

Aaron Gurovich



949-505-0956

aarongur@outlook.com

aarongurovich.com

Education

B.S. in Computer Science | *Texas Tech University, Lubbock, TX*

May 2026

- Minor: Mathematics
- Awards: Chess Scholarship, Presidential Transfer Scholarship
- Hillel President: Lead the student board and secured a \$5,000 multifaith grant for community events.
- Chess Team: Competing at the national collegiate level; USCF rating 1907.

Experience

Clinical Data Engineer Intern | *Johnson & Johnson MedTech Electrophysiology, Irvine, CA*

Jun 2025 – Present

- Accelerated clinical trial site selection and ensured diversity benchmark alignment by building an end-to-end optimization platform with a dual-objective quadratic algorithm and site-activation scheduler (R, Shiny, SQL, quadprog).
- Improved AI protocol deviation classification accuracy from 75% to 98% by optimizing RAG pipeline and cloud-based processing workflows (Python, AWS Lambda, SQS).
- Strengthened clinical study data reliability by designing a scalable data model and developing a CI/CD-compliant, object-oriented ingestion framework with automated testing and quality checks (PyTest, Jenkins, SonarQube).
- Reduced clinical data ingestion errors by 99.95% through a complete redesign of transformation logic and error-handling protocols (Pandas, S3, Redshift).
- Accelerated data insights by building an AI assistant that converts data-related questions into SQL to retrieve and display result tables directly to users (sqldf, dplyr).
- Awarded 2nd place at the J&J Data Science Showcase for the Generative AI-driven protocol deviation platform.

Research Assistant | *Interactive Data Visualization Lab, Texas Tech University, Lubbock, TX*

Oct 2024 – May 2025

- Designed MalScape, a scalable data processing and visualization platform for large-scale network traffic analysis (Flask, JavaScript, Python).
- Engineered 20+ features from billions of network records by building high-throughput pipelines for raw log and PCAP ingestion (NumPy, Scapy).
- Applied graph-based clustering techniques to identify anomalous traffic patterns in structured network datasets (NetworkX).
- Built interactive visualizations enabling rapid exploration of large graph datasets using optimized client-side rendering (D3.js, Cytoscape.js).

Software Engineer Intern | *ProofPerks, Remote*

Aug 2024 – Dec 2024

- Deployed transformer-based verification models as scalable, containerized services supporting high-volume inference workloads (Docker, GCP).
- Built cloud-native data pipelines to automate ingestion, storage, and processing of large biometric datasets (Cloud Functions, Cloud Storage, BigQuery).
- Reduced API latency and database load by implementing multi-layer caching for real-time verification workflows (Redis).

Projects

UStartKit | *React, TypeScript, Supabase, OpenAI, Amazon API*

- Built an AI app that automatically generates essential product lists for new hobbies.
- Engineered a real-time Amazon search system that organizes products into budget tiers.
- Implemented AI logic to tailor product recommendations and explanations to user profiles.

SiftAI | *Python, FastAPI, Node.js, Docker, WebAssembly, Gemini API*

- Built an analytics engine that converts data-related questions into executed visualizations and tables.
- Secured the runtime environment using a sandbox for safe code execution.
- Engineered a self-healing pipeline where the AI debugs and fixes runtime errors.

Elo-Conditioned Chess Predictor | *Python, PyTorch, Flask, JavaScript, NumPy*

- Trained a rating-conditioned CNN in PyTorch to mimic human play styles across Elo levels.
- Deployed a full-stack interface visualizing real-time move probabilities and AI confidence.
- Implemented robust server-side move validation and game state management logic.

Chess GM Twin | *React, Next.js, TypeScript, Tailwind CSS*

- Engineered a heuristics-based analysis engine to quantify chess playstyles (aggression, tactics) from raw game logs.
- Implemented a nearest-neighbor classification system to pair users with their historical Grandmaster counterparts.
- Built a Next.js dashboard visualizing comparative metrics through dynamic radar charts and game breakdowns.

Technical Skills

- **Big Data & Warehousing:** Apache Spark, Snowflake, Kafka, NoSQL (MongoDB), ETL, Data Warehousing
- **AI Data Infrastructure:** GraphQL, MLflow, LangChain, Model Deployment, ML pipelines, REST APIs
- **System Operations:** Kubernetes (K8s), Microsoft Azure, Linux/Unix, Shell Scripting, Cloud & DevOps