Sai Koti Reddy Danda

AI Research and Engineering Leadership

Location: Bengaluru, India
Email: d.saikotireddy@gmail.com
LinkedIn: linkedin.com/in/saikotireddy
Website: saikotireddy.com
Scholar: Google Scholar profile

Profile

AI research and engineering leader with **16 years in AI research** and a record of building enterprise AI products from the ground up in two startups. I have led **multiple products and distributed teams across three geographies**, setting technical direction with founders, executives, and customers. Products I led are used by **Fortune 500 enterprises** including NVIDIA, JPMorgan Chase, CVS Health, Kroger, and T-Mobile, and banking products I led from inception are in proof-of-concept engagements with **SMBC**.

- **Product portfolio:** Originated and led AI products spanning banking process intelligence, agentic AI, Kubernetes autoscaling, GPU orchestration, resource optimization, and intelligent load balancing
- **Teams and organization:** Managed **three AI teams across India, Romania, and Canada**; recruited, onboarded, and mentored engineers; built research partnerships with IIT Bombay and IIIT Bangalore
- **Outcomes: More than 30% lower cloud costs** through reinforcement learning (RL) driven autoscaling, and **3x LLM inference throughput** through GPU orchestration
- **Research and IP:** Co-wrote **two funded five-year research grants** (IBM Research Math Council; Department of Science & Technology, India and Robert Bosch Centre, IISc); peer-reviewed publications at AAMAS, IEEE CDC, and ECAI; contributions to multiple U.S. patents and patent applications; IBM awards for research and client value

Research depth in multi-agent systems, agentic AI, large language models (LLMs), and reinforcement learning (including constrained and risk-averse multi-agent RL), with foundations in game theory and stochastic optimization.

Experience

kAIgentic Inc
Bengaluru, India

Principal Architect – Agentic AI for banking process transformation
Dec 2025 – Present

Leadership and Scope

- Lead AI research and architecture across kAIgentic's banking product suite from early ideation; set technical direction across teams with product, technology, and customer success leaders, turning customer priorities into research and engineering decisions

Key Contributions

- Originated products for process mapping, process transformation, and agentic AI, translating enterprise banking problems into research priorities, prototypes, and product capabilities
- Architected multi-agent process intelligence that combines multiple information sources, banking ontologies, and knowledge graphs for process mapping and analysis

Impact

- Products I have led from inception are in proof-of-concept engagements with **SMBC** for banking process transformation

Avesha Inc
Bengaluru, India

Principal AI Architect – Employee #5, joined at seed stage
Mar 2021 – Dec 2025

Leadership and Scope

- Led AI research and product creation across a portfolio of **five products**, taking each from concept through proof of concept and production in enterprise environments
- Managed **three AI teams across India, Romania, and Canada**; recruited, onboarded, and mentored engineers; established the company's U.S. AI presence after its Series B
- Led all AI initiatives, working directly with the CEO and executive leadership on product strategy, customer acquisition, and technical direction
- Initiated and led university research collaborations with IIT Bombay and IIIT Bangalore

Impact

- RL-driven autoscaling (Smart Scaler) reduced cloud costs by **more than 30%**; GPU orchestration (Elastic GPU Service) increased LLM inference throughput **3x**
- Products used by **Fortune 500 enterprises** including NVIDIA, JPMorgan Chase, CVS Health, Kroger, T-Mobile, and Supermicro, and by customers across telecom, media, security, cloud infrastructure, and financial services including Cox, Akamai, Nutanix, Amagi, McAfee, and Finvi; technology partners included NVIDIA, Nutanix, and Oracle
- Products scaled across 100+ customers, 56+ clusters, and 10,000+ applications
- Contributed inventions in workload placement (granted U.S. patent), adaptive autoscaling, and RL-based global load balancing (patent filings)

Products

- **Smart Scaler** (RL-based Kubernetes autoscaling), **Elastic GPU Service** (multi-cloud GPU orchestration), **Avesha Obliq** (agentic SRE and QA automation), **SmartSizing** (container resource optimization), and **Smart Application Load Balancer** (RL-based service-mesh routing)

IBM Research India — Bengaluru, India

Research Engineer – AI for Supply Chains Group — Jul 2017 – Feb 2021

Key Contributions

- Co-wrote the five-year IBM Research Math Council grant *Making Game Theory Practical for Decision-Making in Multi-Agent Systems*, and led game-theoretic research in constrained and risk-averse multi-agent reinforcement learning under it
- Core contributor to AI-based analytics for Sterling Commerce Multi-Enterprise Business Networks, automating decisions in multi-party supply chains
- Led development of the IPDC supply chain and dealer finance blockchain platform and contributed to blockchain solutions for Bajaj Auto and M&M Finance
- IBM Research Spotlight Award (2020), Client Value Outstanding Technical Achievement Award (2019), and Research Accomplishment Award (2018)

Robert Bosch Centre for Cyber-Physical Systems, IISc — Bengaluru, India

Research Assistant – Stochastic Systems Lab, Prof. Shalabh Bhatnagar — Dec 2016 – Jun 2017

- Co-wrote the five-year DST and Robert Bosch Centre grant *Distributed Multi-Agent Algorithms for Dynamic Control of Microgrids*; researched distributed reinforcement learning for microgrid control

IBM Research India — Bengaluru, India

Research Intern – Mathematical Modeling and Industrial Solutions Group — May – Nov 2016

- Built the initial blockchain invoice-discounting proof of concept for Mahindra & Mahindra; it led to IBM securing the full project, the first blockchain project in India

Education

Indian Institute of Science — Bengaluru, India

M.Sc. (Engg.) in Computer Science, Advisor: Prof. Shalabh Bhatnagar — 2014 – 2016

NRI Institute of Technology — Guntur, India

B.Tech. in Information Technology, Advisor: Dr. Viswanath Pulabaigari — 2008 – 2012

Expertise

- **Leadership:** Product ideation and strategy, technical direction, team building and mentoring, executive and customer engagement, research partnerships
- **AI:** Multi-agent systems, agentic AI, LLMs, reinforcement learning, knowledge graphs and ontologies, stochastic optimization, game theory
- **Platforms:** Kubernetes, GPU orchestration, multi-cloud infrastructure, Hyperledger Fabric
- **Engineering:** Python, Go, PyTorch, Neo4j, MongoDB

Awards

- IBM Research Spotlight Award — 2020
- IBM Client Value Outstanding Technical Achievement Award — 2019
- IBM Research Accomplishment Award — 2018
- Top 1% of 155,190 candidates in GATE — 2014

Patents

- Controlling Placement of Workloads of an Application within an Application Environment — US11805073B2
- Data Quality Control — US11169976B2

Plus four published patent applications and two filings in autoscaling, RL-based load balancing, multi-agent AI, and blockchain; full list on saikotireddy.com/research

Selected Publications

- Truthful and Equitable Lateral Transshipment in Multi-Retailer Systems — ECAI 2023
- Inventory Pooling using Deep Reinforcement Learning — IEEE SCC 2022
- Attention Actor-Critic Algorithm for Multi-Agent Constrained Co-operative RL (joint first author) — AAMAS 2021
- Risk Averse Reinforcement Learning for Mixed Multi-Agent Environments (first author) — AAMAS 2019
- Actor-Critic Algorithms for Constrained Multi-Agent RL (joint first author) — AAMAS 2019

14 publications in total; full list on saikotireddy.com/research and Google Scholar.