

DEEPAN CHAKARAVARTHI J

AI Engineer | Generative AI, LLMs & Retrieval-Augmented Generation (RAG) | Document AI

Chennai, India | deepanchakaravarthij@gmail.com | +91 63790 47473
LinkedIn: deepanchakaravarthij | GitHub: DEEPANCHAKKARAVARTHIJ

PROFESSIONAL SUMMARY

AI Engineer with 1.5 years building production Generative AI applications on AWS. Designs and ships Retrieval-Augmented Generation (RAG) systems using Large Language Models (LLMs), AWS Bedrock, Weaviate vector search and FastAPI. Built a Legal AI assistant achieving **85–90% response relevance** in evaluation testing, and is developing a Document AI platform that uses OCR and computer vision to validate PDFs, scanned documents and images for regulated workflows. Backend engineering foundation in multi-tenant SaaS, RBAC and tenant isolation enables AI systems that are secure, scalable and cost-efficient, including a **~20% reduction in AWS operational costs**.

TECHNICAL SKILLS

Generative AI & LLMs: Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), AWS Bedrock, Prompt Engineering, Embeddings, Semantic Search, Vector Databases (Weaviate), Context Optimization, LLM Evaluation

Machine Learning: Natural Language Processing (NLP), Computer Vision, CNN Object Detection, Scikit-learn, TensorFlow/Keras

Document AI: Optical Character Recognition (OCR), Document Comparison, Layout Analysis, Barcode/QR Validation

Backend & Cloud: Python, FastAPI, REST APIs, Microservices, Docker, AWS, GCP, SQL, Next.js, Multi-Tenant SaaS, RBAC, MFA, Audit Logging, Tenant Isolation

Geospatial & Scale: LiDAR/COPC, GeoTIFF, Raster Imagery, WGS84 Mapping, Batch Inference, GPU Acceleration

Analytics: Power BI, SAP Analytics Cloud, Matplotlib

PROFESSIONAL EXPERIENCE

Software Engineer — Solidpro Engineering Support Pvt Ltd., Chennai Mar 2025 – Present

- Architected a RAG-based Legal AI assistant using AWS Bedrock LLMs, Weaviate semantic search and FastAPI, achieving **85–90% response relevance** in evaluation tests through embedding optimization, document preprocessing and prompt engineering.
- Reduced AWS cloud operational costs by **approximately 20%** by optimizing infrastructure and LLM usage across production AI and SaaS services, without degrading response quality.
- Engineering a Document AI comparison platform that combines OCR and computer vision to detect text, value, layout, image and barcode/QR changes across PDFs and scanned files, with rule-based severity classification and audit-ready reports.
- Built scalable FastAPI services for multi-tenant SaaS and AI apps with hierarchical RBAC, MFA, audit logs and tenant isolation.

Key Projects

DRT Document Project — Document AI & Comparison Platform 2026 – Present

- Building OCR and computer vision pipelines for regulated document quality control, compliance checks and change validation.
- Designing FastAPI workflows for file upload, comparison job processing, result review and audit-ready report generation.

Legal AI Assistant — Retrieval-Augmented Generation Mar 2025 – Nov 2025

- Built semantic retrieval pipelines in Weaviate with optimized embeddings, tokenization and document preprocessing.
- Developed a FastAPI backend for real-time, multi-turn conversational AI using AWS Bedrock LLMs.

Wetlands Geospatial AI Platform — Adrieluxe Sep 2025 – Feb 2026

- Integrated CNN-based computer vision models to detect, count and classify trees from LiDAR point clouds and raster imagery.
- Mapped detections to WGS84 geocoordinates and scaled large-area inference with batch processing and GPU acceleration.

SolidPro Suite — Multi-Tenant SaaS Platform Dec 2025 – Present

- Architected a Super Admin / Tenant / App Admin hierarchy with secure tenant isolation and role-based dashboards (Next.js, FastAPI).

ISTANA ERP System

- Built backend data pipelines and integrated NAS storage for secure document management and workflow automation.

SAP Analyst — Kaavian System, Chennai Jan 2024 – May 2024

- Developed real-time SAP Analytics Cloud dashboards that **improved reporting efficiency by 30%**.

EDUCATION

M.Sc. Data Science and Business Analytics — Vels Institute (VISTAS), Chennai 2022 – 2024

B.Sc. Computer Science — Aksheyaa College of Arts & Science, Chengalpattu 2019 – 2022

CERTIFICATIONS

AWS Certified Solutions Architect – Associate (In Progress)