

TOM TRAN JR

✉ tranjrtom@gmail.com

🌐 linkedin.com/in/tomtranjr

🐙 github.com/tomtranjr

🌐 tomtranjr.github.io

Education

University of San Francisco

San Francisco, CA

M.S. in Data Science

July 2025 – Expected June 2026

- Relevant Coursework: Deep Learning, Advanced Machine Learning, Linear Regression, Distributed Computing, MLOps

University of California, Los Angeles

Los Angeles, CA

B.S. in Applied Mathematics

June 2023

- Relevant Coursework: Linear Algebra, Probability, Statistics, Numerical Methods, Algorithms, Mathematical Modeling

Professional Experience

Gen AI Engineer

San Francisco, CA (Remote)

ArangoDB

Oct 2025 – Present

- Shipped **4 production features** to ArangoDB's **GraphRAG pipeline** for AI Suite 4.0, enabling domain-specific **knowledge graph construction** with multi-provider LLM support, flexible schema customization, and improved retrieval across **5 collection types**.
- Resolved **2 critical production bugs** and designed a **human-in-the-loop** interruption system validated across **20+ Kubernetes test cases**, reducing post-deploy investigation time by **35%**.
- Engineered **streaming vector search** for the open-source **langchain-arangodb** package, reducing memory overhead by **60%** on large-corpus retrieval workloads through unified batch and streaming query execution.
- Co-authored a retrieval benchmarking report with **NVIDIA** as an intern, validating **99%+ accuracy** over **226K+ documents** and outperforming **3 competing retrieval strategies** on complex technical queries.

Data Analyst

Washington, D.C. (Remote)

U.S. Department of State, Public Diplomacy & Public Affairs

Sept 2022 – June 2023

- Built a **Python ETL pipeline** to scrape, clean, and consolidate structured and unstructured data from **20+ sources**, generating **182+ country profiles** for a congressionally mandated State Department report, replacing weeks of manual research assistant effort.
- Analyzed **3+ years** of enterprise Zoom activity using **SQL and Python**, applying **NLP techniques** to meeting titles, descriptions, and chat logs that informed department-wide digital engagement strategies.
- Designed interactive **dashboards** in Google Sheets to monitor and forecast the **\$2M Zoom for Government** loan program, enabling **data-driven optimization** of annual license procurement and reducing excess spend.

Projects

canvas-control — Canvas LMS CLI & MCP Server

Mar 2026 – Present

- Architected an open-source **CLI and MCP server** exposing **15 AI-callable tools** for Canvas LMS, enabling natural-language course management in Claude Desktop and Cursor with async HTTP pagination, retry logic, and idempotent downloads backed by **17 test modules**.

DocQFlow — PDF Document Classifier

Apr 2026 – Present

- Building an end-to-end **MLOps pipeline** for AI-powered permit document processing with **TF-IDF + Logistic Regression** classifying incoming PDFs, then **OCR + LLM reasoning** extracts and validates structured fields, and served via **FastAPI**, tracked with **MLflow**, deployed on **GCP Cloud Run** with **Docker/GitHub Actions CI/CD**.

Leadership Experience

Aircraft Electrical and Environmental Systems Specialist

Mountain View, CA

129th Rescue Wing (129 RQW) | Active Secret Clearance

Oct 2023 – Present

- **Support mission readiness** of a fleet of **10+ aircraft** by analyzing electrical and environmental system data, ensuring operational safety for civil search and rescue, CAL FIRE, and counterdrug missions.
- **Collaborate with cross-functional teams** of pilots, engineers, and maintenance crews to diagnose and troubleshoot **15+ aircraft systems**, enabling rapid deployment in mission-critical situations.
- **Synthesize and communicate insights** from **30+ technical orders** into actionable guidance, ensuring compliance and alignment across diverse stakeholders.

Technical Skills

Languages: Python (PyTorch, Pandas, NumPy, Scikit-Learn, Plotly, Matplotlib, BeautifulSoup), SQL, PySpark, Bash

Tools: Git, GitHub, GCP, Docker, MLflow, FastAPI, GitHub Actions, Streamlit