

# Aravind Balaji

Boston, MA | aravindbalaji.com | balaji.ara@northeastern.edu | linkedin.com/in/aravind-balaji-17a7b2115 | github.com/AravindB98

## PROFESSIONAL SUMMARY

---

MS Information Systems candidate at Northeastern University (GPA: 3.90) with hands-on experience building production-grade RAG pipelines, LLM fine-tuning workflows (QLoRA), and multi-agent AI systems. Author of QEMA-G, a quantum-enhanced memory framework for graph AI published on Zenodo and TechRxiv. VP of R&D at AI Skunkworks, recognised by Northeastern University Graduate Student Government. 4+ years prior production engineering experience, with a track record of translating complex AI capabilities into practical, scalable solutions.

## EDUCATION

---

**Northeastern University, Boston, MA** Expected Dec 2026

*MS Information Systems | GPA: 3.90 / 4.0*

*Relevant Coursework: Application Engineering & Development, Web Design & User Experience Engineering, Program Structure and Algorithms, Knowledge Graphs with GenAI/Graph DB, Data Management and Database Design, Prompt Engineering & AI, Business Process Engineering.*

**BITS Pilani, India** Jul 2021 - Jul 2023

*MTech Software Engineering | GPA: 3.49 / 4.0*

*Relevant Coursework: Cloud Computing, Software Architectures, Agile Software Processes, Software Project Management.*

**SRM Institute of Science & Technology, Chennai** Jul 2016 - Jul 2020

*BTech Electronics & Communication Engineering*

*Relevant Coursework: VLSI Design, Computer Architecture, Microprocessors and Microcontrollers, Digital Signal Processing, Embedded Systems.*

## TECHNICAL SKILLS

---

**AI & Machine Learning:** LLM Fine-Tuning (QLoRA/LoRA, TinyLlama), RAG (Retrieval-Augmented Generation), Multi-Agent Systems (LangGraph, CrewAI), NLP, Prompt Engineering, OpenAI API, Anthropic API, Azure OpenAI Service, Reinforcement Learning (UCB-1, REINFORCE)

**Data, BI & Analytics:** Python, Pandas, SQL, Exploratory Data Analysis, ETL/ELT Fundamentals, Databricks SQL, Snowflake, Power BI, Dashboard Development, Microsoft Office 365

**Vector & Graph Databases:** ChromaDB, PostgreSQL, MongoDB, Neo4j, Knowledge Graphs

**DevOps & Cloud:** Docker, AWS, Microsoft Azure, GCP, Git, VS Code, Agile SDLC

**Web Development:** MERN Stack (MongoDB, Express, React, Node.js), JavaScript, TypeScript, FastAPI, Streamlit, RESTful APIs, JWT Authentication, RBAC

## PROFESSIONAL EXPERIENCE

---

**HealthTech International** | Chennai, India | **Senior Software Engineer** Jul 2024 - Dec 2024

- Developed AI-assisted NLP workflows using Python to process, classify, and extract insights from large-scale structured healthcare documents
- Built automated data ingestion and reporting pipelines, reducing manual processing effort by 60%
- Researched and proposed architectural solutions for AI-powered, data-driven applications in a production healthcare environment
- Collaborated cross-functionally with product and data teams to translate business requirements into scalable technical implementations

**Futurenet Technologies** | Chennai, India | **Data Scientist Intern** Jun 2022 - Dec 2022

- Performed data analysis on 10,000+ daily system events using Python to surface trends and actionable insights
- Built Python-based analytical workflows and automated reporting pipelines, reducing time-to-insight for operations teams
- Assisted with cloud-based data pipelines and ML experimentation in Azure environments

**Checkpoint Systems** | Bengaluru, India | **Software Engineer Intern** Jun 2021 - May 2022

- Contributed to enterprise SaaS software supporting end-to-end retail supply chain processes, reducing customer-reported issues by 12%; participated in Agile sprint cycles with cross-functional teams

## ACADEMIC PROJECTS

---

### **CodeSentinel & RL-CodeSentinel: Multi-Agent Code Review System** | [github.com/AravindB98/RL-CodeSentinel](https://github.com/AravindB98/RL-CodeSentinel)

- Built a multi-agent RAG security analysis system using LangGraph directed graph with a 4-stage pipeline (query construction, semantic retrieval, lexical rerank, citation enforcement); reduced false positive rate from 0.789 to 0.111 compared to single-prompt baseline
- Extended the base system as RL-CodeSentinel by integrating **UCB-1 Contextual Bandit and REINFORCE Policy Gradient** for adaptive reviewer-agent selection that improves with feedback; grounded vulnerability detection in OWASP Top 10 2025 and CWE taxonomy

### **CEREBRO: AI Research Assistant (CrewAI + n8n Hybrid)** | [github.com/AravindB98/Cerebro](https://github.com/AravindB98/Cerebro)

- Engineered a 3-agent CrewAI system (Research Coordinator, Search Specialist, Synthesis Analyst) with a 940-line custom CEREBRO tool querying 16 academic databases in parallel; improved source relevance by +27 percentage points over web-only baseline
- Achieved 85% source relevance, 90% factual consistency, and 95% citation completeness across 45 evaluation runs; peer-validated with Cohen's kappa of 0.74 by independent Northeastern graduate evaluators

### **MediGraph AI: LLM-Powered Healthcare Analytics Platform** | [github.com/AravindB98/Projects/MediGraphAI](https://github.com/AravindB98/Projects/MediGraphAI)

- Built a production-grade RAG system to ingest, categorize, and extract insights from large-scale healthcare documents using ChromaDB for semantic search; developed LLM-powered agents to compare documents, identify trends, and generate summaries with SQL-backed evaluation datasets

### **Care Hub: Secure Distributed Web System** | [github.com/AravindB98/Projects/CareHub](https://github.com/AravindB98/Projects/CareHub)

- Developed a MERN-stack application with JWT authentication and RBAC; designed APIs and backend services supporting future system-level automation and optimization

## PUBLICATION

---

### **Quantum-Enhanced Memory Architectures for Graph-Based AI Systems: A Theoretical Framework with Feasibility Analysis**

Aravind Balaji, Nik Bear Brown | Northeastern University | Published on [TechRxiv](https://techrxiv.org/) & [Zenodo](https://zenodo.org/)

- Designed a novel AI memory architecture improving how ML models store, retrieve, and reason over large-scale connected datasets; developed and evaluated ML-driven graph embedding techniques compressing high-dimensional data for efficient AI retrieval, with performance benchmarks and feasibility analysis for real-world deployment

## LEADERSHIP & RESEARCH EXPERIENCE

---

### **Vice President of Research & Development, AI Skunkworks - Northeastern University**

*Graduate AI Research Organization | Faculty Advisor: Prof. Nik Bear Brown*

- Rebuilt the graduate AI research organization from the ground up; authored the constitution, navigated GSG reinstatement processes, and secured formal **recognition by Northeastern University Graduate Student Government** under faculty advisor Prof. Nik Bear Brown
- Contributed to AI-based Citation Checker and Generative AI-driven clinical documentation projects, focusing on data validation and ethical AI in healthcare; mentor graduate cohorts through hands-on research and cross-functional collaboration

### **1st Place, MGEN Hackathon 2025 - Northeastern University (DNATE MSL Practice Gym)**

- Built React + Node.js systems integrated with OpenAI APIs for AI-powered analytics features, collaborating with cross-functional teams under time pressure to deliver a scalable, user-focused solution

## ADDITIONAL INFORMATION

---

- Available May 4, 2026 - December 25, 2026 for full-time co-op / internship (40 hours/week); eligible to work full-time in the U.S. without sponsorship
- Technical writer at [aravindbalaji1.substack.com](https://aravindbalaji1.substack.com) and [substack.com/@aravindbalaji2](https://substack.com/@aravindbalaji2) (Northeastern ISE publication); also produce independent music with AI tools