

PHILIP SIMON DEROCK

contact@philipsimonderock.com | [+91 8300057632](tel:+918300057632) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

AI Engineer specializing in Agentic AI, RAG pipelines, Multi-Agent Systems, and LLM-powered products, skilled in Fine-Tuning, AI Alignment, MCP ecosystems, and Applied AI Research. Proven in deploying scalable production AI systems.

EXPERIENCE

- | | |
|--|---------------|
| Blackcoat AI AI Engineer – Full Time (Aug 2025 – May 2026) | Remote |
| <ul style="list-style-type: none">Engineered production-grade Agentic RAG infrastructure for legal research orchestrating LangGraph ReAct multi-agent pipelines with domain-specific chunking across legal contracts and acts, Zilliz-backed hybrid retrieval, and LangChain.Served the entire system via FastAPI and a dedicated FastMCP server, enabling external AI agents to invoke legal intelligence as MCP-native tools. | |
| IISPPR AI Agent – Internship (Jun 2025 – Aug 2025) | Remote |
| <ul style="list-style-type: none">Built AI-powered workflow automations for real-time document processing & agent-based task execution. | |
| CodingJr AI Research – Internship (Apr 2025 – Aug 2025) | Remote |
| <ul style="list-style-type: none">Developed an AI coding copilot using LLM-backed code generation to help developers write cleaner code. | |
| Edunet & AICTE AI & Data Analytics – Internship (Nov 2024 – Dec 2024) | Remote |
| <ul style="list-style-type: none">Optimized diabetic patient healthcare prediction via AI-driven feature engineering and statistical modeling. | |
| Cisco Cybersecurity – Internship (May 2023 – Aug 2023) | Remote |
| <ul style="list-style-type: none">Certified in Cybersecurity Essentials and Network Security Fundamentals. | |

PROJECTS

- | | |
|--|----------------------|
| Lunarbit – Evidence-Verifiable Commerce GraphRAG Neo4j, Python, LangGraph, Cohere, FastAPI, HNSW, BM25 | Link |
| <ul style="list-style-type: none">Built Lunarbit, a finance-first GraphRAG product that reconstructs six years of Swiggy and Zomato commerce into an auditable Neo4j knowledge graph with 454 orders, 48K+ nodes, and 69K+ relationships.Designed agentic, multi-document extraction that connects receipts, invoices, fees, discounts, delivery events, and item histories into temporal fintech intelligence with deterministic reconciliation, privacy-safe identity resolution, and citation-level provenance.Served a LangGraph-orchestrated GraphRAG workflow through FastAPI, combining Neo4j BM25/Lucene, HNSW vector search, Reciprocal Rank Fusion, Cohere Embed/Rerank, and test-driven evaluation gates for grounded multi-hop economic questions. | |
| XPHIL - Governed Clinical-Reasoning AI Python, Gemma 4 E4B, QDORA, Unsloth, GGUF, Ollama, llama.cpp | Link |
| <ul style="list-style-type: none">Fine-tuned Gemma 4 E4B via QDORA (QLORA + DORA) using Unsloth on 37k truth-grounded clinical samples across 11 real-patient dataset families – building a governed clinical reasoning AI with Tamil/Tanglish support, calibrated confidence outputs, and full offline deployment on an Android phone via GGUF quantization with llama.cpp and Ollama. | |
| DEX Jobs - Zero-Cost AI Job Hunting Agent Python, Supabase, Qdrant, Jina, Cerebras, Groq, Mistral | Link |
| <ul style="list-style-type: none">Engineered an asynchronous ReAct orchestration engine with a custom Think-Act-Observe loop (sub-200ms latency) integrating 8 specialized agents, multi-tier memory (Postgres + Qdrant), & dynamic multi-provider LLM routing across Cerebras, Mistral, & Groq ensuring 100% runtime.Built a concurrent job pipeline querying 19 sources with Qdrant vector deduplication, LLM relevance scoring, Jina Embeddings v3, job formatting, & Telegram cover letter generation. | |

TECHNICAL SKILLS

Languages & Core: Python, Java, SQL, FastAPI, REST API, PostgreSQL, Neo4j, System Design, Git, GitHub, GitHub Actions, CI/CD, pytest, Linux, Docker, VPS, Google Cloud, AWS, Firebase, Supabase

Data Science & ML: NumPy, Pandas, scikit-learn, Transformers, Hugging Face, NLTK, Data Curation, Data Annotation, Data Labeling, Data Augmentation, Data Preprocessing, AI Feedback, Medical Image Processing

LLM & Agentic AI: LangChain, LangGraph, LangSmith, LlamaIndex, Pydantic, PydanticAI, OpenAI SDK, OpenAI Agents SDK, CrewAI, AutoGen, Semantic Kernel, DSPy, Strands Agents SDK, Google ADK, ReAct Agent, Browser Agent, Multi-Agent Systems, Sub-Agent Orchestration, Tool Calling, Function Calling, Prompt Engineering, Context Engineering, Agentic RAG, GraphRAG, Hybrid RAG, BM25, Reranking, Reciprocal Rank Fusion (RRF), Knowledge Graph, Grounding & Citation, RAGAS, DeepEval, A/B Testing, Embeddings, Chunking Strategies, Dense Retrieval, Sparse Retrieval, Hybrid Search, Vector Search, Zilliz, Milvus, ChromaDB, Qdrant, Pinecone, Weaviate, MCP, FastMCP, MCP Gateway, A2A, SSE (Server-Sent Events), LLM Guardrails, Prompt Injection Defense, AI Governance, Ollama, vLLM, llama.cpp, GGUF, Google AI Studio (Gemini), Claude, Multi-Provider LLM Routing, LLM API Integration, Windsurf, Cursor

AI Research & Alignment: Full Fine-Tuning, LoRA, QLoRA, DoRA, QDORA, PEFT, Unsloth, Supervised Fine-Tuning (SFT), Reinforcement Learning (RL), RLHF, RLAIF, RLVR, AI Safety, AI Alignment, Safety Evaluation, Agent Benchmarking, Human-in-the-Loop Systems, Multimodal AI, Autonomous Reasoning Systems, LLM Research

EDUCATION

Karunya University (Sep 2021 – May 2025)
B.Tech in Artificial Intelligence and Data Science

Coimbatore, India