

Prashant Shukla

Sr. Data Engineer

+44 78763-42409

prashant_978737@yahoo.co.in

SKILLS

Programming Language: Python, Java, JavaScript (JS), Scala

Database and Object Storage: Oracle, MySQL, MongoDB, DynamoDB, Presto, AWS-S3

Data Engineering (Techniques and Libraries): Data Warehouse and Data Lake management, Extract-Transform-Load (ETL), Data Modeling, Data Transformation, Schema evolution, SCD1 from Dimension Tables, Data Catalog, Query Optimisation, PySpark, Numpy, Pandas

Backend Development Framework: Restful, Soap, Springboot, Flask

Cloud Infrastructure- AWS: EC2, Lambda, Elastic Load Balancer, Scaling Groups, VPC, NAT Gateway, Amazon Managed Workflows for Apache Airflow MWAA, EMR (Elastic MapReduce), Streaming- Kinesis, Kafka

Containerisation: Kubernetes, Dockerw

Business-intelligence (BI) Tools: Looker, BigQuery

Others: Kibana, ElasticSearch, DataDog, Agile, Scrum, JIRA, Confluence, Git/GitHub

Domains: Banking, Financial Services and Insurance, Travel, Infrastructure, Communication and Media Streaming

CERTIFICATIONS

- AWS Certified Cloud Practitioner- Aug, 2021
- TestDome (Public Test)- Python (Top 10%), Git, SQL

EXPERIENCE

Twilio Inc., London UK - SrData Engineer (January 2022 - Present)

Project- Data Analytics and Data Platform

- **Improvado Decommissioning:** Oversaw end-to-end decommissioning of Improvado, a third-party tool extracting data from **GoogleAds, Meta Ads, LinkedIn Ads, Bing Ads, and Twitter Ads**. Implemented a custom data pipeline using Singer SDK, Meltano, Python- Pandas, and Airflow, resulting in a cost-saving of \$100,000.

Key Contributions & Achievements:

- Formulated a comprehensive design strategy for the custom data pipeline.
- Led the hands-on development using Singer SDK, Meltano, Python, and Airflow, AWS- S3, EC2, Hive-presto.
- Ensured seamless integration with diverse ad platforms, optimizing data extraction efficiency considering Variety, Velocity and Volume of data.
- Engineered the solution to accommodate future ad analytics integrations, enhancing overall system scalability.
- This tailored solution not only eliminated the substantial cost and maintenance overhead but also set the foundation for a flexible,

scalable, and efficient data extraction, transformation and load process.

- **Reverse ETL** addresses the need for immediate access to analytics-derived insights within operational tools, fostering real-time decision-making and collaboration across platforms.

Key Contributions & Achievements:

- Engineered robust data pipelines for efficient data extraction and transformation from the Data Lake using Python, Scala, PySpark, Pandas, S3, MWA, EMR (Elastic MapReduce) pipelines for deriving SCD1s.
- Integrations: Airtable, Salesforce Integration, ZenDesk
- **Business Data Upload:** Efficiently integrated data from Airtable, Google Sheets, and CSVs to a centralized Data Lake, streamlining processes and optimizing data accessibility.
- **Data Platforms Kubernetes Migration (OTK):** Led the onboarding of Data Platforms services to a Kubernetes Cluster, overseeing the seamless containerization of applications and services. Elevating systems scalability and resiliency.
- **[Hackathon] Slackbot-SQL-Assistant-AI** Significantly streamlined the SQL query generation process for business users, enhancing efficiency and user experience within the Slack platform using Generative AI.

CapitalOne, Nottingham UK - Sr. Associate Engineer (DATA)

February 2021 - December 2021

Project- Data Analytics and Engineering

- **ETLs (Internal and External Sources) Development: Transaction Data, Customer Informations and Risk & Fraud Data-** to analyze spending patterns and customer behavior, personalized insights and Gather data related to fraud detection and risk assessments. **Marketing and Campaign, Credit Bureau Data-** credit scoring and risk assessment, Explore social media and web sources for sentiment analysis and customer feedback. 3rd Party aggregators data integrations.

Key Contributions & Achievements:

- Led the seamless integration of internal transaction and customer data using Apache Kafka, enhancing real-time processing capabilities.
- Successfully migrated on-prem data storage to Amazon S3, optimizing scalability and reducing significant on-prem infrastructure costs.
- Enhanced data quality and integrity by implementing automated testing using Great Expectations.
- Ensured compliance and security by mitigating risks and strengthening data governance.

Dell EMC Technologies, Bangalore, India - Principal Engineer

July 2019 - January 2021

Project- POWERONE (Next Generation Converged Infrastructure) - DELL EMC

- Led system and component validation for the Next Generation Converged Infrastructure, ensuring reliability. Orchestrated installations, validated Ansible playbooks, tested REST APIs, and automated UI processes, contributing to robust design and decision-making processes.

Amadeus Labs, Bangalore, India - Senior Software Engineer

August 2016 - July 2019

Project- Altea Inventory (Seating)

- Engineered a Python-based back-end testing framework, automating test creation, translation, and dynamic test data management for evolving SEATING business rules. Conducted weekly regression tests using CI/CD Jenkins Pipelines, maintained test environments, and implemented a Discrepancy Framework using Flask, Kafka, and Python for real-time detection and resolution of seating discrepancies, delivering reports for 30+ airlines.

Wells Fargo, Bangalore, India - Senior Software Engineer

July 2014 - August 2016

Project- Early resolution loss mitigation Facility- Web Apps

- Revamped a VB-based QTP test framework, overseeing test data, and contributing to the complete Software Testing Life Cycle (STLC) for the project. Innovated the "Login Test Role" tool, reducing manual efforts for testers and BAs by addressing login issues and backend credential expiry.

Accenture Services, Bangalore, India - Senior Software Engineer

June 2011 - July 2014

Project- Bank Of America and Aflac Inc. (USA)

- Worked on ETL Development, Reporting Dashboard Development and UI validation. Enhanced and maintained a QTP framework, managed release test estimations, and developed automated tools for environment health checks and real-time signature simulation.

EDUCATION

Amity University, Lucknow, India - Bachelor of Technology

July 2007 - June 2011

References: On request.