Tiffany Wang

Data Engineering Lead | Cloud-Native Architect | Real-Time Analytics Expert
📍 Redondo Beach, CA | 📞 (626)-223-6123 | ✉️ tiffany.wang.engineer@gmail.com
🔗 LinkedIn

SUMMARY

8+ years architecting **multi-cloud data platforms** (AWS†, Azure†, GCP) for high-growth enterprises. Expert in **Spark†**, **Delta Lake†**, and **event-driven architectures**, delivering **40% cost savings** and **50% latency reduction**. Certified Scrum Master leading Agile teams to build **PB-scale data lakes** and **real-time APIs** with 99.9% uptime. Passionate about IaC automation and AI/ML-driven data solutions.

TECHNICAL SKILLS

Core Stack:
AWS† (Glue†, S3†, Lambda), Azure† (Databricks†, Data Factory), GCP (BigQuery†), Spark† (Structured Streaming)
Data Tools: Delta Lake†, Unity Catalog†, Kafka†, Snowflake, DBT
Languages: Python† (PySpark), Scala†, SQL† (Advanced Optimization), Terraform†
DevOps: GitHub Actions†, Datadog, Airflow†, Kubernetes
Leadership: SAFe Agile†, Cross-functional Team Mentorship, Technical Roadmapping

PROFESSIONAL EXPERIENCE

Capital Group | Lead Data Engineer

Irvine, CA | 2019–2023

Multi-Cloud Retail Analytics Platform (AWS + Azure†)

- Built **Delta Lake†** pipelines ingesting **5TB/day** of clickstream data, reducing query latency by **50%** through Z-Order optimization
- **Technical Choice**: Migrated from Redshift to **Databricks SQL†**, cutting monthly costs by **$120K** via photon-engine acceleration
- Orchestrated **Kafka†**-to-S3 streaming with **Spark Structured Streaming†**, achieving **<500ms** event processing for real-time recommendations

IaC-Driven Infrastructure (Terraform†)

- Automated cloud resource provisioning for **200+ data pipelines**, reducing deployment errors by **35%**
- **Hook**: Implemented **Unity Catalog†** for centralized governance, enabling fine-grained access control across 10+ teams

Online AI & ML Program | Cloud Architect

Remote | 2023–Present

Event-Driven AdTech Architecture (Kafka† + Lambda)

- Designed **serverless event processing** handling **1M+ events/sec**, triggering personalized ad campaigns with **90% accuracy**
- **Impact**: Reduced infrastructure costs by **40%** via spot instance optimization (AWS Batch + Auto Scaling)

AI-Powered Inventory Forecasting

- Integrated **MLflow** with **Delta Lake†** to track 50+ model versions, improving forecast accuracy by **25%**
- Delivered **Power BI†** dashboards for executive stakeholders, reducing decision latency from days → hours

EDUCATION & CERTIFICATIONS

M.S. Software Engineering | Embry-Riddle Aeronautical University | 2015–2017
AWS Certified Solutions Architect | 2023
Databricks Lakehouse Specialist | 2024

Open Source:

- Contributed to **Delta Sharing**† Python SDK (GitHub: 1K+ stars)
- Authored **"Terraform for Data Engineers"** guide (Medium: 15K+ views)

† **JD-Aligned Keywords**
AWS/Azure/GCP | Databricks/Spark | Delta Lake/Unity Catalog | Kafka | Terraform | Power BI