

Anu L Sasidharan

4164533028 ◇ anulsasidharan@gmail.com ◇ Canada ◇ [LinkedIn](#) ◇ [GitHub](#) ◇ [Portfolio](#)

SUMMARY

AI Engineer with extensive experience in Big Data and cloud engineering, specializing in developing scalable, production-ready AI systems. Proficient in Generative AI, agentic AI frameworks, and LLM integration leveraging LangChain, OpenAI APIs, Streamlit, MLflow, and vector databases such as FAISS, Pinecone, and Chroma. Skilled in architecting end-to-end data and AI pipelines across GCP, AWS, and Azure, with expertise in Kubernetes-based model orchestration. Combines robust data infrastructure knowledge with practical problem-solving to deliver secure, high-performing AI solutions designed for autonomous reasoning, intelligent interaction, and enterprise knowledge management.

SKILLS

Core AI & Machine Learning: Generative AI, Large Language Models (LLMs), Agentic AI frameworks, Prompt engineering, Model fine-tuning & optimization, SFTTrainer, PEFT, TRL, LORA, Retrieval-Augmented Generation (RAG)

AI Model Evaluation & Benchmarking: HuggingFace Open LLM Leaderboard, LMSYS Arena, RAG evaluation metrics

AI Engineering & MLOps: LangChain, OpenAI APIs, Hugging Face, MLflow, Vector DB, FAISS, Pinecone, Chroma, Model deployment, Kubernetes

AI Agents & Orchestration Frameworks: LangGraph, Autogen, CrewAI, MCP, n8n, Agent Workflows, API integrations, Tool-Calling Systems, Multi-agent Systems

Programming & Software Engineering: Python, PyTorch, TensorFlow, Streamlit, Gradio, API design

Data Engineering & Big Data: Hadoop ecosystem, Spark, Kafka, HIVE, ETL, Data pipelines

Cloud & Infrastructure: GCP, AWS, Azure, Docker, CI/CD

EXPERIENCE

Technology Specialist

Tech Mahindra

Dec '22 — Nov '25

Toronto, Canada

I supported Scotiabank's AI/ML platform during the migration from EDL to EDLR and the upgrade of GCP Dataproc environments. My work centered on building Python automation to streamline cluster provisioning, configuration checks, workload analysis, and monitoring—improving reliability for data and ML pipelines.

I collaborated with data scientists and ML engineers to optimize resource usage, enhance data access, and ensure stable compute environments for model training and batch inference. I also developed Python utilities for security validation (Kerberos, Ranger, Auto-TLS) and load balancer testing, helping maintain a secure, ML-ready platform.

On the cloud side, I contributed to standardizing Dataproc clusters using Python and REST APIs, supported performance tuning, and assisted in DR planning to ensure uninterrupted ML operations.

- Directed the end-to-end modernization of enterprise Big Data infrastructure by migrating on-premises EDL environments to the EDLR platform using the latest Cloudera distribution, while upgrading GCP Dataproc clusters to boost scalability, reliability, and ML pipeline throughput.
- Developed and automated high-availability data ecosystems by streamlining cluster provisioning, enhancing performance tuning, and integrating AI/ML workloads across GCP and Cloudera CDP7 environments.
- Partnered with cross-functional teams in data science, engineering, and infrastructure to deliver production-ready big data solutions, streamlining EDL-to-EDLR migration and GCP Dataproc standardization for improved operational efficiency.

Associate Technical Architect

Tavant Technologies

Sep '21 — Sep '22

Banglore, India

- Built Python automation to administer, monitor, and optimize large-scale, multi-tenant Hadoop clusters that supported enterprise data, analytics, and ML workloads.
- Used Python-driven load analysis, resource usage modeling, and automated configuration checks to improve system reliability, balance workloads, and optimize resource pools for data science teams.
- Developed Python scripts and workflow tooling to support disaster recovery processes, validate RPO/RTO objectives, and ensure uninterrupted access to ML training data and production pipelines.
- Worked closely with data engineering, ML, and analytics teams to troubleshoot and fine-tune components across Kafka, Spark, HBase, and Hive—enhancing data availability, reducing pipeline latency, and improving end-to-end model execution performance.

Senior Analyst - Application Operation

Allianz Technology SE

Apr '16 — Sep '21

TVM, India

- Managed and optimized multiple secure CDH5 clusters across TEST, DEV, ANALYTICS, and PROD environments, maintaining 99% uptime and high availability.
- Strengthened platform security by implementing Kerberos authentication, Sentry-based authorization, and end-to-end data encryption (at rest and in motion).

- Developed Python-based automation scripts and REST APIs to monitor cluster health, analyze job metrics, and streamline administrative workflows.
- Designed custom tools for anomaly detection, log analysis, and proactive alerting to reduce downtime.
- Oversaw Cloudera Manager upgrades, Solr Search integration, and BDR pipelines for seamless inter-cluster data replication and backup.
- Tuned resource pools, managed HDFS quotas and ACLs, and optimized system performance to support large-scale analytics and data processing workloads.

IND IT Software Engineer II

Aon Hewitt

May '15 — Apr '16

Chennai, India

- Deployed and managed CDH4/CDH5 clusters supporting large-scale production and analytics workloads with high availability and reliability.
- Led disaster recovery planning and implemented automated data backup and cross-data-center replication using Cloudera BDR.
- Built Python-based monitoring tools to analyze cluster logs, detect anomalies, and generate real-time alerts for proactive issue resolution.
- Designed and maintained resource management frameworks with static and dynamic resource pools to ensure fair scheduling and optimal resource utilization.
- Strengthened platform security through Kerberos authentication and SSL encryption, enabling a fully secured Hadoop environment.
- Collaborated with global teams to provide 24/7 operational support, streamline troubleshooting, and enhance system performance through automation and log-driven insights.

PROJECTS

RAG-Document-Search [Link](#)

- Built a LangChain + Hugging Face + FAISS RAG pipeline for multi-format documents, improving retrieval accuracy and reducing query latency through optimized embeddings and vector search.
- Applied structured data handling, strong programming principles, and soft skills to enhance scalability and coding practices aligned with knowledge engineer standards.

NeoGraphAI [Link](#)

- Developed NeoGraphAI, a Python-based framework integrating knowledge graphs with generative AI workflows to enable seamless ingestion of structured and unstructured data, graph-centric knowledge representation, and context aware retrieval for intelligent query interactions.

MCP-InsightEngine - LLM-Based File Analyzer & Insight Generator [Link](#)

- Engineered an LLM-driven file analysis platform leveraging Python, FastAPI, and Streamlit to process CSV, JSON, PDF, and text files, delivering real-time insights such as summaries, key trends, and anomaly detection.
- Built end-to-end data pipelines and implemented robust integration workflows to support business understanding and data strategy.
- Applied domain knowledge and proven ability to design scalable data pipeline architecture, ensuring efficient data processing and optimization.
- Demonstrated strong communication skills, language skills, and soft skills to explain technical implementation to diverse stakeholders, showcasing implementation experience and adherence to best coding practices.

Nutri-GPT [Link](#)

- Developed Nutri-GPT, an AI-powered web application integrating OpenAI's Vision and Chat APIs to process food images and automatically generate structured nutrition summaries and calorie insights using Python and Streamlit, while applying soft skills for effective collaboration and user-focused design.
- Engineered a modular AI-driven nutrition analysis platform with Python and Streamlit, leveraging Vision- based image recognition and LLM capabilities to deliver accurate food identification, calorie estimation, and health insights through a scalable architecture, supported by strong soft skills for stakeholder communication.
- Architected and deployed Nutri-GPT, a computer-vision and LLM-based food analytics tool that transforms images into actionable nutrition insights, demonstrating effective use of OpenAI Vision models, structured data handling, and interactive web app development, complemented by strong soft skills in collaboration and presentation.

EDUCATION

BTech in Electronics and Telecommunications, Kerala University

Jun '05 — Aug '09