

# Alexander Joos

+1-978-799-9242 | [aaaj46@cornell.edu](mailto:aaaj46@cornell.edu) | [linkedin.com/in/alexanderjoos1](https://www.linkedin.com/in/alexanderjoos1) | [github.com/alexanderjoos](https://github.com/alexanderjoos)

## EDUCATION

---

**Cornell University** Ithaca, NY  
*MEng Computer Science* Jan. 2026 – Dec. 2026

**Cornell University** Ithaca, NY  
*B.A. Information Science (Data Science)* Aug. 2022 – Dec. 2025

**National University of Singapore** Singapore  
*Yale NUS College Exchange Program* Jan. 2025 – May 2025

**Relevant Coursework:** Graduate Machine Learning, Graduate Software Engineering, Data Science, Backend Development, Graduate Algorithm Design, Graduate Deep Learning, Statistics, Linear Algebra, Computer Organization

## EXPERIENCE

---

**Fidelity Investments** June 2026 – Current  
*Graduate Data Engineering Intern, Database Modernization* Merrimack, NH

- Building capital markets streaming pipelines and SingleStore schemas, ingesting data via Kafka & Aeron.

**Fidelity Investments** June 2025 – Aug. 2025  
*Data Engineering Intern, Cloud and Platform Engineering* Merrimack, NH

- Built and launched a Flask/AJAX web app on an Ubuntu VM to automate internal database model comparison workflows.
- Supported brokerage platform model analysis across Oracle, PostgreSQL, CockroachDB, and Snowflake databases.
- Implemented Python/SQL backend storage and retrieval logic with JavaScript/HTML/CSS frontend workflows in Flask.
- Deployed the application through Apache, SSL certificate setup, firewall whitelisting, and DEV/PROD rollout.

**American Family Insurance: Homesite** May 2024 – Aug. 2024  
*Data Analytics Intern* Boston, MA

- Built BigQuery pipelines from raw source tables for a Fortune 50 retailer's Partner Management reporting layer in SQL.
- Built 8 Looker dashboards for Partner & Executive teams; automated SQL query generation via Excel/Sheets scripts.

**American Family Insurance: Homesite** June 2023 – Aug. 2023  
*Partner Distribution Intern* Boston, MA

- Conducted Excel, Power Pivot, and Power Query analysis on top U.S. auto carrier to infer competition feature importance.

## PROJECTS

---

**Apache Airflow Anomaly Detection** – *Agile Open Source Systems & Backend Feature Development* May 2026

- Built Airflow-native task-runtime anomaly detection with DAG-file config, custom detectors, and 4 built-in algorithms.
- Implemented backend/task-sdk logic, structured result details, task-log observability, CI tests, and example DAG/docs.

**LOB-Driven Optimal Execution Strategy** – *Quantitative Trading & Market Microstructure* May 2026

- Designed TWAP-beating execution strategy using order-book imbalance & spread-threshold signals on tick-level LOB data.
- Backtested on INTC/MSFT/AMZN/GOOG with smoothed grid search and separate parameter fits by spread regime.

**LLM Text Watermarking** – *ML Research Reimplementation* May 2026

- Reimplemented Kirchenbauer et al. LLM watermarking in PyTorch with soft green-list logit biasing and z-test detection.

**Clustering Analysis on Workplace Gender & Power** – *NLP & Unsupervised ML* Dec. 2024

- Cleaned email bodies into a high-dimensional TF-IDF matrix, capturing sentiment across 15,000+ messages.
- Used DBSCAN, K-means, RiveterNLP, and cross-tabulation approaches to analyze gender, power, and positional clusters.

**Old Bailey Verdict Prediction Model** – *LLM Fine-Tuning & Classification* Nov. 2024

- Fine-tuned a Hugging Face LLM on 300+ years of > 100,000 Old Bailey case briefs across 8 legal categories.
- Performed "GUILTY" vs. "NOT GUILTY" classification, achieving 71% weighted-average F1 and outperforming baseline.

## SKILLS

---

**Languages:** Python, Java, OCaml, C/C++, SQL, R, JavaScript, HTML/CSS

**Frameworks & Libraries:** Flask, REST APIs, PyTorch, transformers, scikit-learn, pandas, NumPy, matplotlib, SQLAlchemy

**Data & Cloud Tools:** BigQuery, Vertex AI, Looker, Power BI, SQLite, Oracle, PostgreSQL, Snowflake, DBEaver

**Developer Tools:** Git, GitHub, Docker Compose, Apache Airflow, Bash, Jupyter, Google Colab, Jira, Confluence