Rajat Ghosh, Ph.D.

+1 (404) 697-5789 · rajat.ghosh11@gmail.com · linkedin.com/in/i-am-rajat

Professional Summary

Senior technical leader building **company-scale** GenAI platforms from research to production. Led LLM agents, post-training, alignment pipelines, evaluation/safety frameworks, and AI infrastructure optimization, delivering **multi-million-dollar impact** and measurable productivity gains. Known for setting long-term technical **vision in high-ambiguity**, frontier AI domains (agentic systems, RAG, reasoning models). Built and scaled **high-performing teams** and mentored talent to industry-leading outcomes. Trusted **cross-functional partner** across Product, Infra, Legal, and Exec leadership. Active contributor to the **ML community** with publications and open-source impact.

Professional Experience

Technical Director, AI Nutanix, *September 2025 - Present*

As Technical Director, I lead the post-training, evaluation, and advanced inference product portfolio for the Nutanix GenAI platform, driving quality, reliability, and readiness for production.

- Led a team of 7 engineers to build LLM post-training and evaluation pipelines used across Nutanix GenAI.
- Owned production pipeline for internal developer-facing AI agents, including reliability and evaluation.
- Designed resource-efficient post-training, agentic productivity apps, and dataset curation.

Staff Data Scientist Nutanix, *May 2021 - August 2025*

As Tech Lead Data Scientist, I spearheaded the design and delivery of two production-ready generative AI agents for customer service and developer productivity. I defined the architectures, built the initial implementations, and established rigorous evaluation pipelines to validate quality and reliability.

- Architected internal coding assistant agents utilizing open-source LLMs and advanced agentic patterns such as RAG, tool-calling, Reflexion ($100M annual impact from 20% improvement in developer productivity).
- Developed customer support assistants using open source LLMs and retrieval augmented generation (RAG) with open-source VectorDB ($20M annual impact from from support engineering productivity).
- Engineered reinforcement learning strategies for distributed infrastructure management including Kubernetes, Cassandra, storage tiering, data warehousing, and load balancing.

Co-Founder & CEO AdeptDC, Inc. *July 2015 – April 2021*

I founded a company to commercialize my PhD research in reinforcement learning and data center energy efficiency.

- Built and led a 7-member team delivering AI-driven optimization solutions to enterprises.
- Secured $1M in venture and NSF funding; exited via acquisition.

Education

- **Ph.D.**, Georgia Institute of Technology, Atlanta, 2013
- **M.S.**, Georgia Institute of Technology, Atlanta, 2010

- **B.S.**, Indian Institute of Technology, Kharagpur, 2008

Honors and Awards

- NSF Small Business Innovation Research Award, 2015 (Link)
- Georgia Research Alliance Phase-IA Award, 2014
- NSF I-Corps Award, 2014

Selected Recent Publications on LLM

Research contributions spanning LLM post-training, evaluation, and alignment.

- D. Nimmaturi, V. Bhargava, **R. Ghosh**, J. George, and D. Dutta, *Predictive Scaling Laws for Efficient GRPO Training of Large Reasoning Models*, arXiv:2507.18014, 2025. (Link)
- A. Singhal, **R. Ghosh**, R. Mundra, H. Dadlani, and D. Dutta, *Code2JSON: Can a Zero-Shot LLM Extract Code Features for Code RAG?*, ICLR 2025 Workshop on Deep Learning for Code. (Link)
- V. Bhargava, **R. Ghosh**, and D. Dutta, *CPP-UT-Bench: Can LLMs Write Complex Unit Tests in C++?*, NeurIPS Statistics in LLMs Workshop, 2024. (Link)
- A. Khera, **R. Ghosh**, and D. Dutta, *Efficient Alignment of Large Language Models via Data Sampling*, NeurIPS ENLSP Workshop (PMLR), 2024. (Link)

Consortium Publications from MLCommons

Research contributions in LLM Safety research.

- One of 82 consortium authors. Contributed to the evaluation pipeline development, *AILuminate: Introducing v1.0 of the AI Risk and Reliability Benchmark from MLCommons.* (Link)
- One of 100 consortium authors. Contributed to the dataset curation, *Introducing v0.5 of the AI Safety Benchmark from MLCommons.* (Link)

Selected Publications from Doctoral Research

Research contributions spanning optimal control, design of experiments, and data center energy efficiency.

- **R. Ghosh** and Y. Joshi, *Proper Orthogonal Decomposition-Based Modeling Framework for Improving Spatial Resolution of Measured Temperature Data*, IEEE Transactions on Components, Packaging and Manufacturing Technology, 2014.
- **R. Ghosh** and Y. Joshi, *Rapid Temperature Predictions in Data Centers Using Multi-Parameter Proper Orthogonal Decomposition*, Numerical Heat Transfer, Part A: Applications, 2014.
- **R. Ghosh** and Y. Joshi, *Error Estimation in POD-Based Dynamic Reduced-Order Thermal Modeling of Data Centers*, International Journal of Heat and Mass Transfer, 2013.
- **R. Ghosh**, G. A. Buxton, O. B. Usta, A. C. Balazs, and A. Alexeev, *Designing Oscillating Cilia That Capture or Release Microscopic Particles*, Langmuir, 2010.

Patents

- **R. Ghosh** and Y. K. Joshi, *Systems and Methods for Intelligent Controls for Optimal Resource Allocation for Data Center Operations*, US Patent 10,439,912, Oct. 2019. (Link)

Invited Talks

- Agentic AI: Future Challenges and Opportunities, AAAI Workshop, March 2025. (Link)

Teaching Experience

- Thermodynamics (ME 3332), Georgia Tech, 2013
- Experimental Methodology (ME 3057), Georgia Tech, 2012–2013
- Introduction to Energy Systems Engineering (ME 4803), Georgia Tech, 2012

- Creative Decisions and Design (ME 2110), Georgia Tech, 2009

Review & Service

Reviewer for NeurIPS, AAAI, IEEE CPMT Journal, InterPACK, Semi-Therm, I-Therm.