

Atharva Sarvesh Devasthali

Software Engineer | Full-Stack & AI/ML Development

Dallas, TX | 682-340-7219 | atharvadevasthali22@gmail.com | linkedin.com/in/atharva-devasthali | atharvadevasthali.com

AI & automation engineer with **3+ years** building production **LLM pipelines, distributed n8n workflows, and FastAPI microservices**. Worked on **Google properties serving 10M+ users**. Currently designing **agentic AI systems** with **99.5% uptime, structured LLM outputs**, and observability via **Promptail/Loki**.

Education

Masters in Computer Science

University of Texas at Arlington

Arlington, TX

August 2023 - May 2025

Bachelor of Engineering in Information Technology

Maharashtra Institute of Technology

Pune, India

July 2016 - April 2020

Work Experience

Software Engineer | Backend & AI Automation

Smart Cookie Rewards

Remote, USA

August 2025 - Present

- Instrumented **n8n workflow observability** using **Promptail and Loki** with a custom **React dashboard** for real-time failure tracing, CPU load, and job execution metrics across **3+ distributed servers**.
- Engineered **LLM evaluation pipeline** in n8n using **structured outputs** to assess script quality across completeness, tone, and relevance metrics, reducing manual review by **85%**.
- Designed **distributed system architecture** using **n8n queue mode** across **3+ Contabo servers** with **MinIO object storage**, enabling horizontal scaling and fault-tolerant workflow processing.
- Built **Redis-based distributed task queue** across 4 servers with **CPU-gated job execution** and **EWMA-smoothed load monitoring**, scheduling **18+ long-running video generation jobs** daily without resource exhaustion.
- Developed **FastAPI microservices** with **Docker containerization**, processing **100+ automation requests daily** with real-time alerting and retry mechanisms, achieving **99.5% uptime** over 6 months.
- Led engineering team to architect a **24/7 streaming platform** generating **60+ videos weekly** with **multi-channel YouTube** delivery via **YouTube Data API** and **FFmpeg** encoding.

Software Engineer | Google Client Projects

Cybage Software Pvt. Ltd.

Pune, India

December 2020 - July 2023

- Developed and maintained **8+ high-traffic Google marketing sites** ([Google Lens](#), [Google Translate](#), [Google Sustainability](#)) serving **millions of daily users**, built with **TypeScript, Node.js, Angular**, and **Python/Flask**.
- Reduced page load times by **90%** through code-splitting, lazy loading, image optimization, and HTTP/HTTPS protocol optimization, achieving **Lighthouse scores of 92+** across all properties.
- Analyzed **Google Analytics data** to drive UI/UX improvements, increasing **user engagement by 80%** through data-driven optimization of click patterns and scroll behavior.
- Ensured **WCAG 2.1 AA compliance** for **10M+ monthly users** by implementing **semantic HTML, ARIA labels**, and comprehensive **keyboard navigation**, and testing with assistive technologies including **screen readers**.
- Conducted internal **GSAP animation training** for engineering team, upskilling peers on **performance-conscious motion design** across **Google properties**.

Technical Skills

Programming Languages: Python, Go, TypeScript, JavaScript, Java, SQL, C++

Backend & APIs: FastAPI (OpenAPI), Flask, Django, Node.js, Express.js, Gin, Spring Boot, REST APIs, GraphQL, n8n, Micro-services, WebSockets

Frontend: React, Next.js (SSR/SSG), Electron, Angular, Astro, TypeScript, HTML/CSS, Sass/SCSS, Tailwind CSS, shadcn/ui, Storybook.js, GSAP, Webpack, Vite, Figma, Component-Driven Development

State Management & Tools: Zustand, Redux, React Query, React Context

AI/ML: LLM Integration (GPT-4o, Gemini 2.5 Pro, Claude 3.5 Sonnet), RAG Systems, Vector Databases (Supabase), TensorFlow, Keras, Hugging Face, Multi-Agent Systems, Prompt Engineering, LangChain, LangGraph, MCP, Structured Outputs

Databases: PostgreSQL, MySQL, MongoDB, Redis, Firebase

Cloud & DevOps: Docker, Kubernetes, Ansible, GitHub Actions, AWS (S3), MinIO, Promptail, Loki, GCP, CI/CD, Git, NGINX, Agile, Scrum, Linux/UNIX

Testing: pytest, Jest, React Testing Library, Playwright, Selenium, TDD, Unit Testing, Integration Testing, E2E Testing

Projects

StackPilot - Self-Hosted DevOps Control Plane | Go, Astro, React, Zustand, TypeScript, Docker, Kubernetes

- Architected comprehensive **DevOps platform** enabling deployment, monitoring, and scaling of **50+ applications** by implementing **Docker** orchestration, **Prometheus/Loki monitoring**, and automated **CI/CD workflows**.

- Built **full-stack control plane** with **25+ REST API** endpoints serving real-time deployment status with **99% uptime** by developing **Go backend APIs and React/TypeScript dashboard** with WebSocket connections for live updates.
- Implemented **Zustand** for efficient client-side state management, handling complex application state including deployment status, resource metrics, and real-time **WebSocket** updates across 50+ applications.
- Designed secure secrets management with **JWT authentication, AES-256 encryption, RBAC**, and **audit logging**, meeting **SOC2/HIPAA compliance** requirements for regulated environments.

ADVI-UI Component Library | React, TypeScript, Tailwind CSS, Storybook, Figma, Claude Code, Cursor, Vercel | [GitHub](#)

- Designed and built **production-ready React component library** (Advi-ui) following **MVVM architecture patterns** using **Figma** for design system, **Storybook** for component documentation, and **Claude Code with Cursor**.
- Created **20+ reusable, accessible components** with React, TypeScript, and Tailwind CSS, ensuring **WCAG 2.1 AA compliance**, optimal **Core Web Vitals**, and **semantic HTML** throughout for **isolated component development and documentation**.
- Implemented comprehensive **design tokens and theming system** for consistent visual language across components, including proper **focus states, reduced motion preferences**, and **ARIA attributes** for screen readers.
- Published to **npm** as public package for cross-project reusability, deployed interactive Storybook documentation to **Vercel**, demonstrating complete **design-to-code workflow** from Figma prototypes to production components.

AI-Powered Lip-Sync Video Generation System | Python, FastAPI, Wav2Lip, LivePortrait, FFmpeg | [GitHub Link](#)

- Engineered **advanced video synthesis pipeline** combining **Wav2Lip** for audio-driven lip-synchronization with **LivePortrait** for realistic face movement, creating photorealistic talking-head videos from static images and audio.
- Implemented **multi-GPU acceleration**: optimized for **macOS MPS** and **NVIDIA CUDA**, achieving **6x speedup** in video generation throughput over CPU baseline.
- Architected **FastAPI microservice** exposing video synthesis as scalable REST API with async generation, webhook callbacks, and real-time progress tracking.

RAG-Powered Blog Content Assistant | LangChain, Django, PostgreSQL, OpenAI | [GitHub Link](#) | [Live Demo](#)

- Built **retrieval-augmented generation (RAG)** system using **LangChain** to enable **semantic search** and **Q&A** across development blog articles, integrating with **PostgreSQL pgvector** vector store for efficient embedding-based retrieval.
- Implemented full **LLM pipeline: document chunking, embedding generation** (OpenAI), **semantic similarity search**, and **context-aware prompt engineering** for intelligent content recommendations based on user queries.
- Deployed as **REST API** endpoints in **Django** backend with **real-time LLM responses** and **source attribution**, achieving 95%+ accuracy in retrieving contextually relevant articles.

Publications

First-author publication in Springer ICECMSN Conference (LNDECT Vol. 53): ["IoT Based Inventory Management System with Recipe Recommendation using Collaborative Filtering"](#) - 3 citations