

Libin Guo

832-668-4022 | Austin, TX | libinguo89@gmail.com | [LinkedIn](#) | [GitHub](#)

Senior Software Engineer (Data Intelligence) with 7+ years of experience designing and scaling backend systems for large-scale for large-scale data pipelines and platforms with Python/Spark/SQL

EXPERIENCE

Apple, inc

Senior Software Engineer (Data Intelligence)

Austin, TX | June 2024 - Present

- Scaled and optimized Microservice app to automate maintenance spark jobs across 20 PB iceberg datalake, improving query performance and driving a 20% YoY increase in internal lakehouse adoption
- Built workflow to convert Snowflake SQL to Spark SQL workflows with LLM judge eval framework to improve accuracy to reduce compute cost
- Launched a RAG enhanced chatbot w/ Lakehouse&snowflake MCP servers for runbook search, Q&A, table creation and answer data insight questions

Software Engineer(Data)

Austin, TX | Dec 2021 - May 2024

- Built and optimized large-scale data pipelines (Airflow + Spark on AWS EKS) from Hadoop, reducing runtimes by 75% and saving \$1.7M annually by optimizing Spark micro-batch execution for ads
- Drove adoption of Snowflake + DBT across the analytics org by establishing standards and best practices — boosting team productivity and strengthening stakeholder trust in data quality

JPMorganChase - Software Engineer(Data)

Dallas, TX | Oct 2020 – Nov 2021

- Enhanced Kafka streaming application to ingest data into Oracle in microservice architecture
- Designed and implemented end to end streaming data pipelines w/ AWS Glue jobs and Lambda functions for ETL processes from AWS Kinesis to DynamoDB with Terraform

Healthmine - Data Engineer

Dallas, TX | Sep 2018 – Oct 2020

- Built ETL pipelines (Python/SQL/AWS) loading multi-tenant data into Oracle; migrated Talend to Airflow for performance and maintainability
- Stood up an EDW on Postgres with Airflow + Spark on EMR; delivered daily ingestion and Superset dashboards for business-critical reporting

Qlarant Integrity Solutions, LLC - Senior Data Analyst

Dallas, TX | Mar 2016 – Aug 2018

- Analyze datasets to identify multi-million-dollar Medicare overpayments; automated SAS templates, cutting query runtime ~75%.
- Led proactive data analyses using SQL, SAS, and Tableau, uncovering hundreds of healthcare fraud leads

SKILLS

- **Languages:** Python, SQL, Java, Linux

- **Data:** Apache Spark, Iceberg, Trino, Pandas, Airflow, FastAPI, Kafka, Flink, Streamlit, DynamoDB, Snowflake, DBT

- **Cloud Infra:** AWS, Kubernetes, Docker, CI/CD, Helm Chart, Terraform

EDUCATION

University of Texas at Dallas - M.S., Information Technology and Management

GPA 3.82 | Dec 2015

Shanghai Normal University - M.A, English Language and Literature

GPA 3.4 | July 2014

Shandong Agricultural University - B.S, Computer Science; Minor in English

GPA 3.25 | Jul 2011

CERTIFICATIONS

AWS Certified Solution Architect Associate; SAS Certified Predictive Modeler and Advanced Programmer;