

ARNOL P S

Senior Data Science Specialist | AI/ML Researcher

+91 9048357172 | arnolps172@gmail.com | arnol.in | [LinkedIn](#) | [GitHub](#)

Summary

Senior Data Science Specialist with 5+ years building production AI/ML systems. Expert in multi-agent architectures, hybrid RAG (vector + graph fusion), and LLM-powered platforms with security guardrails. Led patent-pending CV research (92% accuracy). Ships AI systems processing 10K+ daily events at sub-second latency.

Technical Skills

AI/ML: PyTorch, Transformers, Scikit-learn, ONNX, LoRA/QLoRA, vLLM, DINOv2, YOLO, OpenCV | **LLMs & RAG:** GPT-4o, Claude, Gemini, Qwen, LangChain, LangGraph, MCP Protocol, OpenRouter, Cross-Encoders, RRF Ranking | **AI Safety:** LlamaGuard, Presidio, Langfuse, PII Detection, Prompt Injection Prevention | **Databases:** Qdrant, Neo4j, Pinecone, ChromaDB, Elasticsearch, FAISS, PostgreSQL, Redis | **Backend:** FastAPI, Pydantic, AsyncIO, WebSockets, Celery, SQLAlchemy | **Infra:** AWS (Bedrock, EC2, S3), Docker, CUDA, Next.js, React, TypeScript | **Languages:** Python, TypeScript, JavaScript, SQL

Experience

Senior Engineer – Data Science

Aug 2025 – Present

Reflections Info Systems Pvt. Ltd.

Kerala, India

- Built **hybrid RAG platform** fusing Qdrant vector search + Neo4j graph traversal + Reciprocal Rank Fusion with 3-path query routing (Cypher metadata / semantic content / general), serving enterprise sales intelligence with 5-layer security (rate limiting → validation → LlamaGuard → RAG guardrails → PII sanitization)
- Architected **multi-agent log analytics** — 4-agent event-driven pipeline processing **10K+ daily logs** with semantic error grouping (0.85 similarity), anomaly detection (4-point scoring), and 5-workflow decision tree routing (200ms–2s budgets)
- Implemented **real-time WebSocket streaming** with buffer-then-sanitize pattern and **Langfuse observability** (24 trace points) for LLM cost tracking across all pipelines
- Built **VLM document extraction** for 13+ formats with multi-provider failover (OpenRouter → OpenAI → Anthropic → Google) and automatic VLM escalation below OCR threshold
- Deployed **compliance verification** automating 149+ regulatory criteria across 39 articles via multi-agent LangGraph orchestration with bilingual (EN/AR) support
- Built **agentic outreach platform** with custom MCP server (11 tools), 5-phase LLM workflow (discovery → enrichment → scoring → document generation → outreach), and template-driven PPTX/DOCX/XLSX pipeline
- Contributed to **ISO 42001:2023** AI Management System surveillance audit — received Shout-Out award for documentation quality and audit preparedness

Senior Data Scientist – Consultant

Dec 2024 – Jul 2025

Digital University Kerala

Kerala, India

- Led **patent-pending** cattle biometric ID: YOLOv11 → novel wavelet ridge enhancement → DINOv2 1024-dim embeddings → Pinecone vector search — **92% accuracy, 49ms inference, 533ms E2E**
- Built semantic document search with vector DBs; deployed **LLM Q&A** via AWS Bedrock + LangChain; led team of 4

Senior Software Engineer – AI/ML

Jun 2024 – Nov 2024

Techversant Infotech

Kerala, India

- Built **RAG with conversational memory** for multi-turn enterprise interactions; developed face recognition and **YOLO-based proctoring** systems

Senior Engineer – Data Science

Sep 2023 – Jun 2024

Digital University Kerala

Kerala, India

- Led 3-person team building **ML search infrastructure** with Elasticsearch ETL; built AWS backend for “Fun With AI” at Global Science Fest Kerala

Data Analyst

Jun 2021 – Aug 2023

Digital University Kerala

Kerala, India

- Built automated data pipelines and visualizations; optimized DB performance via indexing strategies

Research Fellow

Sep 2019 – Sep 2020

ICFOSS

Kerala, India

- Developed Malayalam Morphological Analyzer and sentiment analysis systems; managed large-scale NLP datasets

Education

M.Sc. Computer Science (Data Analytics) — IITM-K, Kerala

2017 – 2019

B.Sc. Computer Science, Statistics & Mathematics — Kristu Jayanti College, Bengaluru

2014 – 2017

Certifications

Google Data Analytics Professional (Google) | Real-Time Video AI (NVIDIA DLI) | Deep Learning (NVIDIA DLI)