

Andrew Le

andrewle713@gmail.com | Remote (based in Vietnam) | LinkedIn | GitHub

EDUCATION

University of California, Irvine

Bachelor of Science, Computer Science

Irvine, CA

June 2024

SKILLS

- **Programming Languages:** Python • C/C++ • Java • JavaScript • HTML/CSS • SQL
- **Tools:** Git • Docker • AWS (EC2) • DBever • Lightup • Jira • Confluence • Lucidchart • Kafka • Qlik Replicate
- **Applications:** MySQL Workbench • MongoDB Compass

PROJECTS

Smart Calendar

Personal

Python, MongoDB, OpenAPI

July 2025 - Present

- Developed backend services to manage calendar events and task data, with MongoDB for storage and query
- Integrated OpenAI API to parse natural language commands into structured scheduling actions (e.g., “Remind me to book tickets tomorrow”)
- Built an intent-mapping system to route interpreted user goals to the appropriate API endpoints and database operations
- Architected a modular command-processing framework that supports dynamic task types like scheduling, summarizing, and behavioral recommendations

Data Lineage PoC

Skechers

Python, Airflow, OpenLineage, Marquez, Bash

June 2024 - Oct. 2024

- Developed an internal PoC to demonstrate automated end-to-end data lineage tracking across batch workflows using OpenLineage and Marquez
- Integrated lineage tracking into existing DAGs on Astronomer’s managed Airflow platform, enabling visibility into job-level and field-level transformations
- Configured lineage emitters and metadata collection via Python scripts and YAML-based pipeline configs on a secure bastion-host environment
- Collaborated with data engineering stakeholders to validate feasibility and future production integration of open-source observability tools

WORK EXPERIENCE

Skechers

Manhattan Beach, CA

Data Engineering Intern

June 2024 – Oct. 2024

- Engineered real-time data pipelines using Airflow, Kafka, and Databricks, reducing latency in e-commerce systems
- Refactored 40+ batch tasks into streaming workflows, improving system responsiveness and observability
- Implemented OpenLineage & Marquez PoC, enabling end-to-end data lineage tracking across 300+ DAG tasks
- Developed automated alerting and monitoring systems with Lightup and PostgreSQL
- Collaborated with vendors including Amazon, Astronomer, and Marquez to evaluate and deploy data governance tools
- Authored technical documentation for pipeline configurations and monitoring processes