

Konrad Trestka

Chicago, IL | ktrestka05@gmail.com | github.com/KonradT22 | linkedin.com/in/konrad-trestka-dev

EDUCATION

University of Minnesota Twin Cities, College of Science and Engineering

B.S. Computer Science & B.S. Data Science

Minneapolis, MN

Expected May 2027

Languages: English, Polish | Relevant Coursework: Practice of Database Systems, Data Structures & Algorithms, Regression & Statistical Computing, Software Engineering

SKILLS

Languages: SQL, Python, JavaScript/TypeScript, Ruby, Java, R, Bash

Data & Cloud: Azure Data Factory (ADF), Azure SQL, Azure Blob Storage, ETL/ELT, data validation, stored procedures, REST APIs, JSON, Docker, pandas, SQLite, distributed computing

Tools: Git/GitHub, Linux/Ubuntu, SLURM (HPC batch scheduling), Conda, rsync, Nginx, Transparent Data Encryption (TDE)

WORK EXPERIENCE

Data Engineer Intern

PURIS

Aug 2025 – Present

Minneapolis, MN

- Designed Azure SQL tables and stored procedures for complex JSON parsing and idempotent upserts to support an evolving production schema.
- Built and maintained a production-grade ETL pipeline in Azure Data Factory (ADF) that triggers daily, pulling JSON payloads via REST API into Azure SQL.
- Engineered an OT/IT data pipeline extracting high-frequency time-series metrics from a VantagePoint historian database into an encrypted SQL backend.
- Orchestrated a zero-downtime migration of a critical Operational Technology (OT) testing application from a legacy MSP-hosted environment to a new internal SQL Server, securely transferring user logins and TDE encryption keys via a parallel-run cutover.
- Preserved 12.3 GB / 119,000+ files of historical corporate data by staging and uploading to Azure Blob Storage with a validated local audit vault.

Undergraduate Researcher – Data & Simulation Pipelines for Volatility Modeling

University of Minnesota Twin Cities (Faculty-Supervised UROP)

May 2026 – Aug 2026

Minneapolis, MN

- Engineered an end-to-end research data pipeline ingesting SPX/SPXW options data via the ThetaData API across 20 trading dates, interpolating irregular market quotes onto a fixed 7x11 grid to produce standardized, ML-ready datasets with reproducible experiment manifests.
- Parallelized synthetic dataset generation across 100 SLURM batch jobs on the Minnesota Supercomputing Institute (MSI), evaluating 10,000 Monte Carlo scenarios with deterministic per-scenario seeding and checkpointable outputs supporting reliable merge/restart behavior.
- Built automated data validation gates screening distributed outputs for missing IDs, duplicates, nonfinite values, and invalid prices, merging results into a verified 9,667-scenario dataset with zero duplicate or missing records.
- Trained an inverse neural network (held-out R^2 0.97–0.99) and implemented an 81-dimensional Mahalanobis out-of-distribution monitor with automatic fallback logic, detecting 100% of unseen market states at a ~1% false-positive rate.

Software Engineer Intern

Fine Global Assets

May 2024 – Dec 2024

Chicago, IL

- Designed and developed the corporate website using React and Node.js, improving user engagement by 25% and retention by 15%.

PROJECTS

AJAP – Self-Hosted Data Pipeline

Python, Flask, APScheduler, SQLite, LLM API

- Built an ingest, classify, store, and alert pipeline that pulls roles from multiple sources, normalizes and classifies them via an LLM API into SQLite, and syncs to Google Sheets with batched Discord alerts.
- Scheduled the pipeline as a continuous Linux daemon (APScheduler) with persistent state to avoid duplicate processing across restarts.

AWARDS & ACTIVITIES

1st Place, Quantitative Trading Competition (Financial Mathematics Association)

Google Developer Student Club | Social Coding | Lettuce Club | Ski and Snowboarding Club