

Weijie (Jack) Zhang

wzhang3912@gmail.com | 747-283-5852 | linkedin.com/in/wzhang3912 | U.S. Permanent Resident

Education

University of California San Diego
Bachelor of Science in Data Science

Expected: March 2026
GPA: 4.0/4.0

- Coursework: Database Management, Scalable Data System, Data Visualization, Machine Learning, Deep Learning, Data Structure & Algorithm, Probability & Statistics, Optimization, Linear Algebra

Experience

Data Infrastructure Engineer Intern
Boundary RSS

June 2025 – Present
Los Angeles, CA (Remote)

- Implementing modular **ETL pipelines** in Python to automate extraction, transformation, and delivery of **geospatial** datasets into a data mesh infrastructure for ML-ready applications and analytical dashboarding
- Engineering an **automated dataset upload system** that allows users to onboard new datasets with minimal engineering support, improving extensibility and reducing manual intervention
- Developing secure, serverless data ingestion workflows by integrating **Python scripts** with existing **AWS services** such as AWS Lambdas and API Gateway to support API-based geospatial data pipelines

Software Engineer Intern
Green Optic Films

Mar 2025 – Present
Philadelphia, PA (Remote)

- Developing a surgical video analysis desktop application in **Python** and **PyQt** to streamline downstream analytics
- **Collaborating** with cross-functional teams to improve existing UI/UX and engineer a new scoring system for surgical instrument misuses, translating business needs into technical product features
- Refactoring a legacy codebase using OOP principles and introducing **PyTest** with **Github Actions** for CI/CD, reducing debugging time by **25%** and accelerating software release by **2x**
- **Automating** extraction and transformation of annotation data into structured formats and building **interactive dashboards** in Tableau to visualize instrument usage trend, supporting reporting and decision-making

Instructional Assistant
Halcioğlu Data Science Institute UC San Diego

Sep 2024 – June 2025
La Jolla, CA

- Hosted weekly office hours, providing individual and group support to **100+** students for understanding **data science course concepts** (Statistics, Data Structure and Algorithm, Practice of Data Science)

Research & Projects

Research Project: Meta Learning as Planning | *Python (PyTorch, OpenAI's Gym), Docker, Linux Shell*

- **Engineering** software infrastructure with a research team of 5, developing algorithms upon the Decision Transformer models for solving planning problems in Reinforcement Learning
- Building training and inference pipelines in **Python** and **PyTorch**, containerizing environments with **Docker**, and automating workflow via **Linux shell scripting** on computer clusters, reduced experiment setup time by 50%

AI-Driven SQL Database Tool | *Python, PostgreSQL, SQL, Large Language Model, Prompt Engineering*

- **Simplified database access** by building an AI agentic pipeline with a chatbot interface to automatically generate SQL queries for PostgreSQL databases, reducing query-writing **time by 75% per task**
- Utilized prompt engineering techniques and integrated **Large Language Models (LLM) API** with real-time streaming to ensure smooth and robust user experience

Technical Skills

Programming: Python (Pandas, Numpy, PySpark, Scikit-learn, PyTorch, TensorFlow, Plotly, BeautifulSoup, PyQt, PyTest), SQL, Java, JavaScript, D3.js, HTML/CSS

Technologies: Git, Docker, AWS, Linux Shell, PostgreSQL, Tableau

Concepts: Software Engineering, Agile, CI/CD, Machine Learning, Infrastructure-as-code, REST APIs, ETL Pipeline