

VINOD KUMAR G

Senior Backend Engineer

AI Engineering • Gen AI • Agentic AI • Distributed Systems • Platform Engineering

Dubai, UAE | Bengaluru, India

☎ +91 910814 3704 ✉ imvinodkumarg@gmail.com 🔗 linkedin.com/in/vinod-kumar-g 🐙 github.com/VinodKumarGgithub

Professional Summary

Senior Backend & Platform Engineer with 5+ years of experience designing distributed systems, enterprise SaaS platforms, and AI-powered applications. Experienced in building scalable backend architectures, high-performance APIs, workflow automation, multi-tenant systems, and data-intensive platforms using Go, Node.js, TypeScript, Redis, and modern cloud technologies.

Recently focused on AI platform engineering, building production-ready LLM applications with Retrieval-Augmented Generation (RAG), Model Context Protocol (MCP), agentic workflows, tool calling, semantic search, and AI workflow automation using LangChain, LangGraph, LangSmith, LangFuse, Pinecone, Tavily Search, OpenAI, Gemini, and GroqCloud.

Experienced in taking products from architecture and implementation through production deployment, performance optimization, and operational ownership.

Education

University of Maine at Presque Isle (UMPI)

Presque Isle, Maine, USA

Bachelor of Liberal Studies (B.L.S.), Management Information Systems (MIS)

Experience

MedImpact Healthcare Systems

Aug 2023 – Present

Senior Backend / Platform Engineer

Dubai, UAE

- Partnered with product managers, business stakeholders, and engineering teams to translate complex healthcare and insurance workflows into scalable backend solutions supporting enterprise Pharmacy Benefit Management (PBM) platforms.
- Owned the design, implementation, deployment, and operational support of business-critical backend services, ensuring reliability, scalability, and long-term maintainability.
- Architected scalable backend services and APIs supporting enterprise PBM workflows through distributed processing, caching, and asynchronous execution.
- Architected a configurable multi-tenant platform supporting tenant isolation, role-based access control (RBAC), metadata-driven configuration, and reusable backend components.
- Optimized backend performance through query optimization, caching strategies, asynchronous processing, and production profiling, improving scalability and operational reliability.
- Developed internal automation tools that eliminated repetitive operational workflows, improving engineering productivity and reducing manual effort across teams.
- Led production issue investigations, root cause analysis, and observability improvements across production systems to ensure platform stability.

ToXSL Technologies

May 2022 – Jul 2023

Software Engineer

Chandigarh, India

- Designed and developed scalable web applications and backend services using Node.js, React, Express.js, and REST APIs for enterprise clients.
- Collaborated with clients and internal teams to understand business requirements, design technical solutions, and deliver production-ready software.
- Integrated backend services, optimized application performance, and participated in production support, debugging, and Agile software delivery.

Illuminz

Jun 2020 – May 2022

Software Developer

Mohali, India

- Built full-stack web applications using Node.js, Express.js, React, and MySQL, delivering scalable business applications for multiple clients.
- Developed RESTful APIs, implemented backend business logic, and optimized SQL queries to improve application performance and reliability.
- Collaborated throughout the software development lifecycle, from requirement analysis and solution design to deployment and production support.

Key Projects

Enterprise PBM Platform | *Node.js, Redis, MySQL, BullMQ*

- Designed and delivered backend services supporting enterprise Pharmacy Benefit Management (PBM) platforms handling complex healthcare and insurance workflows.
- Built scalable REST APIs, distributed job processing pipelines, and reusable backend services using Node.js, BullMQ, Redis, and MySQL.
- Automated high-volume operational workflows using asynchronous processing, reducing manual effort and improving overall platform efficiency.
- Improved backend performance through query optimization, caching strategies, and production performance tuning.

Enterprise Config-Driven Multi-Tenant Low-Code Platform | *Node.js, React, TypeScript, MySQL, Redis*

- Architected a configuration-driven enterprise platform enabling dynamic generation of operational pages, forms, tables, and CRUD workflows without additional application development.
- Designed a scalable multi-tenant architecture supporting tenant isolation, multi-database switching, RBAC, tenant lifecycle management, configurable data sources, and centralized user administration.
- Implemented data governance capabilities including audit logging, role-based access control, download permissions, and secure tenant-level data access policies.
- Built reusable platform components that accelerated feature delivery while supporting enterprise-scale data processing and high-volume operational workloads.

LARA – AI Assistant Platform | *Node.js, LangChain, LangGraph, LangSmith, RAG, Tavily Search*

- Designed and developed an AI assistant platform integrating OpenAI-compatible LLMs, LangChain, LangGraph, Retrieval-Augmented Generation (RAG), Model Context Protocol (MCP), and Tavily Search.
- Built agentic workflows supporting tool calling, document retrieval, multi-step reasoning, and workflow orchestration.
- Integrated AI capabilities into backend platforms to automate repetitive engineering and business workflows while maintaining scalability and reliability.
- Developed reusable AI orchestration components enabling enterprise AI integrations across multiple business use cases.

Technical Skills

Languages: JavaScript, TypeScript, Go, SQL

AI: LangChain, LangGraph, LangSmith, LangFuse, MCP, RAG, Pinecone, Tavily Search, OpenAI API, Gemini REST API, GroqCloud

Backend: Express.js, REST APIs, GraphQL, gRPC, Protocol Buffers, Microservices, API Design, BullMQ

Databases: MySQL, PostgreSQL, Redis, MongoDB, MSSQL

Cloud & DevOps: AWS, GCP, Azure, Docker, Docker Compose, Git, Linux, PM2

Architecture: Advanced Multi-Tenant SaaS, Distributed Systems, Platform Engineering, System Design, Event-Driven Architecture, High-Performance APIs, Performance Optimization, Caching, Fault Tolerance