

Deepak Kushwaha

AI Architect & Engineering Leader — Scaling AI-Native Products from 0 to Millions

Bangalore, India +91-9901549150 kdeepakkushwaha@gmail.com deepakkushwaha.tech

linkedin.com/in/deepak8896484842

Accomplished **AI Architect & Engineering Leader** with **9+ years** building and scaling high-impact SaaS platforms. Founder/co-founder of multiple ventures, having architected and shipped **20+ products** serving millions globally. Deep expertise across the modern AI stack — RAG pipelines, agentic systems, LLM evaluation, and responsible AI — combined with full-stack, cloud, and DevOps mastery. Proven track record of leading **35+ person teams** and aligning AI technical strategy with business growth.

AI & ML SKILLS

LLM INTEGRATION

OpenAI / GPT-4, Gemini, Claude, Mistral, LLaMA

AI ARCHITECTURE

RAG pipelines, Vector DBs (Pinecone, Weaviate, pgvector), LangChain, LlamaIndex, Multi-agent systems

EVAL & OBSERVABILITY

LangSmith, Helicone, Braintrust, A/B testing AI outputs, hallucination mitigation

AI INFRA

vLLM, token cost optimization, fine-tuning workflows, embedding strategies

RESPONSIBLE AI

Bias auditing, PII handling, EU AI Act awareness, AI governance frameworks

DEV TOOLING

Cursor, GitHub Copilot, prompt engineering, chain-of-thought, few-shot

ENGINEERING

ARCHITECTURE

Microservices, Design Patterns, High-Scale Infra, Sharding, Database Design

BACKEND

Python, Django, Node.js, REST, GraphQL

EXPERIENCE

The Branding Club

May 2024 – Present

Head of Engineering & AI Architect · Netherlands, Europe (Remote)

- Established and scaled a **20-member cross-functional team** (Engineering, QA, DevOps) from zero, driving hiring pipelines, engineering culture, and technical standards.
- Architected **Brandhub** — a microservices SaaS platform processing **100,000 read/write ops/second** and generating 100,000 SKUs per batch using multiprocessing, multithreaded queuing, and sharding.
- Designed an **AI-powered content and SKU generation layer** using RAG pipelines and LLM APIs (GPT-4, Claude), cutting manual content effort by **70%**.
- Introduced LLM evaluation and observability tooling (Helicone, LangSmith) to monitor output quality, reduce hallucinations, and optimize token costs across all AI workloads.

RAG LangChain GPT-4 API Claude API Microservices AWS Kafka
Kubernetes

Ceodekho

Dec 2023 – Jun 2024

Co-founder & CTO Concurrent venture

- Built an AI-native SaaS blogging platform using **LLM pipelines and agentic content workflows**, reducing content production time by **80%**.
- Achieved **PageSpeed 99+ on desktop, 90+ on mobile** via Next.js SSR, edge caching, and Technical SEO optimization.
- Designed multi-tenant architecture with cost-optimized token management strategies for AI content generation at scale.

LLM Agents OpenAI API Next.js Technical SEO Multi-tenant SaaS

Khaatadekho

Nov 2023 – Jun 2024

Founder Concurrent venture

- Conceptualized and built a **mini-Salesforce CRM platform** from scratch, owning product research, backend, and frontend architecture end-to-end.
- Secured **10+ stable clients** with average monthly revenue of ₹80,000; integrated AI-assisted lead scoring using LLM APIs, reducing manual CRM data entry by **60%**.

CRM SaaS LLM APIs Python React 0-to-1 Product

FRONTEND

React, Next.js, Svelte,
TypeScript, Flutter, Technical
SEO

CLOUD & DEVOPS

AWS, Azure, Kubernetes,
Kafka, CI/CD, Docker

LEADERSHIP

Team building (35+
employees), Tech hiring, Cross-
functional leadership, Budget
management, OKR planning,
Board-level communication

AWARDS

Best Performer (Engineering) ×2

Turno · Sep & Mar 2022

Super Coder

Tally Solutions · Aug 2021

Spot Award + Rookie Award

Tally Solutions · 2020–21

Google India Challenge Scholarship

Google · Feb 2018

Int'l Science Olympiad AIR 38

2012

EDUCATION

B.Tech, Computer Science

Visvesvaraya Technological
University (VTU) · Apr 2019

LANGUAGES

English (Professional) · Hindi
(Native)

Intellectyx

Nov 2023 – May 2024

Technical Lead — AI & SaaS Consultant · USA (Remote)

- Led **4 cross-functional teams** architecting scalable SaaS products for US-based enterprise clients.
- Designed a Udemey-style platform with an **AI-powered course recommendation engine** using vector embeddings, improving learner engagement by **40%**.
- Increased engineering productivity by **35%** through AI-assisted development tooling, contributing to **\$2M in additional client revenue**.

Vector Embeddings AI Recommendations Python Node.js SaaS Architecture

Turno

May 2022 – Dec 2023

Technical Lead — Founding Team

- Led partnership initiative with **Facebook and Google**, engineering a dynamic website deployment app that launched **20** tailored landing pages, driving conversions from **1% to 25%**.
- Integrated AI-driven personalization into the sales funnel using LLM-assisted copy generation, significantly reducing CAC.
- Built and owned the full-stack Agent App and EWS for internal and field sales teams, overseeing backend, frontend, and DevOps end-to-end.

AI Personalization React Node.js DevOps CI/CD

Wizcommerce (Sourcewiz)

Sep 2021 – Jun 2022

Senior Software Engineer — Founding Team

- Built an onboarding automation tool that compressed customer onboarding from **10 days to 1 day**.
- Led a 3–4 member team to build Sourcewiz's Customer App from scratch across design, frontend, backend, and deployment.

B2B SaaS React Python Process Automation

Tally Solutions Pvt Ltd

Aug 2019 – Sep 2021

Software Development Engineer I & II

- Authored code design for **TallyPrime 1.x** and GST web components; developed UI and backend APIs independently from architecture through production.
- Led R&D on integrating WhatsApp and Payment Gateway with Tally — an early conversational automation initiative.

TallyPrime GST API Design Conversational Automation

PixelAI.in

Jan 2017 – Mar 2020

Co-founder Side venture alongside Tally Solutions

- Co-founded an AI & software consultancy delivering **10+ SaaS products** for Indian and US-based clients, specializing in digital performance and operational efficiency.