

Santosh Varma

AI Architect | Enterprise AI, Voice & RAG Platforms

santoshvarma4u@gmail.com | +91 7702393981 | LinkedIn: Santosh Varma | Hyderabad, India

Executive Profile

- A results-driven engineering leader with **14+ years of experience** in full-stack development, distributed systems, and scalable cloud architecture, currently serving as an **AI Architect at Kore.ai**. Specializes in designing enterprise AI platforms with a focus on **AI-enabled Voice Gateways, Retrieval-Augmented Generation (RAG)**, and agentic systems.
- Experienced in **AI engineering**, including the design and deployment of LLM-powered systems, agent-based architectures, and intelligent automation workflows. Skilled in building pipelines for RAG (vector databases, hybrid retrieval, re-ranking), local and cloud-hosted LLM integrations, and custom micro-agents capable of autonomous decision-making and task orchestration.
- Hands-on with **Kore.ai's enterprise AI suite** — including the **XO Platform, AgentAssist, SmartAssist, SearchAssist**, and **Voice AI** capabilities — embedding conversational and generative AI within real-world products to enhance productivity and customer experience.
- Previously led architecture at **Storable Auctions** and engineering / QA at **Savor Brands**, driving end-to-end delivery of high-performance enterprise and mobile applications. Expert in evaluating As-Is and To-Be processes and architecting scalable solutions aligned with strategic goals.
- Directed technology planning, strategy development, leadership and implementation, business and operational transformation execution, and business solution delivery.
- Technically proficient across **JavaScript / TypeScript, Node.js, React, Python, AWS, DevOps, MySQL / PostgreSQL, microservices, event streaming, Unity**, and modern AI stacks. Recognized for fostering engineering excellence and mentoring teams toward strong execution and innovation.

Education

- Bachelor of Technology:** Computer Science from JNTU Kakinada — Bhimavaram in 2012

Technical Skills

- JavaScript / TypeScript
- Node.js
- React
- Mobile Development (Native & Hybrid)
- AI & LLM Engineering
- RAG & Vector Databases
- MySQL / PG SQL / NoSQL
- Python
- AWS / DevOps / K8s
- Voice AI (ASR, TTS, SIP/WebRTC)
- Kore.ai XO, AgentAssist
- Unity

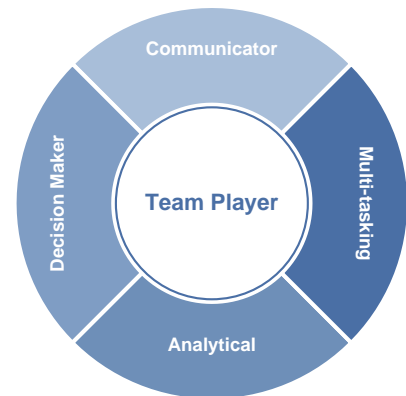
Achievements

- At **Kore.ai**, contributing to AI architecture for enterprise voice gateways, RAG pipelines, and agentic experiences across the **XO Platform** and **AgentAssist**.
- At **Storable Auctions**, contributed to architecture modernization, data integrations, and improvements to core auction workflows through scalable and event-driven design.
- Contributed to major feature development and scalability improvements in **Savor Brands'** multi-party white-label eCommerce platform.
- Built AI-powered agent solutions for **OPSTECH** and **TGQuest** using LLMs, RAG, and autonomous micro-agents to automate workflows and enhance decision-making.

Core Competencies

Software Development Lifecycle	Cloud Computing
Agile Methodologies	AI / LLM Architecture
Database Management	RAG & Vector Search
Project Management	Voice AI Gateways
Machine Learning	Artificial Intelligence

Soft Skills



Professional Experience

AI Architect | Kore.ai Software India Pvt. Ltd., Hyderabad, IN | Apr 2026 — Present

Kore.ai is a global leader in enterprise conversational and generative AI, providing a no-code AI platform — the XO Platform, AgentAssist, SmartAssist, SearchAssist, and Voice AI — that powers virtual assistants and agentic workflows for Fortune 500 enterprises across banking, healthcare, retail, and telecom.

- **Architecting AI-enabled voice gateways** on the Kore.ai platform — integrating ASR, TTS, SIP/WebRTC telephony, real-time streaming, and LLM-driven intent resolution to power enterprise voice agents and contact-center automation.
- **Designing and optimizing Retrieval-Augmented Generation (RAG) pipelines** — semantic chunking, embedding strategy, hybrid retrieval, re-ranking, and grounded-response evaluation — across enterprise knowledge bases on vector stores such as Milvus, Weaviate, and Vespa.
- Defining reference architectures across the **Kore.ai XO Platform, AgentAssist, SmartAssist, SearchAssist, and Voice AI** products, covering tenant isolation, model routing across hosted and private LLM endpoints, and observability for production AI workloads.
- Driving architectural reviews for **agentic AI workflows** — tool-use orchestration, policy-based action authorization, conversation state management, and runtime observability for autonomous agents.
- Collaborating with product and platform teams to extend the **XO Platform** with new generative AI capabilities — knowledge-graph grounding, dialog automation, and LLM-assisted dialog design.
- Partnering with engineering leadership on AI evaluation frameworks, model selection, latency optimization, and production readiness of generative AI features across BFSI, healthcare, and retail customers.
- Mentoring engineering teams on AI architecture patterns, LLM application design, and production best practices for conversational and voice AI systems.

Staff Software Engineer | Storable, Hyderabad, IN | Nov 2024 — Apr 2026

- **Architected scalable and resilient solutions for the Storable Auctions platform**, ensuring high availability, seamless bidder experience, and improved system reliability across core auction workflows.
- **Leading modernization of legacy components into microservices and event-driven systems**, resulting in improved system performance, faster data synchronization, and reduced operational overhead.
- **Driving the redesign of critical integrations between Sitelink (FMS) and Auctions**, enhancing data accuracy, reducing sync failures, and improving end-to-end visibility of unit and facility data.
- Partnering with cross-functional teams across engineering, product, and DevOps to deliver architecture improvements that meet business goals for scalability, performance, and platform stability.
- **Implementing structured system design processes and code quality standards**, improving engineering efficiency and reducing defects across auction-related services.
- **Enhancing the migration architecture for beta and production customers**, resolving login and onboarding issues while ensuring a seamless transition to the new Auctions platform.
- Contributing to next-generation auction capabilities, including async workflows and high-throughput APIs, resulting in a more robust and scalable auction ecosystem.

Senior Fullstack Engineer | Savor Brands, Honolulu, Hawaii, US | May 2021 — Nov 2024

- Partnering with software development and testing team members to create effective solutions that meet client needs for functionality, scalability, and performance.
- Supervising technologists on a \$10M software development project, ensuring the upgrade of the quality management system.
- Spearheading the design and deployment of a fully-featured video-based patient-doctor consultation mobile application.
- Leading the development of a new user authentication system, resulting in increased user engagement and a decrease in login-related issues.
- Implementing a streamlined code review process, reducing the time taken for code reviews and improving overall code quality.
- Orchestrating the implementation of innovative features, enhancing the multi-party white-label eCommerce application for Savor Brands, increasing user engagement and driving a boost in sales.

Senior Software Engineer | Skillsoft, Hyderabad, IN | Feb 2018 — May 2021

- Exhibited hands-on programming experience in Java, JavaScript, React Native, and Node.js.
- Collaborated with cross-functional teams to resolve technical issues and deliver high-quality software solutions.
- Worked as subject matter expert and primary point-of-contact for project management staff, leading critical software development initiatives.
- Played a pivotal role in the successful development and launch of core features for the Ed-tech app **Percipio**, contributing to its widespread adoption and positive user feedback.

Senior System Analyst | Mahathi Software Pvt. Ltd., Vizag, IN | Aug 2016 — Feb 2018

- Developed comprehensive troubleshooting documentation and distributed it to team members and key stakeholders.
- Employed various methods to troubleshoot software and hardware issues, identifying areas for improvement and implementing new software solutions.
- Escalated project-related issues, ensuring timely resolution and preventing delays in project timelines.
- Implemented process improvements that enhanced the efficiency of troubleshooting software and hardware issues, resulting in a **20% reduction in issue resolution time**.

Software Engineer | Schemax Expert Techno Crafts Pvt. Ltd., Vizag, IN | Apr 2012 — Jul 2016

- Collaborated with software development and testing teams to design and develop robust solutions meeting client requirements for functionality, scalability, and performance.
- Reviewed project specifications and designed technology solutions that exceeded performance expectations.
- Coordinated with engineers to evaluate and enhance software and hardware interfaces, ensuring seamless integration and optimal performance.

- Led the evaluation and enhancement of software and hardware interfaces, resulting in a **30% improvement in system performance and reliability**.

Internship

Intern / Trainee Engineer | Genex Software Solutions, Vizag, IN | Jan 2012 — Apr 2012

- Engaged in hands-on training and practical application of software development methodologies and best practices.
- Assisted in the evaluation and testing of proposed technical solutions based on customer requirements.

Personal Details

Date of Birth: 22 / Aug / 1991

Nationality: Indian

Languages Known: English

Address: Hyderabad, India