MOHAK CHUGH

# Software Engineer at Amazon

+91 9810178257  [me.mohakchugh@gmail.com  [https://www.linkedin.com/in/mohakchugh  Delhi NCR](https://www.linkedin.com/in/mohakchugh)](mailto:me.mohakchugh@gmail.com)

# Work Experience

## Amazon Gurugram, Haryana, India · Hybrid

Software Development Engineer 2 – **[Level 5]** 10/2024 – Present

* Worked on Developing APIs to vend out Amazon Financial data for Transactional and Non-Transactional collections use-cases with **100% reconciliation** coverage.

**Used:** AWS EMR (spark), AWS Big Data, APIGateway, AWS Lambda, DynamoDB, Scala, Hadoop, Glue, Athena

Software Development Engineer 1 – **[Level 4]** 12/2022 - 09/2024

* Developed **Spark** based **Big Data reconciliation engine with multiple data source support reconciling 1 billion+ records daily**. Provides capabilities to reconcile between SQL and No-SQL data sources with custom hooks for integration with data pipelines. This solution reconciles **100% of internal financial data** flowing through a central data lake.
* Led the creation of an award-winning **RAG-based LLM QnA bot, engaging 300+ users** and processing 5K questions daily; leveraged advanced prompting techniques and guardrails, resulting in **93% accuracy** from rich text documents

**Used:** LLM ClaudeV2.1, AWS cloud, Streamlit, DynamoDB, PostgreSQL (RDS), ECS Fargate, Airflow, Agentic AI

Data Engineer 1 – **[Level 4]** 09/2021 - 12/2022

* Engineered AI-driven headcount management tools by leveraging **100+ work driver datasets**, enabling precise ML models for scalable headcount predictions; boosted operational efficiency by forecasting resource allocation with **95% accuracy**. Integrated functionalities like email notification systems, auth management and ML data pipeline orchestration using **Airflow**.
* Launched e-invoicing compliance for Amazon by enabling VAT IDs encryption for Brexit, AU, CA regions, impacting **3,000+ global vendors.**
* Collaborated with senior data engineering teams on a re-architecture of our financial analytical mesh data-lake with database consolidation, automating schema-level data sharing across Redshift clusters. Designed and developed to **ensuring SLAs of 1 sec is met**. Authored with **90% test coverage**. Deployed efficient, secure solution, enhancing org productivity with security reviews for **critical data storage**.
* Mentored 3 interns and got 2 converted to full time engineers.
  + **Used:** Airflow, Redshift, AWS EMR (spark), Scala, Hadoop, Glue, Athena, AWS Serverless, CI-CD pipelines, AWS Big Data,

# Education

Maharaja Surajmal Institute of Technology

* Bachelor of Technology - Computer Science | 8.7 | [University Rank 3 in Final year](https://www.linkedin.com/feed/update/urn:li:activity:7055093560970256384/) 08/2017 - 12/2021

# Publications

## [Benchmarking geospatial database on Kubernetes cluster](https://link.springer.com/article/10.1186/s13634-021-00754-2)

[Analysis and Optimizations of PostgreSQL/PostGIS in Kubernetes for](https://link.springer.com/article/10.1186/s13634-021-00754-2)  [improved geospatial query performance.](https://link.springer.com/article/10.1186/s13634-021-00754-2)

## [Generative AI Q&A chat assistant using Amazon Bedrock](https://aws.amazon.com/blogs/machine-learning/how-amazon-finance-automation-built-a-generative-ai-qa-chat-assistant-using-amazon-bedrock/)

[Built a custom RAG based QnA system for a niche use case which gave accuracy of 86% as compared to 49% from of the shelf tools](https://aws.amazon.com/blogs/machine-learning/how-amazon-finance-automation-built-a-generative-ai-qa-chat-assistant-using-amazon-bedrock/)

# Languages and Skills

Work Domains: AWS Cloud (**Development** and **Deployment** for **Serverless** and Server based solutions),

Applied Artificial Intelligence - **LLM** **RAG** Architecture and Natural Language Processing (**NLP**),

**Big Data** (Analytical and Transactional) and Advanced System Design on AWS

Languages: Python, Java, NodeJS, Angular, **Advanced SQL**

Tools and Technologies: API Development, Distributed Systems (**Spark**) and Data pipeline management via **Airflow**, Data Structures and Algorithms

# Projects

[DroneDash](https://mohakchugh.github.io/website/#/dronedash)

## <https://mohakchugh>.github.io/website/#/dronedash

* [An Online Web Interface for surveillance in large scale construction sites. Supports up to 8 livestreams at once. Currently being used by Aditya Birla Group.](https://mohakchugh.github.io/website/%23/dronedash)

# [ContentCollab](https://mohakchugh.github.io/website/#/blogger)

## <https://mohakchugh>.github.io/website/#/blogger

* [A real-time online document editor that allows a group of users to collaborate on webpage contents, see and make changes in real-time.](https://mohakchugh.github.io/website/%23/blogger)

# Achievements

[1st Runner’s up in Digital India Makeathon](https://drive.google.com/file/d/1YEMPcosx8COkOxqFn3vXHiugsIIhkZMX/view)

* Certificate Link: https://drive.google.com/file/d/1YEMPcosx8COkOxqFn3vXHiugsIIhkZMX/view

# [Soft Skills Certification from AICTE](https://drive.google.com/file/d/1aBjeD3gCGHphhdfNAyZrXADcOZyBjLQ-/view)

* Certificate Link: https://drive.google.com/file/d/1aBjeD3gCGHphhdfNAyZrXADcOZyBjLQ-/view

# [University Rank 3 in BTech Final year](https://www.linkedin.com/feed/update/urn:li:activity:7055093560970256384/)

## Certificate Link: https://www.linkedin.com/feed/update/urn:li:activity:7055093560970256384/