

# Anthony J. Qin

a.j.qin@wustl.edu | (971) 275-7637 | [github.com/anthonyq7](https://github.com/anthonyq7) | [linkedin.com/in/anthony-qin-1947942b4/](https://linkedin.com/in/anthony-qin-1947942b4/)

## EDUCATION

**Washington University in St. Louis** | St. Louis, MO  
Computer Science, B.S.

Expected Grad. 2027  
GPA: 3.97/4.00

## TECHNICAL SKILLS

**Languages:** Python, C++, Java, Swift

**ML/AI:** PyTorch, Transformers, SBERT (Sentence Transformers), CodeBERT, LoRA (Low-Rank Adaption), Scikit-learn, NumPy, Pandas, RAG (Retrieval-Augmented Generation)

**Backend:** FastAPI, PostgreSQL, MongoDB, Redis, SQLAlchemy, Docker Containerization, REST APIs, Microservices, JSON, SQL

**Tools/Other:** Git (Version Control), GitHub Actions, Pytest, Ruff, Linux, CI/CD, Unit Testing, Data Structures, Algorithms, Agile Methodology, Cloud Deployment (Vercel and Render)

## EXPERIENCE

### Portal Career Discovery

Sep. 2025 - Present

*Software Engineer Intern*

- Built **SBERT**-based **RAG** microservice for resume-job matching using bi-encoder retrieval and cross-encoder re-ranking, reducing infrastructure costs by **90%** compared to GPT-based approaches while matching baseline relevance on an internal query-job test set.
- Automated data pipeline (ETL) processing **10,000+** monthly job postings using Python and MongoDB, reducing manual curation time from **~20 hours/week to zero**.

### McKelvey School of Engineering

Aug. 2025 - Present

*Teaching Assistant*

- Selected from top 10% of class to mentor **120+** students in proof-based discrete math (logic, sets, combinatorics, graph theory) via weekly office hours, increasing exam score averages by **5%** compared to previous semesters

## PROJECTS

### Aegis | Python, PyTorch, LoRA, Transformers

Oct. 2025 - Present

- Fine-tuned **CodeBERT** with **LoRA** for AI code detection, achieving **81% accuracy** (21pp improvement vs. zero-shot baseline) on 4,000 code samples (2K human-written from APPS dataset, 2K via OpenAI calls).
- Streamlined synthetic training dataset generation pipeline from days to **~3 hours** via async OpenAI calls.
- Packaged inference engine as CLI tool with adjustable classification threshold, achieving **0.94 recall** and **0.75 precision**.

### TriLive | FastAPI, PostgreSQL, SwiftUI

Jun. 2025 - Present

- Developed a live iOS transit tracking app for the Portland metro area to help commuters visualize ride status and receive arrival notifications in real-time, managing a beta program with **20** users via TestFlight.
- Added Redis caching for transit routes, reducing repeat-call latency by **97%**.
- Designed an MVVM mobile architecture capable of monthly data refreshes of **6,300+** transit stops into PostgreSQL.

## ACTIVITIES

### Google Developer Group WashU

Sep. 2025 - Present

*Project Lead*

- Leading cross-functional team in building a full-stack SaaS (Next.js + FastAPI) platform for media production logistics, establishing CI/CD best practices via GitHub Actions.
- Translating client requirements from Show Media Services into technical architecture specifications, resulting in a modular system design that handles scheduling and inventory management.

### WashU Robotics Club

Feb. 2025 - Sep. 2025

*Software Developer*

- Built RViz2 visualizer with color-coded pose vectors to speed up multi-robot debugging by **50%**.
- Explored Reinforcement Learning approaches for swarm robot path-planning in NVIDIA Omniverse.