NAGARJUNA K

Lead AI Architect | AI Solutions Architect | GenAI Lead — Agentic AI & RAG

Currently in Dubai, UAE · Open to relocation · Requires work visa sponsorship · Available immediately
+971 50 742 4361 +91 93913 47406 gmailtonagarjuna@gmail.com
linkedin.com/in/nagarjuna-kamireddy github.com/nagarjunak-pixel

PROFILE

Lead AI architect with **20+ years in enterprise data and AI**: a decade of production ML and GenAI (John Deere, Persistent, Vertiv), built on a decade embedded at client sites as an SAP BW / BPC / HANA consultant across **India, Malaysia, Singapore, the UK and the US**. Most recently owned architecture and led a 12+ member team for **23+ AI initiatives at John Deere — 750K+ automated requests and 42K+ hours saved in FY26**. Hands-on with agentic AI, RAG, LLM evaluation and AWS. **AWS Certified Solutions Architect – Professional** · **NVIDIA-Certified Professional: Agentic AI**.

EXPERIENCE

John Deere — Pune, India
AI Architect & Technical Lead (official title: Senior Data Scientist) — **01/2023 – 10/2026**
Analytics, Automation & Process Improvement — Customer & Product Support

- **Owned architecture, AWS platform and delivery** for 23+ AI initiatives, working directly with technical writers, engineers and support teams to find the problem, prove the value and take each solution from pilot to production — FY26: 750K+ automated requests, 42K+ hours saved.
- **Built StoneAge, an AI writing copilot inside the company's engineering system of record:** authors draft and check service manuals with AI that searches approved content, picks the best model for each task (GPT-5, GPT-4o, Gemini) and flags safety-critical wording automatically. Stack: LangGraph agents, hybrid search on OpenSearch, React, FastAPI, Okta.
- **Turned around a failing AI product:** a single-model design was only ~30% accurate; rebuilt it as a step-by-step agent workflow that reads engineering diagrams with GPT-4o vision, routes by product family and checks every output against rules — accuracy rose to ~70%, now used by 46 people.
- **Connected the company's knowledge for AI:** made engineering manuals searchable without losing their structure, and designed a 6-agent authoring team backed by a Neo4j knowledge graph so every answer cites its source.
- **Made AI quality measurable:** built AIFlow, an in-house library of 255 quality checks that blocks releases when answers get worse, plus live monitoring of accuracy, drift, cost and response time — with no slowdown for users.
- **Kept AI safe in production:** agents can only change official records after a fresh user approval, and every attempt is logged; chose search-plus-rules over model fine-tuning so content stays auditable as products change.
- **Automated high-volume work:** repaired ~171K broken links (~8,600 hours, ~3 full-time staff a year); generated ~110K standards-compliant documents (100% adoption in the Large Tractor unit); automated ~25% of support tickets end to end with ~83% accuracy, sending uncertain cases to people (~4,700 hours saved).
- **Platform and leadership:** built a self-service AWS deployment platform (Terraform, GitHub Actions) adopted department-wide and by three partner organisations, and owned AWS cost; led a 12+ member team, ran design reviews and hiring panels, mentored 3 engineers; sat on the AI Governance Council (NIST AI RMF / EU AI Act) deciding which AI projects to scale, fix or retire; led 3 years of digital transformation enabling 300+ employees.

Persistent Systems — Pune, India
Engineering Lead — Machine Learning — **01/2021 – 01/2023**

- Delivered end-to-end client ML projects — development, optimisation, static and dynamic deployment, MLOps and secure database connections — for PIAM access recommendation, iDataCleanse identity de-duplication, face recognition, VMS recommendation, VectorFlow POD and Vertina NLP; mentored engineers across the portfolio.

Vertiv — Hyderabad, India
Data Scientist — **03/2020 – 01/2021**

- Built ML / DL models on production data for compliance analytics and Customer 360, pushing filtering and aggregation down to AWS Athena; delivered Tableau dashboards for business decisions.

Soham Global — India
Software Engineer — Machine Learning — **03/2016 – 02/2020**

- Built account-balance and price-tag forecasting with leakage-safe time-based validation, plus market-basket analysis, recommendation and inventory tracking for finance and e-commerce clients.

Earlier career — enterprise data & analytics consulting, 2004 – 2016 (India, Malaysia, Singapore, UK, US)

GlaxoSmithKline (GSK) — Brentford, United Kingdom
HANA Lead Consultant — Data Platform & Predictive Analytics — **02/2014 – 02/2016**

- **In-database machine learning (my move into ML):** built regression, random forest, SVM, clustering, ARIMA / Holt-Winters forecasting and churn-scoring models inside HANA (PAL, APL, R); moved MRP and financial consolidation batch runs from hours to minutes.
- **GSK India & HCL Axon UK** — HANA Consultant, 2012 – 2014: turned finance KPIs into data models and specifications with business users.
- **PETRONAS Energy Trading & Deloitte**, Kuala Lumpur — HANA / BPC Architect, 2011 – 2012: proof of concept cut a data load from **48 hours to 12 minutes** and won an internal rollout.
- **DHL & Symantec**, Malaysia · Singapore — BPC / BusinessObjects Consultant, 2008 – 2011: financial consolidation and a billing data warehouse.
- **Lenovo & High Industries**, Chennai · United States — BW / BI Consultant, 2004 – 2008: CRM sales analytics data models.

SELECTED PROJECTS — github.com/nagarjunak-pixel · 2,249 contributions in the last year

CompanyOS: 7 executive AI agents plan, assign and track a company's work from one config file — Kanban, live logs, KPI engine (Python).
Surya Shiksha: agentic learning platform with 12 GPT-4o agents for curriculum, tutoring and assessment (Next.js, PostgreSQL, Prisma).
LLM Knowledge Base: turns raw sources into an LLM-maintained wiki with full-text search and Q&A (FastAPI, SQLite FTS5, Claude).
LangGraph LocalAI TV: 3-stage news production pipeline with human approval gates, checkpointing and error recovery (LangGraph).

CORE SKILLS

Generative AI & Agents: Generative AI · LLMs · agentic workflows · multi-agent orchestration · LangGraph · LangChain · LlamaIndex · CrewAI · OpenAI Agents SDK · MCP · tool use / function calling · structured outputs (Pydantic) · prompt & context engineering · multi-model routing (GPT-5, GPT-4o, Gemini) · vision LLMs
Retrieval & Knowledge: RAG · hybrid search · GraphRAG · reranking · embeddings · OpenSearch · FAISS · pgvector · Neo4j · knowledge graphs · citations
ML: Python · PyTorch · TensorFlow · Hugging Face · scikit-learn · XGBoost · NLP · computer vision · forecasting · recommendation
Evaluation, LLMOps & Responsible AI: LLM evaluation · golden datasets · LLM-as-judge · RAGAS · CI regression gates · monitoring (accuracy, drift, cost, latency) · MLOps · guardrails · human-in-the-loop · NIST AI RMF · EU AI Act
Cloud & Engineering: AWS (ECS Fargate, Lambda, S3, DynamoDB, Athena, IAM) · Azure OpenAI · Terraform · GitHub Actions · Docker · FastAPI · React · SSE · SQL · REST APIs · Okta / OAuth 2.0 · Tableau
Enterprise Systems: SAP (HANA, BW, BPC, Data Services) · BusinessObjects · Windchill PLM APIs
Leadership: AI strategy & roadmap · solution architecture · PoC-to-production · ROI & value tracking · AI governance · stakeholder management · team leadership · hiring · mentoring

CERTIFICATIONS, RECOGNITION & EDUCATION

Certifications: AWS Certified Solutions Architect – Professional (SAP-C02) · NVIDIA-Certified Professional: Agentic AI (NCP-AAI)
Training: Anthropic — Advanced Prompt Engineering, MCP (Advanced) · OpenAI — Agents & Workflows
Recognition: Enterprise SME for ML & GenAI · AI Governance Council member · BI&A 2026 showcase
Education: M.Sc. (Integrated) Applied Science, IT — Periyar University · 2004 · First Class | **Languages:** English, Hindi, Telugu, Tamil