Aaron Gurovich

aargurov@ttu.edu • +1 949-505-0956 • aarongurovich.com

Education

Texas Tech University *Lubbock, TX*
B.S. in Computer Science, Minor in Mathematics *Expected: May 2026*

- **GPA:** 3.5
- **Relevant Coursework:** Software Engineering, Applied Data Science, Concepts of Programming Languages, Design and Analysis of Algorithms, Data Structures, Object-Oriented Programming, Computer Architecture.
- **Activities:** President of Hillel, Member of Chess Team (USCF: 1907).

Experience

Johnson & Johnson Electrophysiology *Irvine, CA*
Clinical Data Engineer Intern (Extended) *June 2025 – November 2025*

- **Architected** an **R Shiny** site selection tool for clinical trials using a custom **genetic algorithm**, automating optimal site selection to **accelerate trials**, **ensure diverse participant representation**, and **support FDA compliance**.
- Boosted a **Generative AI** pipeline's accuracy in classifying **protocol deviations** from **75% to 98%** by optimizing data ingestion in **Python**, applying advanced **prompt engineering**, and refining the **AWS** infrastructure.
- **Conceptualized** and developed an **AI-powered assistant** in **R** that converts **natural language** into optimized **SQL queries**, reducing costly AI ingestion and enabling **faster, more reliable insights** in **R Shiny dashboards**.
- **Transformed and ingested clinical device data** by refactoring legacy code into an **OOP Python** framework across multiple **Lambda functions**. Built **unit tests**, used **Bitbucket** for version control, and **Jenkins** for **CI/CD** tracking.
- **Co-authored** 3 abstracts based on this work, submitted to the **Johnson & Johnson Data Science Showcase**.

Texas Tech University *Lubbock, TX*
Research Assistant *Oct 2024 – May 2025*

- Developed a full-stack network analysis platform using **Flask**, **JavaScript**, **D3.js**, and **Cytoscape.js** for multi-layered, interactive cluster visualizations of complex network traffic data.
- Built a high-performance data pipeline with **Python** to process and enrich raw packet captures by engineering over 20 analytical features on **billions** of rows of data.
- Applied unsupervised **machine learning** (**Louvain Clustering**) to interactively discover anomalous traffic patterns in network data using **Python**'s **NetworkX** library.

ProofPerks *Remote*
Software Engineer Intern *Aug 2024 – Dec 2024*

- Fine-tuned transformer models from **Hugging Face** for race/ethnicity verification, improving authentication reliability, and deployed them as scalable microservices on **GCP**.
- Designed and implemented a cloud-native **data pipeline** on **GCP** to automate the ingestion, preprocessing, and augmentation of large-scale biometric datasets for model training and real-time inference.
- Implemented a multi-layer caching strategy with **Redis**, significantly reducing database load and decreasing API latency for real-time biometric verification services.

Projects

UStartKit `https://ustartkit.com/`

- Developed a full-stack web app using **React** and **TypeScript** to generate curated starter kits for a wide range of hobbies, enhancing user discovery and engagement.
- Integrated a **Deno** serverless backend using **Supabase** with **OpenAI GPT-3.5-turbo** and the **Amazon Product Advertising API** to dynamically source and recommend products.

ArgumateAI `https://react-argument-gpt.pages.dev/`

- Engineered a **React** frontend to record and transcribe user debates in real-time by directly integrating the **AssemblyAI** API.
- Leveraged **OpenAI GPT-4-turbo-preview** within the client-side application to provide impartial, AI-driven judging and feedback on the transcribed arguments.

ChessGMTwin `https://chessgmtwin.com/`

- Implemented a PGN parser and playstyle analyzer using **Next.js** to calculate stylistic metrics from chess games, matching users to Grandmasters based on their playstyle.
- Integrated the **Stockfish** chess engine to perform in-depth game analysis, processing moves to extract key tactical and strategic patterns.

Skills

Programming:	Python, R, Java, SQL, Pandas, Numpy, Scikit-learn, PySpark, Scapy, dplyr, ggplot2, tidyr
Cloud & Data:	AWS (S3, Lambda, SQS, Redshift, Glue), GCP, Azure, Docker, PostgreSQL, Cloudflare, Supabase
Frameworks:	Flask, React, Shiny, D3.js, REST APIs, Cytoscape.js, Node.js, Next.js
Dev Tools:	Git, Linux, Bash, CI/CD, GenAI, RStudio, VS Code, Tableau