

NINAD DIGHE

dighe.ninad7@gmail.com ♦ +91 9175348865 ♦ GitHub ♦ LinkedIn

PROFESSIONAL SUMMARY

Founding Engineer and Ex-Google SRE with 6+ years of experience building scalable, AI-driven systems. Expert in bridging the gap between rapid product development and reliability engineering. Proven track record of leading engineering teams, architecting agentic RAG pipelines, and ensuring 99.99% availability for high-traffic distributed systems. Passionate about leveraging Large Language Models (LLMs) to solve complex business problems.

PROFESSIONAL EXPERIENCE

Krazimo

Bangalore, Karnataka

Founding Engineer

10/2025 - Present

- Architected Agentic RAG pipelines using LangChain and Knowledge Graphs to enable efficient long-term and short-term memory retrieval, significantly enhancing context accuracy in AI systems.
- Orchestrated the technical roadmap for a client's engineering team, establishing code review standards, CI/CD pipelines, and mentorship protocols. Resolved critical scalability bottlenecks and data leaks, directly enabling a successful go-to-market strategy.
- Spearheaded the development of full-stack Agentic AI solutions (Python, GoLang, Web3), including an on-chain information retrieval system and an AI-driven sales training platform for high-ticket client acquisition.

Google

Bangalore, Karnataka

Software Engineer - Site Reliability Engineering

10/2023 - 10/2025

- Engineered AI workflows to automate the launch review process, gathering production readiness requirements and providing proactive feedback, which reduced manual toil by 70%.
- Achieved a 100% reduction in Recovery Time Objective (RTO) for data incidents by designing and implementing "Point-in-Time" restore capabilities for critical databases.
- Developed a schema canarying framework for internal databases to prevent bad schema pushes, eliminating 90% of production data incidents.
- Designed and developed "Detectr," a tool for automated discovery of software bugs from social platforms (Reddit, X), automatically triaging tickets to relevant dev teams with actionable steps.
- Served as a reliability gatekeeper for major service launches, collaborating with product teams to enforce production best practices and conducting postmortems to identify architectural improvements.
- Improved reliability of critical Android/iOS workflows by implementing priority-based load shedding, reducing task overloads by 80% during high-traffic events.

Google

Bangalore, Karnataka

Technical Solutions Engineer, Google Cloud

08/2020 - 09/2023

- Developed an automated GCP incident handler for GKE and Anthos deployments in GoLang, reducing incident response time by 50%.
- Automated support workflows for GKE autoscaling using "Troublemaker," enabling automatic detection of 80% of scaling issues in customer clusters.
- Acted as Subject Matter Expert (SME) for GKE and Anthos, resolving high-severity incidents for premium enterprise customers including Spotify, Snap, and Flipkart.

Hermitcrabs

Bangalore, Karnataka

Software Engineer

1/2020 - 06/2020

- Built a full stack website for task assignment and intern management, streamlining interaction between interns and team leads.
- Designed performance assessment dashboards for administrators to track task completion and intern progress.

EDUCATION

PES University

Bangalore, Karnataka

B.Tech., Computer Science Engineering, 8.5 CGPA

08/2016 - 08/2020

DAV Public School

Pune, Maharashtra

12th, Science

09/2010 - 07/2016

TECHNICAL SKILLS

Core Competencies	Languages & Tools
<ul style="list-style-type: none">• Software Engineering & System Design• Site Reliability Engineering (SRE)• AI/ML & Agentic Workflows• Cloud Architecture (GCP)• Problem Solving & Algorithms	<ul style="list-style-type: none">• Python, GoLang, C++• JavaScript/TypeScript, Flutter• Kubernetes, Docker, Terraform• LangChain, TensorFlow

PROJECTS

Foreverkind - Therapeutic AI Assistant & Life Coach	Python, LLMs, NLP
Personal Project	
<ul style="list-style-type: none">• Built an empathetic AI companion acting as a therapeutic assistant and life coach, utilizing Large Language Models (LLMs) for personalized mental wellness support.• Implemented Retrieval-Augmented Generation (RAG) with Vector Databases (e.g., Pinecone) to maintain long-term context of user conversations and track emotional progress.• Designed a privacy-focused architecture with sentiment analysis guardrails to ensure safe interactions, leveraging prompt engineering to handle sensitive topics responsibly.	
Agentic Automated Instagram Content Pipeline	Python, LangChain, Instagram Graph API
Personal Project	
<ul style="list-style-type: none">• Architected an end-to-end autonomous pipeline using Python and LangChain to generate, schedule, and publish high-engagement content on Instagram without manual intervention.• Integrated multimodal AI agents to analyze trending topics, generate captions, and create visual assets via generative models, streamlining the content creation lifecycle.• Implemented robust error handling and rate-limiting safeguards to interact reliably with the Instagram Graph API for seamless automated posting.	
Dating App for Professionals	Flutter, Firebase, AI
Personal Project	
<ul style="list-style-type: none">• Developed a secure, cross-platform dating application using Flutter, designed specifically for verified working professionals.• Designed a scalable microservices architecture to handle real-time messaging via WebSockets and secure user verification workflows.• Engineered AI-driven matchmaking algorithms that analyze professional trajectories and personality compatibility to provide highly relevant partner suggestions.	