KUMAR GYANAM   Chief AI Architect | AI Strategy Leader

📍 Gurgaon |📞9953682525|✉ gyanamc@gmail.com|🔗 LinkedIn Profile | 🌐 Live
Portfolio: gyanam.store

STOP READING. START INTERVIEWING. *"I built an Agentic AI bot to deprecate my own traditional resume. Instead of reading, if you would like to talk to my resume, then please click"*   👉 CLICK HERE

EXECUTIVE SUMMARY

Transformational AI Leader with **~20 years of experience** architecting enterprise-scale AI systems across fintech, healthcare, and media. Proven expert in building and deploying **sovereign, on-premise GenAI infrastructure** within highly regulated environments. Trusted advisor to C-level stakeholders with board-approved AI roadmaps, managing full-cycle AI transformation, multi-million dollar budgets, and vendor ecosystems. **Guest Faculty at IIM and ISB.**

TECHNICAL ARCHITECTURE EXPERTISE

- **Agentic AI & Orchestration:** LangGraph (Stateful Agents), LangChain, n8n, Agentic RAG Workflows.

- **On-Premise Generative AI Stack:** Ray Serve (Distributed Serving), vLLM (Inference Optimization), ELK Stack (Vector Search), Open Source LLMs (Llama-3, Mistral, OSS).

- **MLOps & Observability: Ragas** (RAG Evaluation), **MLflow** (Experiment Tracking & Registry), Airflow (Pipeline Orchestration), Prometheus/Grafana.

- **Cloud & Platforms:** Railway.app (Expert—30+ deployments), AWS, Firebase, Docker, Kubernetes (K8s).

- **Data & Automation:** n8n.io (Expert—30+ workflows), Python (scikit-learn, TensorFlow), FAST APIs, SQL/NoSQL.

PROFESSIONAL EXPERIENCE

SBI Card | Deputy Vice President - Conversational AI   *April 2023 – Present*

Strategic AI Infrastructure & Platform Leadership

- **Sovereign On-Premise Agentic RAG Platform:** Architected and operationalised a secure, air-gapped Retrieval-Augmented Generation (RAG) ecosystem, ensuring 100% data sovereignty and financial-grade security.

- **Orchestration & Logic:** Leveraged **LangChain** and **LangGraph** to build complex, stateful agentic workflows, enabling multi-turn reasoning and tool-integrated AI capabilities.

- **High-Performance Stack:** Integrated **45+ enterprise components** including **Ray Serve** for distributed serving and **vLLM** for high-throughput inference, managed via **Airflow** for robust data pipeline orchestration.

- **Infrastructure Depth:** Deployed open-source LLMs (**Mistral, Llama-3**) using **ELK (Elasticsearch) as a high-scale Vector Database** for efficient semantic retrieval.

- **Continuous Evaluation:** Established a rigorous evaluation loop using **Ragas** (for metric-based assessment of faithfulness/relevancy) and **MLflow** for end-to-end experiment tracking and versioning.

- **Multi-Modal Strategic Implementations:** Utilized the on-premise RAG engine to deliver high-impact business applications:

  1. **Credit Card Recommendation Engine:** Real-time discovery engine tailored to customer financial profiles and spending behavior.

  2. **Generative Agent-Assist:** Real-time reply generation for live chat agents, improving response accuracy and reducing average handle time (AHT).

  3. **SBI Card Internal GPT:** A secure, internal-only portal for employees to interact with massive knowledge bases and internal documentation.

  4. **Self-Healing & HITL Architecture:** Engineered a **Human-in-the-Loop (HITL)** framework to allow real-time correction of model misses or biases, ensuring reliability.

- **Platform Scale & Governance:** Scaled the conversational platform to support **22M+ users** and **21M+ monthly transactions** with a **99.9% uptime SLA**.

  - Authored the board-approved AI roadmap, driving a **400% growth in bill payments** via AI channels and leading enterprise-wide AI procurement and vendor evaluation.

Brane Group | Analytics Delivery Head   *January 2022 – April 2023*

- Architected ML-powered solutions for Healthcare (ambulance dispatch optimisation), Financial Risk (fraud detection), and Ad-tech (campaign ROI optimisation).

- Led a cross-functional team of data scientists and engineers to deliver production-ready AI models for global clients.

Pixstory | Head of Technology Architecture   *January 2019 – January 2022*

- Built a custom ML-based recommendation engine from scratch, resulting in a **180% improvement in user engagement**.

- Led technical product strategy and executed go-to-market strategies supporting 100K+ user acquisition.

- Managed the full-stack development of semantic analysis and automated tagging workflows for 1M+ posts.

The Indian Express | Chief Manager - Business Intelligence   *July 2009 – January 2019*

- Transformed subscription architecture and built predictive analytics for collection management.

- Reduced user detractors by **50%** through text mining and continuous consumer research.

Axis Bank | Sales Manager   August 2008 - July 2009

Managed distributed sales operation achieving Rs 200 Cr+ disbursement targets through process optimisation and team leadership.

ICICI Bank (Sales Manager) | June 2006 - August 2008:

Built franchise network across multiple districts managing Rs 60 Cr+ lending portfolio and regional expansion strategy.

EDUCATION & ACADEMIC LEADERSHIP

- **Academic Leadership:** Faculty at **IIM-Indore** and **ISB** for AI Strategy, Technology Leadership, and AI Governance.

- **Indian Institute of Technology (IIT), Madras:** Advanced Diploma in Programming (2022–2025, Dropped Out).

- **Indian Institute of Management (IIM), Calcutta:** AI-Powered Marketing (2021–2022).

- **Institute of Rural Management (IRM):** MBA in Marketing & Strategic Management (2004–2006).

- **Ranchi University:** B.Sc. in Mathematics & Statistics (2001–2004).

KEY ACHIEVEMENTS & STRATEGIC IMPACT

- **Production Excellence:** Maintains a live portfolio at **gyanam.store** featuring 30+ production AI systems.

- **Operational Transformation:** Achieved **75+ FTE cost savings** through live chat integration and AI-driven automation.

- **Thought Leadership:** Recognised expert in **Railway.app** and **n8n.io** for enterprise-scale AI automation.

- **Governance:** Expert in fintech AI governance, regulatory compliance, and large-scale system security audits.