM Forms.  
( Scala, C++, Java, JavaScript, Python, Apache Sling/OSGI )
- Automatic Paper Form Capture: Lead project from inception. Created pipelines for gathering datasets, defined protocol for communication between different knowledge sources, heuristically extract structure and semantics from "flat documents", created framework for measuring KPIs, implemented the merge and presentations engine which generated the final "Adaptive Form".
- HTML5 and Adaptive forms: Optimised browser footprint to allow large forms to be efficiently loaded. Built a framework for auto importing xfa-pdf controls as html widgets. Worked on layout and formatting of elements in both html and pdf rendition. Worked on backend management, formatting and validation of user data, and generation of xml/json output for persistent storage.

## EDUCATION

### IIT GUWAHATI | M.TECH. COMPUTER SCIENCE AND ENGINEERING

Guwahati | 2011 - 2013

GPA: 8.92

## PROJECTS

### GENERIC DATA INGESTION FRAMEWORK | SCALA, SPARK

[github.com/saswata-dutta/spark-seed](https://github.com/saswata-dutta/spark-seed)

Ingests data from source, transform and flush to sink, based on configs. Motivation was to allow easy configuration of simple ingestion and ETL. Supports AWS S3, Jdbc enabled SQL db, MongoDB.

### SCHEMA AWARE FEATURE STORE | PYTHON, PYSPARK

[github.com/saswata-dutta/feature-store](https://github.com/saswata-dutta/feature-store)

A simple Feature-Store which allows schema evolution. Uses S3 for storage, Glue for data catalog, and Athena as SQL engine. Motivation was the lack of cost effective tooling for small to medium scale data processing.

### EXPRESSION EVALUATOR | SCALA

[github.com/saswata-dutta/eval-exp](https://github.com/saswata-dutta/eval-exp)

A generic expression parser and evaluator with support for variables and allowing extensions to operations. Input can be expressed in JSON or Lisp.