Randy Wang

Los Angeles CA • randywang127@gmail.com • (408) 598-0001

PROFILE

Experienced Senior Data Engineer with 8+ years of expertise in data infrastructure, large-scale distributed processing, and ML pipeline enablement. Proven track record of architecting petabyte-scale data lakehouses, optimizing ETL performance, and building production-grade GenAI applications. Passionate about transforming raw data into business intelligence through modern cloud solutions (AWS), scalable systems, and clean code.

TECHNICAL SKILLS

- Programming: Python, SQL, Scala, PySpark
- Cloud & DevOps: AWS (EKS, Lambda, Bedrock, EC2, S3, Glue, Redshift, Athena, Fargate, API Gateway, IAM, CDK), GCP, Terraform, Docker, GitLab CI/CD
- Data Engineering: Apache Spark, Kafka, Airflow, Apache Iceberg, EMR, Glue, Lakehouse Architecture
- Databases & Warehousing: Snowflake, BigQuery, DynamoDB, StarRocks, ClickHouse, MySQL, RDS, Redshift Spectrum
- AI/ML & GenAI: LangChain, AWS Bedrock, FAISS (Vector DB), Sentence Transformers, LangFuse, RAG Pipelines, MLflow, GenAI experimentation

EDUCATION

University of California, Los Angeles   **Los Angeles, CA**
Bachelor of Science, Statistics   Graduation Date: March 2015

PROFESSIONAL EXPERIENCE

Demandbase   **Los Angeles, CA**
Senior SDE – Core Data Platform   Nov 2023 – Present

- Enterprise Data Platform & Architecture: Architected and built a modern, large-scale ETL platform on AWS to process terabytes of data, leveraging a sophisticated stack including Python, Apache Iceberg, and dynamically generated Airflow DAGs. Automated infrastructure management with Terraform, resulting in a 30% reduction in costs and a 40% improvement in data processing efficiency.
- Lakehouse Implementation: Led the end-to-end design and implementation of a petabyte-scale analytics platform by architecting an Apache Iceberg data lakehouse on AWS, developing optimized Scala/Spark ETL pipelines, and automating the entire data lifecycle with robust GitLab CI/CD pipelines.
- Generative AI (Text2Cube): Architected the core AI engine for Text2Cube, a platform that translates natural language into executable data queries. Designed and implemented a production-ready Retrieval-Augmented Generation (RAG) pipeline in Python using LangChain, Sentence Transformers, and a FAISS vector database to ensure high-accuracy semantic understanding of user intent.
- AI Infrastructure & Observability: Engineered the end-to-end serverless deployment on AWS Lambda and containerized the application with Docker. Implemented full system observability with LangFuse to monitor the LLM chain, which reduced debugging time and improved system reliability.
- GenAI Application Development: Spearheaded the development of a generative AI-powered data analyst application on the AWS ecosystem, leveraging AWS Bedrock to translate natural language user queries into real-time business insights and deploying the scalable microservices architecture on Amazon EKS for resilient operation.

- Cloud-Native Integration: Architected the end-to-end cloud-native solution by integrating a suite of AWS services, including AWS Lambda for serverless processing, Amazon DynamoDB for high-performance data retrieval, and a fully automated CI/CD pipeline to ensure rapid and reliable delivery of the AI platform.

Amazon Studio   **Los Angeles, CA**
Data Engineer 2 – Studio Core Data Engineer   May 2021 – Nov 2023
- Architected data infrastructure on AWS platforms, including Airflow, Fargate, and Athena, achieving a 15% reduction in monthly costs and 20% efficiency gains.
- Led a team of 5 engineers to migrate 40+ daily ETL pipelines in 3 months, incorporating Python and SQL to streamline workflows and ensure data integrity.
- Built a robust microservices architecture to support scalable data ingestion and processing for multiple teams.
- Utilized EMR Spark for high-volume data processing, implementing PySpark jobs for complex transformations.

Verishop   **Los Angeles, CA**
Data – Marketing Technology   June 2020 – May 2021
- Designed and developed ETL pipelines to ingest and process web/app traffic data into Redshift, enabling real-time analytics.
- Created a product recommendation model using Spark, increasing product CTR by 10% and saving 30 hours per week for the planning team.
- Enhanced the Looker platform through advanced data modeling, delivering actionable insights to internal teams via 10+ custom reports.

TechStyle Fashion Group   **Los Angeles, CA**
Data Engineer   June 2019 – June 2020
- Developed automated ETL pipelines for social media, third-party vendor, and transactional data using Airflow and SQL.
- Designed Snowflake databases to support 50+ daily users, enabling real-time data visualization through Tableau.
- Implemented Spark-based pipelines on AWS EMR, facilitating advanced data modeling for data scientists.

Lionsgate Entertainment   **Los Angeles, CA**
Data ETL Developer   June 2017 – June 2019
- Enhanced existing ETL workflows and optimized data pipelines using Python and Airflow, improving reliability and performance.
- Standardized data structure and modeling processes across multiple sources, ensuring seamless integration.
- Built Tableau dashboards to provide actionable marketing and revenue insights for executive decision-making.

Sony Picture Entertainment   **Los Angeles, CA**
Advertising Data Analyst   June 2016 – June 2017
- Designed and managed Snowflake databases, improving data retrieval efficiency by 15%.
- Conducted advanced SQL queries to profile customer behavior and platform performance, delivering key insights.
- Created Tableau dashboards for revenue forecasting and campaign analysis, supporting strategic planning.