

Tiffany Wang

Senior Data Engineer | Multi-Cloud & HIPAA Compliance Expert

📍 Redondo Beach, CA | 📞 (626)-223-6123 | ✉️ tiffany.wang.engineer@gmail.com

🔗 [LinkedIn](#)

SUMMARY

8+ years of experience designing **multi-cloud data architectures** (AWS⁺, Azure⁺, Databricks⁺) for healthcare analytics. Delivered **\$3.8M+ cost savings** through Delta Lake optimizations and IaC-driven scaling. Certified in **AWS Solutions Architect** with deep expertise in **HIPAA⁺ compliance**, **Spark performance tuning**, and **end-to-end data governance**.

TECHNICAL SKILLS

Core Platforms:

AWS⁺ (Glue⁺, S3⁺, Redshift⁺, Lambda), Azure⁺ (Data Factory⁺, ADLS⁺, Synapse), Databricks⁺ (Delta Lake⁺, Spark⁺)

ETL/ELT: dbt⁺, Airflow⁺, Dagster, Great Expectations

Data Modeling: Delta Lake/Iceberg⁺, Star Schema, Data Vault 2.0

DevOps: Terraform⁺, Docker⁺, Kubernetes, GitHub Actions

Compliance: HIPAA⁺, Column-Level Encryption, Audit Logging

Languages: SQL⁺ (Advanced), Python⁺ (PySpark, Pandas), Scala

PROFESSIONAL EXPERIENCE

Capital Group | Lead Data Engineer

Irvine, CA | 2019–2023

Multi-Cloud Healthcare Lakehouse (AWS + Azure)

- Architected **Delta Lake**⁺ pipelines processing **10TB/day** of PHI data, achieving **HIPAA**⁺ **compliance** via AES-256 encryption and RBAC policies
- **Technical Choice:** Migrated from Redshift to **Databricks SQL**⁺, reducing complex query latency by **65%** (\$450K/year cost savings)
- Orchestrated **Airflow**⁺ DAGs with **Terraform**⁺-managed infrastructure, enabling auto-scaling clusters for 99.99% SLA

Real-Time ML Feature Engineering (Databricks⁺)

- Optimized **Spark**⁺ jobs using adaptive query execution, cutting batch processing time from 4h → 22min
- **Hook:** Implemented **Delta Sharing**⁺ for secure data collaboration with 15+ external research partners

Online AI & ML Program | Cloud Architect

Remote | 2023–Present

IaC-Driven Data Platform (Azure⁺ + AWS)

- Deployed **Terraform**⁺ modules for multi-cloud resource provisioning, reducing environment setup time by **80%**
- Built **dbt**⁺ models with **Great Expectations** checks, improving data quality from 92% → 99.8% accuracy
- **Cost Optimization:** Transitioned cold data to **S3 Glacier** + **Azure Archive**, saving **\$220K/month**

Healthcare Compliance Automation

- Designed **Python**⁺-based audit system tracking 50+ HIPAA⁺ controls, reducing compliance review time by **70%**
- Integrated **Datadog**⁺ alerts for unauthorized data access attempts, achieving **100% audit readiness**

EDUCATION & ACHIEVEMENTS

M.S. Software Engineering | Embry-Riddle Aeronautical University | 2015–2017

AWS Certified Solutions Architect | 2023

Delta Lake Advanced Certification | Databricks Academy | 2024

Open Source Contributions:

- Authored **Airflow-Databricks[†]** connector (1.2K+ GitHub stars)
- Contributed to **dbt-delta[†]** plugin for seamless Delta Lake integration

[†] JD-Aligned Keywords

AWS/Azure/Databricks | HIPAA | Terraform | Delta Lake | Spark | Airflow