

KUAN-PEI (YUKI) LAI

(617) 470-9920 | yukiii04160@gmail.com | [Personal Website](#) | [LinkedIn](#) | [GitHub](#) | [Tableau](#) | Boston, MA

Experienced Data Engineer with 5+ years of expertise in developing and maintaining cloud-based data ecosystems and data engineering solutions to support various teams, along with strong data science knowledge and extensive programming skills. With a philosophy rooted in establishing FAIR data principles.

PROFESSIONAL SKILLS/ KNOWLEDGE

AWS: EC2, S3, Lambda, EventBridge, DynamoDB, Athena, Neptune, RDS, Glue, CloudFormation, SNS

Database & BI: Data Warehouse, Data Lake, MS SQL Server, MySQL, PostgreSQL, Spark SQL, NoSQL, dbt

Analytics Tools & Programming: Python (boto3, Pyspark, pandas, numpy, scikit-learn), R, R Markdown, R shiny, Databricks (Apache Spark), Git, KNIME, Nextflow, C#, C, Java, JavaScript, Visual Basic

Data Science & ML: EDA, NLP, Linear/Logistic/Regularized Regression, KNN, Decision Tree, Random Forest, GBM, XGBoost, K-means, DBScan, Supervised/Unsupervised Learning, Data Mining

Data Visualization: Tableau, Power BI, Qlik Sense, R (ggplot2), Python (plotly, matplotlib, seaborn, ipywidgets)

PROFESSIONAL EXPERIENCE

Data Engineer II

Jul 2024 - Mar 2025

Apriori Bio

Cambridge, MA

- Developed and deployed a metadata collection and tagging system using AWS S3, Lambda, DynamoDB, and SNS to support data platform and KPI tracking, ensuring effective metadata management and data governance.
- Collaborated with cross-functional teams to design scalable data ingestion pipelines using SDK and PySpark, scheduling job runs to streamline data centralization, processing, validation, and storage.
- Led the design of robust data models to catalog 150+ datasets from external sources, optimizing data discovery and accessibility across the organization, and created detailed documentation and dashboards for team support.

Data Engineer I/II

Oct 2021 - Mar 2024

Omega Therapeutics, Inc.

Cambridge, MA

- Developed and maintained an end-to-end cloud-based genomic experiment registration system based on R shiny interface and integrative AWS services such as EC2, S3, Lambda functions, and DynamoDB, to standardize metadata collection and to facilitate data findability and reusability across 400+ experiments and 19K+ samples.
- Designed and automated scalable ETL and data processing pipelines using AWS EventBridge, Lambda functions, and API, and built RESTful APIs to seamlessly collect data and utilize in the pipeline infrastructure, mitigating manual errors, enhancing efficiency, and ensuring data quality.
- Established a data intelligence platform by integrating large scale internal and external data using AWS services such as CloudFormation, Neptune, and Athena, and built a data environment that supports advanced data analytics.
- Worked closely with data scientists to capture modeling needs and data requirements to design data models, construct reliable and interoperable database systems, and create visualizations and dashboards.
- Collaborated with the IT team to optimize system technology, conduct code reviews, maintain cloud and data infrastructure, and troubleshoot system-related and data infrastructure issues in a timely manner.

Business Intelligence Engineer

Aug 2015 - Aug 2017

Fu Jen Catholic University

Taipei, Taiwan

- Extracted and aggregated data from databases by SQL and developed ETL process and data warehouses using SSIS.
- Established OLAP cubes using SSAS and deployed jobs in SQL Server Agent to automate annual report systems.
- Built and maintained 10+ key business intelligence dashboards by BI tools to support monitoring and tracking KPIs.
- Interfaced with non-technical offices to define data needs and support analytics solutions based on large datasets.
- Performed routine data quality check, resolved data-related issues, and collaborated with DBA on data governance.
- Educated, trained, and provided consultation to the end user group on the ad hoc reporting and analysis tool.

EDUCATION

Northeastern University

Jul 2021

Master's in Analytics | GPA: 4.0 | [Experiential Network Badge](#)

Boston, MA

Soochow University

Jun 2015

Bachelor's in Computer Science & Information Management

Taipei, Taiwan