**KRUTHIN B** [**kruthinneha@gmail.com (+91) 8867040196**](mailto:kruthinneha@gmail.com%20%20%20(+91)%208867040196)

**Data, Microservices and Generative AI Engineer Master of Business Administration| Ramaiah Institute of Technology | 2019**

**CAREER OBJECTIVE:**

A Data and AI engineer with practical experience in GCP (Google Cloud Platform), AWS (Amazon Web Services) and Azure hyperscale’s, eagerly looking for opportunities to apply my expertise in **GCP** services and Generative AI technologies to develop innovative AI-driven solutions and enhance cloud-based applications. Actively seeking a challenging role where I can leverage my technical skills and experience in **Amazon Web Services** to contribute to impactful, results-oriented projects and drive significant advancements in artificial intelligence, Generative AI, and data engineering.

**PROFESSIONAL EXPERIENCE:**

**Data & AI Engineer| Capgemini | MAY 2022 – Present**

**Data Engineering:**

• Formed **federated data lakes by ingesting** data from various source types – RDBMS, Cloud File Systems, Network File Systems and

Big Data systems.

• Developed frameworks for Data Profiling, Validation, and Reconciliation using serverless approach utilizing **Google Cloud Functions**

and **Google Cloud Data Fusion**.

• Created **microservices** frameworks for building seamless integrations between cloud services through SDKs for various phases of data engineering like **ingestion, data processing, data governance etc.** using templatized **Google Data Proc** scripts for a variety of data source types.

• Managed metadata using **Firestore** and **Cloud Spanner** to build audit logs for the accelerator framework.

• Integrated **Data Proc** capabilities with **Data Fusion** to perform data transformations through a serverless approach, via Notebooks.

• Built data migration systems from both **homogeneous and heterogeneous databases** to cloud data warehouses using **AWS Database**

**Migration Service (DMS)** and **Schema Conversion Tool (SCT**) driven by Flask APIs**.**

• Developed frameworks for **Data Profiling, Validation** and **Reconciliation** using both serverless approach and managed approaches. Serverless Approach utilizing AWS Glue and Azure ADF and Managed approach through Spark.

• Integrated with Snowflake Tasks through Airflow orchestration framework for creation of task specific Databrick’s Notebooks.

• Managed metadata using **MongoDB** and **Firestore** to build audit logs for the accelerator framework.

• Integrated **HDInsight** capabilities with **Azure Data Factory** to perform **data transformations** through a serverless approach, via

Notebooks.

• Orchestrated **ADF, Glue** jobs using Airflow to run on scheduled or ad-hoc basis in **sequential and parallel executions.**

• Familiarity with **Google Pub Sub** and **Google Cloud Composer**.

**Generative AI:**

• Implemented a solution for querying databases **by converting natural language** using **prompt to SQL queries**, including output summarization and dynamic chart visualization. Utilized **Prompt engineering techniques** and **a RAG-based** for vector databases using **Google Gemini Model**.

• Created a **self-healing data reconciliation framework** to automatically correct degraded records following templatized, configurable, spark-based **data quality checks**.

• Developed an **AI-based conversational interface(chatbot)** using a **robust knowledge base** and **AWS Bedrock** service.

• Developed Generative AI-powered**, RAG-based schema and DDL conversions** for transforming schemas from various source databases to target databases, utilizing **multi-modal agents**.

**SKILL SET:**

**Programming Languages** : Python, SQL

**Python Packages** : Pyspark, Pandas, NumPy, Boto3, Matplotlib, Flask, PyMongo, Streamlit, PyODBC

**Relational Databases** : Teradata, SQL-Server, Postgres, Oracle, MySQL

**No-SQL Databases** : MongoDB, Document DB, Cosmos DB

**Cloud Services**:

• **GCP** : **Data Proc, GCS, Data Fusion, Big Query, Firestore, Pub Sub, Cloud Composer**

• **AWS** : DMS, Glue, Bedrock, Redshift, S3, IAM, EMR, Lambda

• **Azure** : Databricks, HDInsight, ADF, Synapse Analytics, ADLS

• **Snowflake** : Snowflake DWH, Snowpipe

**Big Data Technologies** : Hive, HDFS

**Development & Collaboration Tools** : GitHub, Linux, Airflow

**Streaming Services** : Confluent Kafka, Kinesis

**CERTIFICATIONS:**

• **GCP Cloud Digital Leader certification**

• **Google Cloud Associate Cloud Engineer**

• **Google Professional Cloud Architect** – **Pursuing**

• AWS Associate Developer certification

• OCI (Oracle Cloud Infrastructure) Generative AI

• SI (Solution Integration) Mongo Architect

• Microsoft Azure Data Fundamentals (DP-900) – Pursuing

• AWS Solutions Architect – Pursuing

**AWARDS AND RECOGNITIONS:**

• Recognized as a **Rising Star** for transforming legacy code in AWS environment from conventional Livy-based jobs to Glue-based for achieving Data Governance.

• **Promoted twice in 3 years** as a testimonial to my technical acumen and diverse contribution.

• Achieved **Innovative Minds** recognition for developing AI-based conversational interface and code conversions using a combination of RAG (Retrieval Augment Generation) and AWS Bedrock agents and Prompt Engineering techniques.

• Led a team of 3 to build a **LSHC (Life Science and Health Care)** use case for **Capgemini’s Annual National Hackathon Symposium** and finished as **runner-up** qualifying **against 50 teams**.

• Attended Internship training at **HHP Diesel Loco Shed** Hubballi.

• Lead the Inter-college **Advitiya Techno Fest for** the entire **Electrical cluster** organized in 2021.