

Quan A. Nguyen

(480) 862-4827 | Houston, TX, 77449 | quannnguyenanhnaq@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION & CERTIFICATION

Arizona State University Bachelor of Science in Computer Science

GPA: 3.7/4.0

Arizona State University Master of Science in Computer Engineering

GPA: 3.78/4.0

Relevant Coursework: Data Structures and Algorithms, Software QA and Testing, Distributed Software Development, Distributed Database, Operating System, Agile Scrum Methodologies, Microservices, Object-Oriented Programming (OOP)

EXPERIENCE

Next Lab

Oct 2025 – Present

Software Developer Associate – AI/ML

Tempe, AZ

- Fine-tuned and deployed Qwen3-1.7B with a **local RAG pipeline on Raspberry Pi 5** (8GB, Ubuntu Linux), supporting offline semantic search and conversational inference at **~6–18 tokens/sec** depending on quantization.
- Implemented vector retrieval and embedding indexing** for long-context Vietnamese datasets, optimizing chunking and retrieval to reduce latency and hallucinations under edge-device constraints.
- Developed a **Python-based ARM profiling framework** to measure token throughput, memory usage, and CPU temperature (**~55–75 °C sustained**), guided optimizations for long-prompt inference workloads.
- Benchmarked GGUF quantization formats (F16, Q8, Q6, Samsung LittleBit ultra-low-bit), achieving **~70–95% model size reduction** while analyzing accuracy degradation, latency, and thermal stability for low-power edge deployment.

Maricopa County Attorney Office

August 2024 – May 2025

Software Engineer Intern

Phoenix, AZ

- Led a greenfield application w/ a team of 5 interns, deploying a report-writing system using **React.js** and **ASP.NET Core**, backed by **SQL Server** database and hosted on-premises with **Windows Server** which streamlined reporting processes.
- Built **20+ RESTful API** endpoints w/ Swagger, integrated Open XML SDK and Aspose.PDF for templates and redaction.
- Used **EF Core** for initial data migrations, designed CI/CD pipelines w/ **Docker** and **Azure DevOps**, and followed Agile Scrum practices which allowed the project to complete **2 weeks ahead of schedule**.

Amazon Web Services Cloud Innovation Center

August 2023 – May 2025

Cloud Developer & Teaching Assistant

Tempe, AZ

- Taught AWS services (S3, Lambda, EC2, load balancers, Kubernetes) through Zoom and hands-on labs for **150+ students**.
- Introduced networking fundamentals: TCP/IP, TLS/SSL, cryptography, and web development basics (HTML, CSS).
- Automated grading workflows** w/ AWS Lambda, serverless scheduling, and Canvas API, **reduced workload by 95%**.

Crown Castle

June 2024 – August 2024

Software Engineer Intern

Houston, TX

- Redesigned the fiber network search platform w/ **Flask** and **Neo4j**, migrating from relational databases to graph databases.
- Optimized **ETL pipelines** processing **500K+ records** by batching and indexing, improving query performance by **47%**.
- Containerized and tested services w/ **Docker** and **GitLab CI**, integrating **Prometheus** and **Grafana** for performance monitoring, error tracking, and performed **Scrum Meeting** weekly for documentation on sprints and retrospectives.

PROJECTS

Fill – Automated Loan Data Entry | React 19, TypeScript, Vite 7, Chrome Manifest V3, Cursor, AntigraVity, GitHub

- Cut loan processing time by **20 minutes/version** for Precious Realty & Mortgage by scraping data with Google Extension.
- Reduced **context window and inference costs** by making custom Agent Skills to isolate tasks and minimize token usage.
- Eliminated **hallucination loops** and unnecessary dependencies by enforcing strict .cursor rules and workflows.

Medical Imaging ML Benchmark | Python, Pytorch, Pandas, NumPy, Linux, SLURM

- Trained **deep learning models** (UNet, ConvNeXt, Transformer) for medical image classification and segmentation.
- Utilized **SLURM on HPC Linux cluster** to run large-scale computations and track performance metrics.

TECHNICAL SKILLS

Languages: Python, C++, Java, JavaScript, C#, bash / shell, SQL, Cypher, TypeScript, HTML, CSS

Frameworks & Cloud: React.js, LangChain.js, Node.js, Neo4j, PostgreSQL, SQL Server, AWS, Microsoft Azure

Operating System & Tools: Linux, Unix, PowerShell, macOS, Git, GitHub, Docker, Postman, Microsoft Office

AI Automation: Cursor, AntigraVity, ChatGPT, Gemini 3, Agent Skills,