Ji-Won Park Staff Data Engineer

Charlotte, NC · jiwon.park@gmail.com · +1 704-555-0149 · linkedin.com/in/jiwonpark

PROFILE SUMMARY

- **Staff-level Data Engineer** with **11 years** at **Wells Fargo and American Express**, leading the **risk data platform** behind regulatory reporting for credit, market, and liquidity risk under **BCBS 239 and SOX**.
- Chooses the platform stack across languages (**Python**, **Scala**, SQL), processing (**Apache Spark** on Dataproc, **dbt**), storage (**BigQuery**, **Apache Iceberg**), streaming (**Kafka**), orchestration (**Airflow** on Cloud Composer), and infrastructure (**Terraform**).
- Expert in **risk data aggregation**, **column-level lineage**, and **migrations off Teradata and Hadoop**, tying every **critical data element** to an owner, a quality check, and a control that auditors can test.
- Partners with **risk reporting, finance, and compliance leaders** and **internal audit** on what the platform must prove, briefs the **technology risk committee** every quarter, and reviews designs across **6 data engineering teams**.
- Grew **4 senior engineers** into lead roles, wrote the **data platform engineering standards** and **ADR template** used across risk technology, and sits on the **principal engineer promotion panel** for data.

TECHNICAL SKILLS

Languages:	Python, SQL, Scala, Java, Bash
Processing & Transformation:	Apache Spark (Dataproc), PySpark, dbt, Apache Beam, Trino
Streaming & Messaging:	Apache Kafka, Spark Structured Streaming, Google Cloud Pub/Sub, Avro, Schema Registry
Orchestration & Ingestion:	Apache Airflow (Cloud Composer), Control-M, Informatica PowerCenter, Datastream
Warehouses & Lakehouses:	BigQuery, Apache Iceberg, BigLake, Teradata, Hadoop (Hive, HDFS), Apache Parquet
Cloud & Infrastructure:	Google Cloud, Microsoft Azure, Terraform, Kubernetes (GKE), Docker, Jenkins
Data Quality & Observability:	OpenLineage, DataHub, Great Expectations, dbt tests, Dataplex, Datadog

EDUCATION

Virginia Tech — **B.S. in Computer Science** — Blacksburg, VA · Aug 2011 to May 2015

WORK EXPERIENCE

Wells Fargo — **Principal Engineer** — Charlotte, NC · Mar 2021 to Present

- Led the move of **1,140 risk feeds** from Teradata and Hadoop to **BigQuery** on Google Cloud, running both stacks in parallel with row-level reconciliation, and missed no regulatory filing dates during cutover.
- Designed a shared risk data model on **Apache Iceberg** with conformed counterparty, product, and legal-entity dimensions as the one source for FR Y-14 and FR 2052a reporting, cutting reconciliation breaks by **81%**.
- Built column-level lineage in **OpenLineage** and **DataHub** from source feeds to each line item in the regulatory reports, cutting internal audit findings against BCBS 239 from 17 to **3**.
- Partitioned and clustered the largest position tables in **BigQuery** and moved each risk team onto its own slot reservation with a monthly cost report, saving **$2.3M a year** in compute.
- Started a monthly data contract review where source-system owners, risk reporting, and compliance agree on schemas, quality thresholds, and SOX control owners, with **52 contracts** signed so far.

American Express — **Senior Data Engineer** — Phoenix, AZ · Jul 2015 to Feb 2021

- Built **Kafka** and **Spark Structured Streaming** jobs that turn card authorization events into real-time spend alerts in the Amex app, with p99 delivery **under 2 seconds** during holiday peaks.
- Replaced **Control-M** schedules with **Apache Airflow** DAGs for merchant reporting, adding sensors on upstream files and SLA alerts, which cut missed file deadlines from 11 a quarter to **1**.
- Wrote **Terraform** modules and Jenkins pipelines so data engineers could create Kafka topics, storage paths, and service accounts by pull request, cutting the wait for each from 3 weeks to **2 days**.