

# Tiffany Wang

---

**AWS Data Engineer | Healthcare Data Specialist**

Redondo Beach, CA | (626)-223-6123 | [tiffany.wang.engineer@gmail.com](mailto:tiffany.wang.engineer@gmail.com)

LinkedIn

## SUMMARY

---

8+ years of experience in data engineering with expertise in AWS ecosystems, ETL development, and healthcare data integration. Proficient in AWS Glue, Athena, PySpark, and Terraform, delivering 99.9% data accuracy for mission-critical systems. Certified in HIPAA compliance with a proven track record of optimizing data workflows in AGILE environments.

## TECHNICAL SKILLS

---

**Cloud Platforms:** AWS (Glue<sup>+</sup>, Athena<sup>+</sup>, Data Lakes<sup>+</sup>, Redshift), Snowflake<sup>+</sup>

**ETL Tools:** Informatica PowerCenter, IICS, Apache Hudi<sup>+</sup>

**Languages:** Python<sup>+</sup> (PySpark), SQL (Complex Joins/CTEs), TypeScript

**Databases:** MySQL<sup>+</sup>, PostgreSQL<sup>+</sup>, Amazon Aurora<sup>+</sup>, SQL Server

**DevOps & IaC:** Terraform<sup>+</sup>, CloudFormation, Git<sup>+</sup>

**Monitoring:** AWS CloudWatch, APM tools

**Certifications:** AWS Certified Data Engineer, Google Cloud Professional Data Engineer, HIPAA Privacy & Security

## PROFESSIONAL EXPERIENCE

---

**Capital Group | AWS Data Engineer**

Irvine, CA | 2019–2023

**AWS Data Lake Integration (Glue<sup>+</sup> + Athena<sup>+</sup>)**

- Migrated **15+ legacy ETL workflows** to **AWS Glue<sup>+</sup>**, reducing data latency by **40%** and operational costs by **\$15K/year**
- **Technical Choice:** Implemented **Apache Hudi<sup>+</sup>** for real-time data ingestion, improving query performance by **35%** compared to traditional batch processing
- Designed **row-level security** policies in **AWS Athena<sup>+</sup>** to ensure HIPAA compliance for **300K+ patient records**

#### ETL Process Optimization with Terraform<sup>+</sup>

- Automated infrastructure provisioning using **Terraform<sup>+</sup>**, reducing deployment time from **2h → 15min**
- Integrated **Python-based data quality checks** (Great Expectations) with **AWS Lambda<sup>+</sup>**, cutting dirty data incidents by **90%**

## Online AI & ML Program | Healthcare Data Engineer

Remote | 2023–Present

#### Real-Time Data Pipeline (AWS Kinesis<sup>+</sup> + Snowflake<sup>+</sup>)

- Built serverless ETL workflows using **AWS Glue<sup>+</sup>** and **Kinesis<sup>+</sup>**, processing **2TB/day** of clinical IoT data with **<500ms latency**
- **Impact:** Reduced ICU alert false positives by **30%** through sensor data enrichment and real-time anomaly detection

#### Legacy System Migration to Snowflake<sup>+</sup>

- Migrated 50+ Excel-based reports to **Snowflake<sup>+</sup>** via **AWS Glue<sup>+</sup>**, saving **200h/month** manual effort
- Trained 20+ analysts on **SQL self-service queries**, increasing data utilization by **150%**

## KEY METRICS & IMPACT

- **40% Latency Reduction** | **90% Data Quality Improvement** | **\$15K/year Cost Savings**
- **35% Query Performance Gain** (Apache Hudi<sup>+</sup>) | **150% Utilization Growth**

† JD-Aligned Keywords

AWS Glue/Athena/Hudi | Terraform | Python | Snowflake | HIPAA | AGILE