

# DATA ENGINEER

## MAMIDI DIKSHITHA

Atlanta, GA | 469-588-4432 | [deekshithareddy573@gmail.com](mailto:deekshithareddy573@gmail.com)  
<https://www.linkedin.com/in/dikshitha-reddy-mamidi-38461b24b/>

---

### SUMMARY

Data Engineer with 3+ years of experience designing and implementing scalable data pipelines, APIs, and cloud-native solutions. Specialized in building ETL workflows on AWS using Lambda, DynamoDB, and Step Functions. Proficient in Python, GraphQL (Strawberry & Apollo), and data validation using Pydantic. Adept at developing federated API architectures and collaborating with cross-functional teams to deliver high-impact, production-ready data systems across multiple environments. Committed to leveraging cloud infrastructure and automation to optimize performance, reliability, and data accessibility.

---

### SKILLS/TOOLS

- **Programming:** Python
- **Cloud & DevOps:** AWS (Lambda, S3, DynamoDB, CloudFormation, Step Functions), GitLab (CI/CD)
- **Databases:** MySQL, SQL Server, Aurora, Teradata
- **Data Engineering Tools:** Informatica (basics), Jupyter, Tableau
- **Tools & Frameworks:** Jupyter, Tableau, Apollo, Strawberry GraphQL
- **Concepts:** ETL, Data Warehousing, Machine Learning (basics), Agile

---

### EXPERIENCE

#### Data Engineer

Delta Airlines – Atlanta, GA

*Sep 2022 – Present*

- Architect, develop, and maintain end-to-end ETL pipelines into AWS Cloud using Python and AWS Lambda functions.
- Design Pydantic models for robust data validation, ensuring data quality before ingestion into DynamoDB tables.
- Build GraphQL APIs with Strawberry GraphQL Gateway, deploying across DEV, SI, UAT, and PROD environments, enabling efficient data access for downstream applications.
- Leverage Apollo Server and Apollo Gateway to federate subgraphs into a unified supergraph, consolidating schemas and simplifying client queries.
- Utilize AWS Step Functions to orchestrate state-machine pipelines, triggering independent Lambda functions for sequential data processing tasks.
- Develop local development workflows with Node.js for API serving, linting and formatting to maintain code consistency.
- Collaborate via GitLab: create Merge Requests, apply labels, and follow GitFlow to merge feature branches into release branches seamlessly.
- Designed and implemented data warehousing solutions to organize and store large volumes of data for analytical purposes.
- Leveraged AWS services, including Lambda and DynamoDB, to build scalable, resilient, and efficient technology solutions.

- Collaborated with cross-functional teams, including data engineers, analysts, and business leaders, to deliver data-driven solutions.
- Architected, implemented, and operated data engineering pipelines and solutions using Agile methodology.
- Provided data scientists and analysts with the necessary datasets to support their analysis and modeling efforts.

### **Data Collection Intern**

**Sciffer Analytics** – Pune, India

***Feb 2021 – Jul 2021***

- Gathered and submitted observational information and required data points for analysis.
- Collaborated with team members on data validation, derived key insights, and coordinated tasks to address issues effectively.
- Ensured accurate and complete data collection to support predictive model development and improve analysis quality.

---

## **EDUCATION**

### **University of North Texas**

*Master of Science, Computer Science* | GPA: 3.6/4.0

Aug 2021 – May 2023

### **Guru Nanak Institutions Technical Campus**

*Bachelor of Technology, Information Technology* | GPA: 8.85/10.0

Aug 2016 – Sep 2020

---

## **PROJECTS**

### **Fake User Identification on Social Media (Twitter)**

- Developed a program to detect fake users and spam on Twitter using the **Random Forest algorithm** deployed in **AWS SageMaker**.
- Analyzed user, content, graph, structure, and time features to classify accounts.
- Compared various techniques and optimized for high accuracy.

---

## **CERTIFICATIONS**

- AWS Cloud Practitioner
- AWS Data Engineer
- Cloud Computing (Red Hat Academy)
- Data Science (Udemy)
- IoT in Smart Bridge with IBM Cloud

---

## **KEY ACHIEVEMENTS**

- Successfully reduced operational costs by implementing scalable cloud-based data solutions.
  - Enhanced data pipeline performance by optimizing Python scripts and leveraging cloud-based infrastructure.
  - Collaborated on a real-time analytics project that significantly improved decision-making processes for stakeholders.
-