

James Liu

312-545-1767
james.liu27@ncf.edu
github.com/jimmmmmmmmmmmmy
Sarasota, FL

EDUCATION

New College of Florida

Expected 2027

Master of Science in Data Science • GPA 3.9

Oregon State University

2024

Bachelor of Science in Computer Science • GPA 3.8

EXPERIENCE

Research Assistant • USDA Science Computing Initiative • Beltsville, MD

2026

- Co-wrote and submitted two manuscripts to Plant Disease and Journal of Fungi with Dr. Ezekiel Ahn.
- Developed an L2-regularized multinomial logistic regression model to classify cacao flower-bud genotypes from hyperspectral reflectance.
- Compiled a sorghum resistance atlas by linking seed-fungus profiles to weather data.

Data Consultant • Lost Forty Operations LLC • Austin, TX

2024 - Present

- Built a financial model from Bloomberg, FactSet, and public filings as the basis for a \$10M allocation.
- Launched a due-diligence RAG system over 150+ filings using Claude and PostgreSQL/pgvector.
- Extended the RAG stack into a search and read tool callable by a local agent harness.

Business Owner • The Plantier • Chicago, IL

2018 - 2022

- Scaled a rare-plant retailer to seven-figure annual revenue with a storefront and greenhouse.
- Modeled exotic-plant prices from Instagram, eBay, Etsy, and nursery catalogues.

PROJECTS

PIT-Tag Analytics MCP (Mote Marine / NOAA) • Python, SQL, LLM tool-calling, MCP

- Analyzed NOAA restoration-site fish habitat data with a 3-student team and Mote researchers.
- Implemented SQL/Python pipelines with predictive ML on 7M+ fish detections.
- Building an MCP server for researchers to query fish data in natural language.

Stylometric NLP Grader • Python, scikit-learn, TF-IDF / n-grams, applied statistics

- Built a stylometric NLP grader to compare two submitted USDA manuscripts with peer papers.

Market Dashboards • Python, FastAPI, PostgreSQL, Prefect, Docker, JS • breadthalyzer.com · straddlyze.com

- Shipped live market-breadth and macro-event dashboards with Prefect ETL and equities REST APIs.

TECHNICAL SKILLS

Python / ML: pandas, NumPy, scikit-learn, PyTorch, Jupyter

NLP / GenAI: RAG, MCP, LLM tool-calling, document ingestion, embeddings, local LLM inference

Engineering: SQL, ETL, Prefect, FastAPI, PostgreSQL/TimescaleDB, Docker, React/TypeScript, Git