

PRASHANT YADAV

Lead AI & Full Stack Engineer | Agentic AI Specialist | LangGraph | RAG | Multi-Agent Systems| Python | Next.js | React.js | AWS | System Design | 11+ Years

Pune, India | prasahnt.yadav567@gmail.com | 7665737211 | [LinkedIn](#) | [Github](#)

SUMMARY

Lead AI & Full Stack Engineer with 11+ years of experience designing and delivering enterprise-scale cloud-native platforms across Banking, FinTech, E-Commerce, and Manufacturing domains.

Specialized in Agentic AI, Multi-Agent Systems, Hybrid RAG Architectures, AWS Cloud Platforms, and Distributed Systems. Currently leading the architecture and implementation of an enterprise AI platform serving 500+ engineering leaders through Microsoft Teams using LangGraph, FastAPI, PostgreSQL, pgvector, and AWS Bedrock.

TECHNICAL SKILLS

AI & GenAI: LangGraph, LangChain, Multi-Agent Systems, RAG, Hybrid Search, MCP, LangSmith, RAGAS, Bedrock, OpenAI, Anthropic

Backend: Python, FastAPI, Node.js, Express.js, REST APIs

Frontend: React.js, Next.js, TypeScript, JavaScript, Redux, HTML5, CSS3

Cloud & DevOps: AWS, ECS, Lambda, S3, SQS, RDS, Terraform, Docker, Jenkins, GitHub Actions,

Databases: MongoDB, PostgreSQL, Redis, Pinecone, FAISS, Chroma DB

Architecture & Design: Solution Architecture, System Design, Microservices, Event-Driven Architecture, Distributed Systems, Platform Engineering

PROFESSIONAL EXPERIENCE

Coforge | Tech Lead

Oct 2021 - Present

Enterprise Delivery Intelligence Platform (Agentic AI)

- Leading architecture and delivery of an enterprise Agentic AI platform serving 500+ engineering leaders across 50+ teams.
- Designed LangGraph-based multi-agent workflows with source verification and citation mechanisms.
- Architected hybrid RAG using PostgreSQL + pgvector across Jira, GitHub, and Confluence data sources.
- Built event-driven ingestion pipelines using AWS SQS, S3, Redis, and webhooks.
- Established RBAC, audit logging, observability, and security controls for enterprise adoption.
- Reduced manual delivery investigation effort by approximately 50%.

Tech Stack: Python, FastAPI, LangGraph, PostgreSQL, pgvector, Redis, AWS ECS, SQS, S3, Terraform

AI Shopping Assistant & Visual Product Discovery

- Architected a multimodal AI assistant enabling technicians to identify tools and vehicle parts through natural language and image uploads.
- Developed a visual search workflow that matched photos of broken tools or missing vehicle components against a product catalog using image embeddings and vector search.
- Built a RAG-powered recommendation engine using LangChain, Pinecone, and AWS Bedrock to provide technical guidance and compatible replacement recommendations.
- Increased recommendation conversion by 15% through AI-powered search and personalized product discovery.

Tech Stack: Python, FastAPI, LangChain, Pinecone, AWS Bedrock, HuggingFace, React, Next.js

MatcoTools – E-Commerce & Financing Platform

- Designed and scaled an e-commerce platform supporting 150000+ technicians and 2000+ distributors across USA and Canada.
- Developed financing and EMI management workflows that increased financed purchases by 25%.
- Optimized critical purchasing APIs reducing checkout latency from 2.5s to 800ms.
- Architected scalable microservices and GraphQL APIs supporting customer onboarding, product catalog management, financing approvals, and payment processing workflows.

Tech Stack: GraphQL, React, Next.js, TypeScript, MongoDB, AWS, Microservices, Jenkins

Wipro, Sr. Project Engineer

Feb 2019 - Oct 2021

Project: HSBC Bank

- Developed enterprise loan management modules for investors, suppliers, and buyers using React, Node.js, and microservices architecture.
- Built and integrated REST APIs supporting loan approval workflows, reducing processing time by 30%.
- Optimized application performance through frontend and API enhancements, improving response times by 25%.
- Contributed to AWS-based deployment, monitoring, and production support activities.

Tech Stack: React, Node.js, Express.js, JavaScript, REST APIs, PostgreSQL, AWS, Microservices, Git

Habilelabs, Software Engineer

June 2015 - Feb 2019

Project: Polls Guru (Survey Platform) & Hiretale (Job Portal)

- Led development of cloud-based polling, analytics, and recruitment platforms serving enterprise users.
- Built real-time dashboards and reporting solutions, reducing processing effort by 60%.
- Improved user engagement by 25% and reduced page load times by 40% through frontend and backend optimization.
- Implemented AWS-based deployments and CI/CD pipelines supporting 99.9% platform availability.
- Collaborated with product owners and senior engineers to deliver scalable features and performance improvements.

Tech Stack: React, JavaScript, Node.js, Express.js, MongoDB, REST APIs, AWS

KEY ACHIEVEMENTS

- AWS Certified Developer Associate
- Google Cloud Certified Associate Cloud Engineer
- NASSCOM - National Association of Software and Service Companies
- Generative AI with Large Language Models from DeepLearning.AI and AWS
- Winner of appathon(hack day codefest), 24 hours app development event

EDUCATION

Jaipur National University, Jaipur Rajasthan
Bachelor of Technology (Computer Science)

Jun 2011 - May 2015