$\bigcirc$  Bellevue, WA  $\checkmark$  +1 (206) 731-9171  $\square$  <u>rishirm@uw.edu</u>  $\square$  linkedin.com/in/rishimullur

## Summary

Software Engineer with 4+ years of experience delivering robust, scalable systems across AI, productivity, and data infrastructure domains. Former founder and AI/ML engineer. Proven ability to ship high-impact features. Experienced in designing AI-enabled systems under real-world constraints, with cross-functional collaboration across product, engineering, and research teams.

## Experience

Software Engineer	April 2025 – Current
Kanary Inc	Bellevue, WA
• Engineered intelligent matching pipelines for high-volume data systems, cutting false positives by 50%.	
• Built LLM-powered browser automation tools for web scraping and data extraction, improving data collect	tion efficiency.
<ul> <li>Led algorithmic improvements using heuristics and statistical tuning, increasing engagement and accuracy.</li> </ul>	
0	June 2024 – December 2024
2ndChair AI	Seattle, WA
<ul> <li>Designed and scaled modular RAG pipelines using Python, FastAPI, and Hugging Face for legal document</li> <li>Implemented performance-optimized parsing pipelines with SQL and async processing, improving through</li> <li>Built agent-based task orchestration with memory and reflexion modules, reducing redundant computation</li> <li>Created LLM evaluation and debugging tools for token-level latency and accuracy tracking in production.</li> </ul>	but by 45%.
Software Engineer, Research - Collaborator	July 2024 – March 2025
Allen Institute for Artificial Intelligence - Capstone Project	Seattle, WA
<ul><li>Engineered Python-based data fusion pipelines for LiDAR, IMU, and imaging on Raspberry Pi 5.</li><li>Enabled forestry field research with ML-ready sensor pipelines and local inference of LLaMA 8B.</li></ul>	,
• Created templates and evaluation datasets for multi-sensor learning and CV model training.	
Lead Software Engineer/CTO	Jan. 2021 – Mar. 2022
Scryorb Technologies <ul> <li>Architected distributed backend systems with Go and GCP to support real-time short-video ingestion and</li> </ul>	Bangalore, India
<ul> <li>Architected distributed backend systems with Go and GCP to support real-time short-video ingestion and</li> <li>Built analytics pipelines for creator activity and engagement using PostgreSQL and Firebase event streams</li> <li>Implemented JWT-based auth gateway; enforced least-privilege IAM on GCP.</li> </ul>	
Software Security Researcher	Mar. $2022 - Dec. 2022$
Freelance	Bangalore, India
• Ranked #1 globally in Twitter Bug Bounty 2022; discovered critical security flaws in production apps used	
Associate Product Manager (Technical)	Jan. 2023 – Aug 2023
<ul> <li>ClearFeed Inc.</li> <li>Drove development and launch of GPT-3.5-powered support automation, reducing response time by 50% a through collaboration with engineering and analytics teams.</li> </ul>	Bangalore, India nd boosting ARR by 55%
	June 2019 – December 2020
Wipro Technologies	Bangalore, India
• Engineered and optimized 5G core network protocols (PFCP, NGAP, NAS) in C++ for telecom, ensuring	
high-throughput performance.	low latency and
• Collaborated in a multi-team agile environment to design modular, reusable components aligned with 3GP	P specs.
<ul> <li>Developed test harnesses in Python for simulating mobile core scenarios and validating protocol behavior u</li> <li>Documented and presented technical design specifications to internal review boards and contributed to cro sharing.</li> </ul>	inder edge-case loads.
Education	
University of Washington, Global Innovation Exchange	Seattle, WA
Master of Science in Computer Science [Technology Innovation]	Sep. 2023 – March 2025
(Courses: Deep Learning, Programming for Digital Interfaces, Planning & Managing Hardware/Software Dev)	
National Institute of Technology, Warangal	Warangal, India
Bachelor of Technology in Mechanical Engineering	Aug. 2015 – May 2019
Projects	11ag. 2010 may 2010

ClaudePlaysPokémon: Human-AI Gameplay System | Creator & Presenter

• Enabled human coaching through live text input, demonstrating collaborative decision-making between LLM and human user.

• Presented AI agent which plays Pokemon Red demo and architecture overview to 150+ attendees at exclusive AI event, showcasing agentic control, strategy abstraction, and emergent behavior.

• Featured as main stage talk at AI Tinkerers Seattle, highlighting innovative human-AI collaboration techniques.

Decoding Prompting for LLMs (Research) | Poster, CSE599 Advanced CS Topics

- Benchmarked GPT-4, Claude, GPT-3.5 on adversarial tasks (30+ examples) to evaluate prompt strategies.
- Published best practices based on user study across reasoning prompt types.

## Skills

Languages: Python, C++, Go, JavaScript, SQL, R AI/ML: PyTorch, Transformers (HF), OpenAI/Anthropic APIs, RAG, Prompt Engineering, LLM Evaluation Systems: REST APIs, Data Pipelines, Docker, CI/CD Cloud: GCP (Vertex AI), AWS (S3, EC2), Firebase Infra: Load Balancing, Caching, API Rate Limiting, WebSockets
Testing: Pytest, Unit Testing, Integration Testing
Databases: PostgreSQL, MongoDB, Redis
Messaging: Kafka, Pub/Sub, Celery
Monitoring: Grafana, Sentry

2025

2024