

PROFESSIONAL SUMMARY

Full-Stack and AI Developer with a proven track record of architecting highly scalable applications, spanning high-performance websites and cross-platform mobile apps. I build from first principles, leveraging microservices, robust DevOps pipelines, and rock-solid software engineering to drive measurable impact. As a National Hackathon winner, India Innovates 2026 Finalist, and technical mentor to 200+ engineering students, I seamlessly bridge the gap between traditional development and the bleeding edge - designing and deploying production-ready AI systems, LLM agent frameworks, and advanced RAG pipelines.

2+

YEARS BUILDING

200+

MENTEES

25+

LIVE PROJECTS

CORE COMPETENCIES

Scalable System Design

Microservices Architecture

RAG / LLM Pipelines

API Design & Backend Eng.

Performance Optimization

DevOps Automation

Prompt Engineering

LLM Integration

WebGL / 3D Engines

MLOps & RL Systems

Developer Tooling

TECHNICAL SKILLS

LANGUAGES

JavaScript

TypeScript

Python

Java

C++

FRAMEWORKS

React.js

Next.js

Node.js

FastAPI

Flask

Fastify

AI / ML

LangChain

LangGraph

LlamaIndex

RAG

NLP

PyTorch

Ollama

DATABASES

MongoDB

PostgreSQL

Firebase

FAISS

Pinecone

Redis

DEVOPS & CLOUD

Docker

Nginx

GitHub Actions

CI/CD

Vercel

Azure

TOOLS & LIBS

Three.js

WebGL

Socket.IO

Prisma

Drizzle

TipTap

PROFESSIONAL EXPERIENCE

AI Developer

Apr 2026 - Present

Dimakh Consultants · Pune, India · Production-grade AI systems for B2B enterprise clients

50%+ Manual Effort Eliminated

60% Latency Reduction

B2B Production AI

- Engineered enterprise-grade **AI Agents** and advanced **RAG pipelines**, leveraging high-level tool chaining to automate complex workflows and cut manual B2B operations by **50%+**.
- Architected scalable AI inference backends deployed on **GCP** and **AWS**, ensuring high availability and seamless connectivity between AI agents and core business applications.
- Designed and developed highly performant, **SEO-optimized** web platforms and cross-platform **React Native** mobile applications.
- Owned the **end-to-end delivery lifecycle** - from gathering client requirements to production handoff and iterative optimizations.

Software Engineer

Nov 2025 - Mar 2026

Bug0 (Hashnode) · Remote · AI-integrated developer tools and globally distributed backend systems

40% Faster Deployments

Global User Base

- Developed the core architecture of Bug0, an AI-agentic testing platform integrating directly with **GitHub Actions** to automate interval-based **AI Playwright testing**.
- Built high-throughput **Node.js** backend services and CI/CD pipelines (**Docker**, **GitHub Actions**), reducing deployment cycle times by **40%**.
- Acted as a **Forward Deployed Engineer**, directly interfacing with enterprise clients to implement custom testing pipelines and drive core product adoption.
- Refactored system architectures to halve **cold-start latency** and significantly improve testing reliability on production deployments.

Full-Stack Developer & DevOps Intern

Aug - Oct 2025

LabsCheck · Pune, India · Healthcare SaaS and AI-driven lead automation

100,000+ Leads Automated

35% Faster APIs

- Spearheaded the development of a **React Native** mobile app and the core business logic for a B2C healthcare aggregator, connecting users to affordable diagnostic labs.
- Architected RESTful APIs using the **MERN stack**, facilitating the processing of **100,000+** inbound leads via AI-driven smart routing.
- Managed end-to-end DevOps operations, configuring **Nginx** reverse proxies, **Contabo VPNs**, and **Docker** containers for secure infrastructure.
- Optimized backend queries to improve API response times by **35%** and implemented integration testing to reduce post-deployment bugs by **60%**.

Web Developer Intern

Jan - Jun 2025

Walnut Solutions · Dhule, India · SEO-optimised, high-performance Next.js applications

30% Engagement Lift

30% Load Time Cut

- Built high-performance, SEO-friendly SSR/SSG web applications using **Next.js**, collaborating with senior engineers to optimize technical architectures.
- Cut page load times by **30%** via strategic code splitting, WebP optimization, and **Cloudflare** edge caching, driving a **30%** lift in user engagement.
- Implemented advanced **technical SEO** strategies - structured data, canonical tags, and **Core Web Vitals** optimizations - measurably boosting organic search rankings.

KEY PROJECTS

01	NoteForge AI - Autonomous Video-to-Notes RAG System AI Systems Architect LangChain Ollama Express Vector DB Advanced RAG - End-to-end RAG pipeline utilizing RecursiveCharacterTextSplitter for semantic chunking of extracted YouTube transcripts and audio streams. - Local Ollama model routing with task-specific inference paths, distributing heavy extraction tasks to high-context models while keeping data private. - Dense vector embeddings with nearest-neighbor search dynamically inject relevant context into dynamic prompt templates for highly accurate synthesis. - Auto-healing JSON generation layer leveraging jsonrepair to enforce strict schema adherence for automated quiz generation and structured study notes. GitHub Live Demo
02	FlowNotes - AI-Powered Visual Knowledge Workspace Frontend Engineer React Flow Zustand TipTap Gemini Infinite Canvas - React + TypeScript mindmap editor built on @xyflow/react with draggable canvas, automatic layout calculations, and localized persistent storage. - AI-assisted mindmap generation utilizing Gemini and Ollama providers to instantly parse and generate validated JSON knowledge structures. - Rich node editing capabilities integrating TipTap for inline text formatting, media attachments, and custom chapter/sub-topic node categorization. - Node-level AI expansion dynamically creates related subtopics on the canvas via a configurable multi-provider API service architecture. GitHub Live Demo
03	WCareer (WCareers.ai) - AI-Powered Career Copilot Full-Stack AI Lead Next.js 14 FastAPI GenAI Drizzle ORM SIH & Techathon Winner - Led microservices architecture bridging Next.js 14 frontend with Python FastAPI LLM inference server - awarded Smart India Hackathon & National Techathon 1st place. - ML-powered ATS Resume Wizard using NLP tokenization and entity extraction, improving structured resume parsing accuracy by 92%. - Real-time Generative AI mock interview engine streaming audio/video responses with semantic similarity scoring and structured feedback reports. - End-to-end authentication via Clerk and Neon PostgreSQL - secure multi-tenant user data with row-level access control. GitHub wcareers.ai
04	Mpulse - Event Micro-Frontend Ecosystem Frontend Architect Next.js GSAP Razorpay Vercel EAS CMS & SEO Engine - Decoupled micro-frontend: GSAP-animated public registration UI, CMS-driven Admin Panel, and Expo mobile check-in app - each independently deployable. - Razorpay payment gateway with idempotent webhook handlers processes ticket purchases reliably under high concurrent event-day load. - Technical SEO via SSR and JSON-LD structured data achieved top Google Event SERP indexing and flawless Core Web Vitals scores. - Automated CI/CD deploys edge functions to Vercel and pushes OTA mobile updates via EAS - zero manual release steps. GitHub Live Site
05	Web Mine Craft - Browser-Native WebGL Voxel Engine Graphics Engineer Three.js WebGL Vite Perlin Noise 60+ FPS Engine - Full 3D voxel game running natively in the browser at 60+ FPS using raw Vanilla Three.js - no game engine or UI framework wrapper. - Custom AABB physics collision system delivers accurate first-person movement, jumping, and mouse-driven block raycasting interactions. - Infinite procedural terrain using Perlin noise with dynamic chunk loading and GPU memory unloading as the player moves through the world. - InstancedMesh draw call batching reduces GPU calls by 90%, sustaining 100k+ simultaneous voxels on-screen without frame drops. GitHub Play Live
06	EduQuest - NLP-Powered Subjective Exam Platform Backend & AI Developer NLP Fastify PostgreSQL MongoDB Dual DB Architecture - Subjective exam platform using compromise.js NLP to algorithmically evaluate open-ended student answers against rubric criteria in real time. - Browser proctoring engine via Page Visibility API and focus-detection events reduces exam malpractice rates by 80% without external software. - Dual-database architecture: Drizzle ORM + PostgreSQL for relational auth/users, Mongoose + MongoDB for unstructured exam payload storage. - Fully containerized with Docker; staging and production environments share identical runtime snapshots for reliable, reproducible deployments. GitHub Live Demo
07	HackMate - Developer Matching CLI Network Systems Engineer React Ink Fastify Socket.IO Redis Published NPM Package - Globally distributed NPM CLI tool that renders a fully interactive terminal UI using React and Ink - works on any developer's machine instantly after install. - Full-duplex real-time in-terminal chat via Socket.IO with Redis pub/sub fan-out - supports high concurrent connections without message loss. - Proprietary 100-point GitHub compatibility algorithm analyses repository diversity, language breadth, and commit consistency to surface ideal collaborators. - JWT session management with Argon2id hashing secures locally-stored credentials in encrypted terminal config files. GitHub npm Package

TECHNICAL DEPTH & ARCHITECTURE

RAG & LLM Systems: End-to-end RAG pipelines (chunking, FAISS/Pinecone vector index, BM25 + dense hybrid retrieval, re-ranking); multi-step LangChain/LangGraph agents with tool-use, memory & conditional branching; fine-tuned Mistral & LLaMA; prompt engineering (CoT, few-shot, JSON output hardening); RAGAS evaluation.

System Design: Event-driven microservices, API gateways (Nginx), circuit breakers, horizontal AI inference scaling with Redis async queues & Docker Compose. RESTful/streaming APIs (OpenAPI spec), JWT auth, rate limiting, versioned endpoints.

DevOps & CI/CD: Docker Compose, Nginx reverse proxy, SSL/Let's Encrypt, GitHub Actions (lint → test → build → rolling deploy, <4 min pipelines). Observability: Winston/Pino structured logging, Prometheus metrics, Grafana dashboards.

Additional Skills: React Native cross-platform mobile apps; WebGL/Three.js 3D engines (60+ FPS, InstancedMesh GPU batching); Socket.IO real-time systems; NPM package publishing; NLP tokenization & entity extraction; MLOps (PPO/Actor-Critic RL, FastAPI tensor serving).

ACHIEVEMENTS & LEADERSHIP

EVENT	RESULT	CONTEXT
Smart India Hackathon 2024	Runner-Up	National - top-2 from 1M+ participants. Built AI-powered disaster response coordination system.
WCareer - National Techathon	Winner	1st place - AI career copilot; microservices architecture bridging Next.js + FastAPI LLM backend.
Innovate You Hackathon 2025	Winner	1st place, AI/automation track - end-to-end process automation platform, judged by industry experts.
India Innovates 2026	Finalist	National finalist - Jalsakhi precision agriculture AI platform, 500+ competing submissions.
Technical Lead	200+ Mentees	Workshops and 1:1 mentorship in full-stack development, system design, and career readiness.
HackMate CLI	NPM Published	Globally distributed NPM CLI developer-matching tool - active community adoption and OSS contributions.

EDUCATION

DEGREE

B.Tech - Information Technology

PES Modern College of Engineering, Pune

CGPA 8.21 / 102022 - 2026

Coursework: Data Structures & Algorithms · Operating Systems · Database Management · Computer Networks · Machine Learning · Software Engineering · Distributed Systems

Open Source & Community

Published multiple NPM packages for SaaS scaffolding and developer tooling - actively maintained with real-world adoption. Open-source PRs reviewed and merged by maintainers across JS/AI ecosystem repositories.

PRODUCTION STACK

DockerNginxLet's EncryptGitHub ActionsGrafanaPrometheusRedisCloudflareVercel EdgeAzurePrismaDrizzle ORM

SAMEER BAGUL
Full-Stack & AI Developer · DevOps Engineer

sameerbagul.com

sameerbagul2004@gmail.com

+91 7841941033 · github.com/Sameer-Bagul