PREM KASHYAP CHILAKAMARTHI

prem.kashyap.0206@gmail.com | +91 9502645945 | prem-kashyap.pages.dev
linkedin.com/in/prem-kashyap-chilakamarthi | github.com/premkashyap004

Professional Summary

Accomplished Full-Stack AI Developer with over two years of experience in designing and deploying AI-driven applications. Proficient in developing scalable full-stack solutions and integrating advanced AI systems into web platforms. Expertise in containerization, CI/CD pipelines, and cloud infrastructure management. Adept at architecting efficient workflows and delivering user-centric interfaces to enhance functionality and performance.

Technical Skills

Languages: Python, JavaScript, TypeScript

Frontend: React, HTML, CSS, Bootstrap

Backend: Node.js, Flask, FastAPI

Cloud & DevOps: AWS, GCP, Docker, GitHub Actions, Kubernetes

GenAI Frameworks: OpenAI, LangChain, LlamaIndex, LangGraph, MCP, Langmem

Databases: MongoDB, ChromaDB, Redis, Qdrant, SQL

Work Experience

Generative AI Developer, **Sify Technologies** July 2023 – Present

- Engineered multiple generative AI proof-of-concepts, including storyboard creation, media enrichment, and assessment generation, and developed Scout AI, a real-time voice-to-voice chatbot for children, leveraging OpenAI frameworks.
- Conducted advanced research and prototyping to integrate and scale AI solutions into learning management systems and standalone applications.
- Deployed full-stack applications using Python, React, and Node.js on EC2 and virtual machines, configuring Nginx, system services, and cloud infrastructure for robust production environments.
- Designed scalable AI workflows for storyboard generation using serverless cloud architectures, optimizing for concurrency, cost efficiency, and multi-tenant functionality.
- Customized and deployed open-source AI models on private cloud infrastructure for scalable inference, utilizing custom scripts, XInference, and vLLM.

Education

B.E. in Computer Science and Engineering Aug 2019 – Apr 2023
Sri Sairam Engineering College, Chennai, Tamil Nadu, India CGPA: 9.28/10.0

Certifications

- **AWS Certified Developer Associate** - View Certificate April 2023 - April 2026
- **Oracle Generative AI Certificate** - View Certificate July 2024 - July 2026
- **AWS Certified AI Practitioner (Udemy)** - View Certificate September 2025

Achievements

- Published a research paper titled "Chronic Kidney Disease Prediction Using Different ML Algorithms" in IEEE Xplore at the 2022 1st International Conference on Computational Science and Technology (ICCST). *View Publication*
- Secured 1st place in the Mini-Project Expo for the "Chronic Kidney Disease Prediction Using Different ML Algorithms" project.

Projects

Ticket Resolution Automation System – Agentic AI Workflow (In Progress)

- Architected an agentic AI workflow using LangGraph (Python) to automate ticket resolution within Sify's cloud management ecosystem, introducing intelligent state-based decision-making.
- Developed a proof-of-concept for automated infrastructure provisioning on AWS using MCPs provided by AWS, significantly streamlining and accelerating ticket handling workflows.
- Creating MCPs like ssh-server, code generator (generator, debugger, executor) for workflow to streamline the communication between the llm and the sify's cloud infrastructure.
- Building a fully autonomous ticket management system leveraging agentic states and dynamic nodes to process, diagnose, and resolve incidents — escalating to human intervention only when necessary.
- Implemented multiple MCP servers to enable seamless communication between AI agents and Sify's cloud infrastructure, ensuring reliable orchestration and data exchange across environments.
- Once fully deployed, the system will automate incident request tickets, reducing human dependency and cutting resolution times by up to 50%.

Skillflo – AI-Powered E-Learning Platform

- Architected a serverless generative AI platform for e-learning using LlamaIndex, LangChain, and AWS, enabling semantic search, automated assessment generation, storyboard creation, and AI-enhanced media libraries.
- Implemented multi-tenant architecture with ChromaDB and Qdrant vector databases, ensuring scalable, high-performance AI service delivery across diverse learning environments.
- Transformed content development workflows through AI-driven personalization and interactive learning features, reducing instructional design time by over 30%.

Sify's Media Library – AI-Powered Media Management Platform

- Engineered a full-stack MERN application with AI-driven media management, featuring automated captioning, object detection, and semantic search powered by LlamaIndex.
- Designed and deployed an advanced semantic search engine enabling intelligent media categorization and fast, context-aware asset retrieval.
- Integrated the AI-powered media library into multiple platforms, including SkillFlo, to deliver scalable, multi-tenant media management across diverse applications.
- Centralized the asset validation process, eliminating the need for separate LLM-based enrichment workflows — cutting content creation and verification time.

Storyboard Generation – AI-Powered Content Creation Tool

- Developed a Python-based automation tool for storyboard generation, transforming raw instructional materials into structured content aligned with Bloom's taxonomy and customizable prompt templates.
- Engineered a robust content parsing and analysis pipeline, enabling accurate extraction, contextual mapping, and full traceability to original source documents.
- Integrated a fact-checking engine within the SkillFlo platform to ensure storyboard accuracy and consistency with verified source content.
- Built an interactive workarea module for SkillFlo, allowing chat-based collaboration with source materials and generated storyboards for faster content refinement.
- Empowered instructional designers (IDs) to directly upload raw documents and auto-generate storyboards, dramatically reducing manual review effort and accelerating content creation cycles by over 40/
- Embedded the application as a modular SkillFlo component, enhancing the platform's instructional design and AI-assisted content development capabilities.

Scout AI – AI-Powered Companion Chatbot for Children

- Developed an AI-powered interactive assistant for children (ages 4–12) using GPT-4o, Python, and FastRTC, delivering real-time bedtime stories, homework help, and engaging conversations.
- Built a safe, age-appropriate interface with dynamic multi-agent context switching and adaptive storytelling for personalized learning experiences.
- Introduced imaginative roleplay and AI-generated quizzes, fostering creativity, engagement, and knowledge retention through immersive interactions.