

# Tiffany Wang

---

**Objective:** Associate Data Engineer - FinThrive

**Focus Areas:** Azure<sup>+</sup> | PySpark<sup>+</sup> | Healthcare Data Systems | Agile Development<sup>+</sup>

Phone: (626)-223-6123 | Location: Redondo Beach, CA 90277

LinkedIn: [linkedin.com/in/tiffanywangengineer](https://www.linkedin.com/in/tiffanywangengineer) | Email: [tiffany.wang.engineer@gmail.com](mailto:tiffany.wang.engineer@gmail.com)

## SUMMARY

---

**Cloud Data Engineer** with 8+ years of experience building **healthcare-focused data solutions**, including real-time pipelines serving 50M+ users. Proficient in Python<sup>+</sup>/PySpark<sup>+</sup> and Azure<sup>+</sup> ecosystem (Data Factory, Databricks), delivering 90% latency reduction and \$500K+ cost optimization. Certified in Agile methodologies with experience mentoring 5+ junior engineers.

## SKILLS

---

**Core Stack:**

Python<sup>+</sup> (PySpark<sup>+</sup>, Pandas), Azure<sup>+</sup> (Data Factory<sup>+</sup>, Databricks<sup>+</sup>, Event Hub), SQL<sup>+</sup>, Spark Streaming<sup>+</sup>, Git<sup>+</sup>

**Data Engineering:** ETL Optimization, HIPAA Compliance, CDC Pipelines, Real-time Analytics

**Collaboration:** Agile Scrum<sup>+</sup>, Cross-functional Team Leadership, Technical Documentation

## EXPERIENCE

---

### Senior Cloud Data Engineer

Capital Group | Irvine, CA | 2019–2023

Healthcare Claims Analytics Platform (Azure<sup>+</sup>)

- Migrated 10TB+ EHR data to **Azure Databricks<sup>+</sup>**, reducing Spark job runtime by **40%** through Delta Lake optimizations
- **Technical Choice:** Implemented Azure Data Factory<sup>+</sup> over Airflow for HIPAA-compliant orchestration, achieving **99.95% SLA**
- Built PySpark<sup>+</sup> pipelines processing 1M+ daily claims, improving data freshness from 24hr → 15min

### Real-Time Patient Monitoring System

- Designed Kafka→Azure Event Hub<sup>+</sup> streaming pipeline with **<200ms latency**, detecting 98% abnormal vital signs
- **Hook:** Integrated Databricks MLflow for model versioning, adopted by 15+ data scientists

## AI Data Infrastructure Consultant

FinTech Innovation Lab | Remote | 2023–Present

### PHI Data Anonymization Framework

- Developed column-level encryption using Azure Purview<sup>+</sup>, reducing PII exposure risk by **70%**
- **Impact:** Enabled GDPR/HIPAA dual compliance for 5M+ patient records

### Clinical Trial Data Lake

- Architected Azure Synapse-based solution unifying 20+ data sources, accelerating insights delivery by **3x**
- Mentored 3 junior engineers on PySpark<sup>+</sup> optimizations and CI/CD practices

## EDUCATION

---

MS Software Engineering | Carnegie Mellon University | 2015–2017

BS Computer Science | Tsinghua University | 2011–2015

Certifications:

- Microsoft Azure Data Engineer Associate
- Certified ScrumMaster® (CSM)

**Quantifiable Impact:**

- 90% latency reduction | \$500K+ cost savings | 99.95% SLA compliance

**† JD-Aligned Keywords**

Azure Data Factory/Databricks | PySpark | Agile | Healthcare Data | HIPAA Compliance