

Rahul Atul Bhope

Sr. Applied Scientist

☎ (+1) 949-562-3750
✉ rbhope@uci.edu
🌐 rahulatulbhope.github.io
in rahul-bhope-3a0976130
📄 rahulatulbhope

Education

- 2021–2026 **PhD in Networked Systems (Computer Science)**, *University of California, Irvine*
Thesis: Middleware Systems for Efficient and Robust Distributed Machine Learning
Advisor: Dr. Nalini Venkatasubramanian. Focus: LLMs, Agentic AI, Federated Learning
- 2016–2020 **B.Tech in Instrumentation & Control Engineering, Minor in CS**, *NIT Tiruchirappalli, India*

Professional Experience

- 2026–Present **Sr Applied Scientist**, *Microsoft M365, Mountain View, CA*
○ Joined Microsoft's Copilot Memory team, advancing retrieval and memory systems for large-scale AI assistants.
- 2021–2026 **Graduate Research Assistant**, *University of California, Irvine*
○ Lead research on adaptive federated AI: (1) shift-aware training with latent matching for model selection, (2) inference memoization via semantic caches for reduced inference latency in agentic AI systems
○ Developed **ShiftEx**, a mixture-of-experts framework for Federated Learning under covariate/label shift, reducing post shift recovery time by 22-95% and improving accuracy by 5.5-12.9% points
○ Architected **CareDex** a WiFi-based localization platform.. Worked with Kafka/RabbitMQ-based real-time ingestion and sensor health heartbeat monitoring; deployed at **Walnut Village** senior living facility
- Summer 2025, 2024, 2022 **Research Scientist Intern**, *IBM Research, Yorktown Heights, NY*
○ Designed **VOILA**, adaptive fidelity selection for multimodal AI systems achieving 90-95% accuracy at 40-43% cost for edge-cloud, IoT and agentic memeory deployments.
○ Developed **OptiSeq** improving LLM-based in-context example ordering for API prediction accuracy by 5.5-10.5% points
○ Developed **FLIPS** for intelligent participant selection in Federated Learning, reducing communication costs by 20-60% and improving accuracy by 17-20% points
- 2020–2021 **Project Associate**, *Indian Institute of Science, Bangalore*
○ Benchmarked Federated Learning on Jetson clusters; deployed **GoCoronaGo** contact-tracing across 4 campuses, improving BLE proximity inference by 20%
- Summer 2019 **Research Intern**, *HESL Lab, Nanyang Technological University, Singapore*
○ Built **CampusNavNTU**, indoor navigation using WiFi fingerprinting and IMU tracking

Selected Publications

- Under Review **VOILA: Value-of-Information Guided Fidelity Selection for Cost-Aware Multimodal Question Answering**. R. A. Bhope, K. R. Jayaram, V. Muthusamy, R. Kumar, V. Isahagian, N. Venkatasubramanian.
- Middleware'25 **Shift Happens: Mixture of Experts based Continual Adaptation in Federated Learning**. R. A. Bhope, K. R. Jayaram, P. Venkateswaran, N. Venkatasubramanian. *ACM/IFIP Middleware 2025*.
- Middleware'25-DS **Middleware Systems for Efficient and Robust Distributed Machine Learning**. R. A. Bhope. *ACM/IFIP Middleware 2025 Doctoral Symposium*.
- Middleware'25 Demos **In-Context Example Ordering for LLM-Based API Sequence Generation**. R. A. Bhope, P. Venkateswaran, K. R. Jayaram, V. Isahagian, V. Muthusamy, N. Venkatasubramanian. *ACM/IFIP Middleware 2025 Demos*.
- EMNLP'25 **OptiSeq: Ordering Examples On-The-Fly for In-Context Learning**. R. A. Bhope, P. Venkateswaran, K. R. Jayaram, V. Isahagian, V. Muthusamy, N. Venkatasubramanian. *EMNLP 2025 Findings*.
- PVLDB'25 **DIM-SUM: Dynamic Imputation Methods for Smart Utility Management**. R. Hildebrant, R. A. Bhope, S. Mehrotra, C. Tull, N. Venkatasubramanian. *PVLDB 2025*.
- Middleware'23 CCS'25-Poster **FLIPS: Federated Learning using Intelligent Participant Selection**. R. A. Bhope, K. R. Jayaram, N. Venkatasubramanian, A. Verma, G. Thomas. *ACM Middleware 2023, Confidential Computing Summit 2025*
- IPDPSW'22 **Don't Miss the Train: A Case for Systems Research into Training on the Edge**. Prashanthi S. K., A. Khochare, S. A. Kesanapalli, R. Bhope, Y. Simmhan. *IPDPS Workshops 2022*.

Patents (Filed)

- P202300541US01 Federated Learning Participant Selection Through Label Distribution Clustering
P202402468US01 Optimizing In-Context Example Ordering for API Sequence Generation using LLMs
P202502973US01 Storage and Memory Management for Multimodal AI Systems

Skills & Leadership

- Programming Python, C#, Java, Shell Scripting, HTML/CSS, Embedded C
Technologies Transformers, PyTorch, TensorFlow, Kafka, Spring Boot, Docker, Kubernetes, AWS, Google Cloud
Teaching TA for Grad level Operating Systems, Distributed Systems; Mentored REU & Capstone students