Denis Craciun

Software & AI Engineer | +393246014086 | deniscraciun18@icloud.com | [LinkedIn Profile](#) | [GitHub](#)

PROFESSIONAL SUMMARY

AI Engineer with 3+ years of software development experience, merging a BSc in Computer Science with an ongoing remote MSc in Finance (UK). Specializes in Agentic RAG systems and QLoRA fine-tuning within AI R&D, delivering production-grade solutions like in-house Bitcoin Lightning integrations. Combines deep technical expertise in Python and LangChain with financial acumen. Seeking an AI Engineer role to develop scalable, high-impact AI-powered products.

EDUCATION

The Open University, *Master of Science in Finance*, Milton Keynes, United Kingdom — Oct 2025 - Oct 2027
University Of Pisa, *Bachelor of Science in Computer Science,* Pisa, Italy — Sep 2022 - Oct 2025

WORK EXPERIENCE

InnovYou, *Software & AI Engineer*, Remote — Mar 2023 – Present
- Architected and scaled microservices handling millions of payment transactions using Django, Node.js, and containerized infrastructure, optimized API response latency, enabling real-time financial data processing for 100K+ concurrent users.
- Designed fault-tolerant microservice architecture with Docker, Kubernetes-ready deployments, and async job queues (Django-Q/Redis), achieving hight stability across payment processing pipelines while handling 10K+ requests/second under peak load.
- Designed cloud-native microservices architecture with composable integration patterns and cross-system connectors; implemented REST/JSON APIs with OAuth2 security and real-time data synchronization, reducing integration latency while enabling modular tool orchestration across enterprise systems.
- Led AI R&D initiatives by implementing agentic RAG pipelines, QLoRA fine-tuning and MCP, improving chatbot retrieval accuracy and significantly reducing model hallucinations whilst reducing costs thanks to open-source models.
- Refactored legacy codebases to modern TypeScript and encryption, reducing technical debt and accelerating feature releases.

Abivet, *Software Developer Junior*, Remote — Sep 2022 – Mar 2023
- Integrated third-party RESTful APIs into mobile applications, expanding feature capabilities and decreasing hardware costs.
- Optimized application performance and memory usage, reducing crash rates and ensuring stability across legacy devices.
- Designed intuitive UI/UX workflows, improving user navigation flows and decreasing drop-off rates during onboarding.
- Streamlined troubleshooting processes, reducing bug resolution time significantly and improving overall release quality.

PROJECTS

Vivia – AI-Powered Tourism Chatbot
- Built production-grade serving backend for AI platform with LangChain and VectorDB, implementing request routing, context retrieval, and observability hooks for monitoring reliability and SLOs. Deployed self-serve capability enabling multiple teams to manage independent data pipelines and agents; optimized architecture for scalability under high concurrency.
- Implemented RAG and QLoRA fine-tuning with a scalable backend (LangChain, NLP, VectorDB, AI agents) enabling real-time merchant recommendations.

DerpGPT – Conversational AI Model (From Scratch)
- Built and trained a 50M-parameter conversational language model on synthetic persona-based, multi-turn dialogue datasets using PyTorch only.
- Designed the full training pipeline and evaluated model behavior, highlighting challenges in coherence and context retention in smaller LLMs.

Genetic Algorithm – Dot Navigation Simulation
- Created a visual demo showing how a genetic algorithm trains a swarm of agents to navigate toward a goal.
- Implemented evolutionary logic (fitness, selection, mutation) using pure HTML and JavaScript.

Qubit Simulator – Visual Quantum Computing Simulation
- Built an interactive qubit and quantum gate simulator using Python and Pygame with real-time state visualization.
- Enabled hands-on exploration of quantum state changes to illustrate core quantum computing principles.

CERTIFICATES AND QUALIFICATIONS

Quantitative and Qualitative Research in Finance — Oct 2025
English C1 Level — Apr 2021

SKILLS

Languages: TypeScript; Python; C; Java; SQL; Assembly ARMv7; OCaml
Frameworks: React; Angular; Django; FastAPI; Flask; Node.js; Next.js; Flutter; Springboot; Wordpress; Figma;
AI & Agents: Agentic RAG; LangChain; Huggingface; MCP (Model Context Protocol); VectorDB (Pinecone/Chroma); QLoRA Fine-tuning; PyTorch; NLP; Hallucination Mitigation; ETL pipelines; Model Evaluation
Cloud & DevOps: Docker; AWS (S3/EC2 concepts); Git; Observability & monitoring; NoSQL; RESTful API design; Jira;