

Prasad Dattaram Deshmukh

+44 7436349955 | prasad21121996@gmail.com | <https://www.linkedin.com/in/prasad-deshmukh-586690130/>

PROFESSIONAL SUMMARY

Results-driven AWS Data Engineer with 6+ years of experience in **data engineering, cloud computing, and big data technologies**. Proficient in **AWS (Glue, Lambda, S3, EventBridge, DMS), Snowflake, SQL, and Python**, with expertise in designing and implementing **scalable data pipelines, ETL workflows, and real-time data processing solutions**. Skilled in **modernizing data architectures**, optimizing performance, and ensuring high availability in **cloud-based data warehousing solutions**. Strong background in **REST APIs, microservices, and CI/CD best practices** for data-driven applications. Proven ability to **collaborate with cross-functional teams**, lead **agile development**, and deliver high-quality, cost-efficient solutions aligned with business objectives.

PROFESSIONAL EXPERIENCE

Mastek UK LTD

Software specialist I

Leeds, UK

Nov 2022 - Present

- Migrated GDPR data processing from Parquet to Delta Tables in S3, upgrading ingestion pipelines to support delta format and automating data depersonalization via AWS Glue.
- Designed and implemented a metadata-driven ingestion framework that automates external table creation in S3, base views, and dynamic masking rules for PII columns.
- Enhanced API performance by 75%, optimizing queries, caching, and indexing, leading to a \$19K annual cost saving.
- Built a Serverless Data Ingestion Framework using AWS Glue, RDS to extract data from SQL Server, FTP, DynamoDB, and S3, transforming data into Parquet format for efficient storage and analytics in Datalake.
- Reduced ETL execution time by 60% through parallel processing, Snowflake materialized views, query optimizations, and resource tuning.
- Developed a Streamlit-based web application for Excel/CSV uploads, allowing users to apply custom transformations and insert data into Snowflake, with automated deployment.
- Integrated S3 with Snowflake by exposing Parquet and Delta tables as external tables, improving query performance and reducing data latency using materialized views.
- Optimized S3 storage costs by 30% through object lifecycle policies, intelligent tiering, and storage class selection.
- Implemented a cost-effective Snowpipe solution using S3, EventBridge, and SNS, ensuring real-time data ingestion and pipeline automation.
- Developed a Data Availability Trigger Framework leveraging AWS EventBridge, Step Functions, Lambda, and Snowflake stored procedures, integrating SES/SNS for real-time notifications.
- Implemented a Test Data Generator (TDG) framework for anonymizing sensitive data, enabling Data Science teams to train models on production-scale datasets without privacy risks.
- Led a team of 5+ engineers, mentoring in best practices, performance tuning, and CI/CD automation using Azure DevOps.

Mastek Ltd

Senior Software Engineer

Mumbai, Maharashtra, India

Dec 2018 - Nov 2022

- Optimized SQL queries and ETL pipelines, reducing data processing time by 45% and improving query performance by 50%.
- Automated data validation and transformation, increasing data accuracy by 98% while streamlining ingestion workflows.
- Implemented monitoring and root cause analysis dashboards, reducing production incidents by 40% and improving troubleshooting efficiency.
- Developed AWS Quicksight reports and CloudFormation templates, enabling cross-environment report automation.
- Collaborated with Data Science teams to optimize data pipelines, improving ML model training efficiency and feature engineering processes.

EDUCATION

Mumbai University

Bachelor's, Computer Engineering

August 2015 - March 2018

Maharashtra State Board

Diploma In Information Technology

June 2012 - May 2015

CERTIFICATIONS

AWS Certified Developer – Associate | Amazon Web Services (AWS) |

SKILLS

Programming Languages: Python, SQL, Java, JSON

Database: RDBMS, MS SQL, MySQL, Postgres, SQLite, NoSQL, MongoDB, Redis

BI Tools: Quicksight, Pyramid Analytics

Hadoop-Ecosystem: HDFS, Spark, Hive, Pyspark

ETL Tool: Pentaho Data Integration

Web Technologies: HTML/CSS, REST APIs

Cloud Platforms: Amazon Web Services (AWS), Snowflake

AWS Services: Glue, Lambda, EMR, Athena, Cloud Formation, API gateway, DynamoDB, Step Function, Event bridge, S3, DMS etc.

Version Control: Git (GitHub), Azure DevOps (Azure Repos), Docker