

MAYANK DEVANGKUMAR TAMAKUWALA

Boston, MA | +1-5108267414 | tamakuwala.m@northeastern.edu | [linkedin.com/in/mayanktamakuwala](https://www.linkedin.com/in/mayanktamakuwala) | github.com/MayankTamakuwala | mayank-tamakuwala.me

EDUCATION

Khoury College of Computer Science - Northeastern University

Jan 2025 - May 2027

Master of Science, Computer Science

Boston, MA

• **GPA:** 3.95/4.00

• **Coursework:** ML Research Assistant (AR/VR ML - Reality Design Labs), Teaching Assistant CS7180, Programming Design Paradigm, Natural Language Processing, Machine Learning Operations, Algorithms, Distributed Systems, Machine Learning

California State University Long Beach - Magna Cum Laude

Jan 2020 - Dec 2023

Bachelor of Science, Computer Science

Long Beach, CA

• **GPA:** 3.75/4.00

• **Achievements:** Magna Cum Laude, President's List, Dean's List, COE Excel Peer Mentor and Tutor

• **Coursework:** Data Structures & Algorithms, Machine Learning, Search Engine Technology, Database Fundamentals, OOP, Software Management and Testing, Advanced C++

EXPERIENCE

Amazon

May 2026 - Present

Software Development Engineer Intern

Irvine, CA

- Built the backbone infrastructure for an **agentic AI platform**, developing a greenfield **multi-tenant backend service** using **Java**, **AWS CDK**, and **DynamoDB**, orchestrating automated onboarding workflows across **3 distributed AI services**.
- Architected a **Model Context Protocol (MCP)** service layer, enabling **AI agents** to execute complex multi-step onboarding through natural language via **10+ agent-callable tools** with zero traditional UI dependency.
- Engineered a production-grade **multi-tenant security model** using **AWS KMS**, **SigV4 cross-account authentication**, and dynamically generated **IAM bucket policies**, enforcing secure tenant isolation across AWS account boundaries.
- Led an **architecture migration** from **AWS Bedrock AgentCore** to a serverless **CORAL Lambda + API Gateway + MCP Gateway** stack, presenting design trade-offs to senior engineering leadership while simplifying deployment and improving scalability.

Grossi Consulting LLC

Feb 2024 - Dec 2024

Full Stack Developer

Raleigh, NC

- Built and deployed **20+ production applications** using **React**, **TypeScript**, **Node.js**, **Next.js**, **GraphQL**, and **REST APIs**, delivering customer-facing features that improved client satisfaction by **30%**.
- Designed reusable **frontend components** and **backend services** using **React**, **Express.js**, and **Flask**, supporting responsive user experiences, workflow automation, and reliable communication across full-stack applications.
- Deployed and operated applications across **AWS** and **GCP**, configuring **S3**, **Elastic Load Balancers**, **Cloudflare**, **Git workflows**, and **CI/CD**, improving scalability and production reliability across **20+ client systems**.
- Instrumented customer experiences using **Google Analytics**, **Tag Manager**, and **Search Console**, translating product usage data into iterative improvements that increased user engagement by **20%**.

PROJECTS

Notion AI Workspace

May 2026 - Present

- Built an **AI-native document workspace** using **React**, **TypeScript**, **Node.js**, **PostgreSQL**, and **OpenAI APIs**, supporting **4 inline AI actions** -rewrite, summarize, expand, and extract action items, while implementing **document-level retrieval**, streaming responses, autosave, and evaluation tracking across **500+ seeded interactions** for latency, quality, and reliability.

HERMES - Hybrid Retrieval and Evaluation System [GitHub](#)

Mar 2025 - Mar 2025

- Designed a **multi-stage retrieval system** indexing **10k+ documents**, combining **dense embeddings**, **BM25 sparse search**, reranking, and automated **evaluation pipelines**, improving **Recall@10** by **25%** while serving low-latency **REST APIs** through separated offline indexing and online retrieval workflows with robust ranking instrumentation.

EdgeGuard - Distributed HTTP Gateway

Oct 2025 - Dec 2025

- Built a **high-performance HTTP gateway** using **Node.js**, **TypeScript**, **Redis**, and **Docker**, implementing **rate limiting**, **smart routing**, **distributed caching**, **asynchronous request handling**, **structured logging**, and end-to-end **observability**, achieving **sub-50 ms p95 latency** under load while improving backend reliability, scalability, and operational visibility.

SKILLS

- **Programming:** Python, TypeScript, JavaScript, Java, SQL, Data Structures, Algorithms, Git
- **AI & LLMs:** PyTorch, LangChain, RAG, Embeddings, AI Agents, Prompt Engineering, Model Evaluation
- **Product Engineering:** React, Node.js, FastAPI, REST APIs, GraphQL, Full-Stack Development
- **ML Infrastructure:** Data Pipelines, Model Deployment, CI/CD, Monitoring, Guardrails, Debugging
- **Databases & Search:** PostgreSQL, MongoDB, Redis, Weaviate, SQL, Vector Databases, Retrieval
- **Cloud & Systems:** AWS, GCP, GKE, Docker, Distributed Systems, Scalability, Observability