

# Aryasomayajula Ram Bharadwaj

ram.bharadwaj.arya@gmail.com | +91-9108832338 | Bengaluru, India | github.com/rokosbasilisk

## ABOUT

---

Versatile ML Engineer and Backend Developer with expertise in deep learning and strong DevOps knowledge. Skilled in multiple programming languages and refactoring legacy codebases, passionate about AI research.

## EDUCATION

---

**B.Tech Electronics and Communications**  
*GMR Institute of Technology, Andhra Pradesh*

2015 - 2019

## PROFESSIONAL EXPERIENCE

---

**AI Resident - Lossfunk AI Residency**  
*Lossfunk AI Residency (Remote)*

April 2025 - May 2025

- Selected for Lossfunk, a 6-week intensive AI residency with 10 highly skilled builders and researchers.
- Developed STU-PID (Steering Token Usage with PID Control), a novel activation steering technique to reduce redundant reasoning tokens in LLMs.
- Conducted experiments on reasoning models like DeepSeek-R1-Distill-Qwen-1.5B with dynamic activation interventions, achieving 32% token reduction and improved reasoning accuracy on grade school math problems.
- A detailed paper on this is accessible at [Steering Token Usage with PID Control](#).

**Associate Technical Architect - MLOps**  
*Quantiphi Analytics, Bengaluru*

Nov 2024 - Present

- Designed and implemented an AI agent system for automated replies for a major telecom company's sales chatbot.
- Experimented with LangGraph agent architectures to improve controllability and modularity in conversational flows.
- Revamped and automated the scheduled RAG data ingestion pipeline and Improved retrieval accuracy by 20%.
- Improved retriever latency by several folds via extensive code refactoring.

**Senior Developer at Innovation & Development Labs**  
*Musigma Business Solutions, Bengaluru*

June 2019 - Nov 2024

- Led development of multiple high-impact projects in LLM operationalization, automated trading, and MLOps.
- Specialized in backend development, DevOps, and ML engineering across various domains.
- Received multiple recognitions including Impact Awards and Star Performer of the Team.

## KEY PROJECTS

---

**LLM Agent Platform**

Dec 2023 - Nov 2024

*Team Lead, ML Engineer, Backend Developer*

- Designed and implemented a semi-autonomous data analysis platform powered by LLM Agents.
- Developed training strategies for LLM Agents via automatic prompt optimization and integrated Retrieval-Augmented Generation (RAG) support.
- Implemented evaluation for LLM Agents' results using an ensemble of locally hosted LLMs.
- Maintained a customized fork of Microsoft's Autogen framework.
- Successfully deployed for two major clients, assisted client teams in creating use cases.

## High Velocity Trading Platform

2021 - Dec 2023

*Team Lead, Backend Developer, DevOps*

- Refactored backend code and migrated deployment from bare-metal to Kubernetes.
- Implemented automated hyperparameter search for ARIMA models.
- Rewrote legacy trade-signal generation from R to Scala using Akka framework.
- Enabled near real-time metric calculation and portfolio visualization.

## Next-Gen Model Operationalization Platform

2019 - 2021

*Backend Developer*

- Developed automatic retraining pipelines for image classification models.
- Designed and implemented Java microservices for creating and serving Jupyter notebooks.
- Engineered a microservice for deploying ML models in a DAG structure with canary and blue-green deployment strategies.

## RESEARCH PAPERS

---

### Understanding Hidden Computations in Transformer Language Models

August 2024

*Independent Researcher*

- Conducted research to decode the "hidden" computations in transformer models during Chain of Thought (COT) reasoning.
- Investigated methods to enhance the faithfulness and internal mechanics of chain-of-thought reasoning in LLMs.

## PERSONAL PROJECTS

---

### Wiserank.io

June 2025

*Creator, Developer*

- Built 'Wiserank.io', a research paper search engine, to rank and retrieve research papers based on relevance and novelty.
- Implemented a feature to automatically retrieve and generate relevant citations for an uploaded research manuscript.

## SKILLS

---

**Programming Languages:** Scala, Python, Java, R

**Machine Learning Frameworks:** PyTorch, JAX

## AWARDS & RECOGNITIONS

---

### AI Alignment Awards

July 2023

[www.alignmentawards.com](http://www.alignmentawards.com)

Winning entry for the "goal misgeneralization" track research proposal, selected among 118 entries in this research competition.

### Honorable Mention

Mar 2022

*Alignment Research Center*

"Eliciting Latent Knowledge" is an open research problem in AI Safety. My proposal was one of the honorable mentions in this research competition.

### SPOT and Impact awards

2019 - 2023

*Musigma Business Solutions*

Received six SPOT (Star Performer) awards and two Impact Awards at Musigma for exceptional contributions, DevOps skills, and LLM-Agent platform design.

### Bronze Position

Feb 2021

*Build-on-Redis Hackathon*

Winning entry in the Build-on-Redis hackathon. Implemented a text2code search tool entirely on Redis Stack using CodeBERT model's embeddings for a private code repository.