

Romain Sebire

Portfolio: romainsebire.fr | GitHub | LinkedIn | romain.sebire1@gmail.com

EDUCATION

École des Mines d'Alès (IMT) Master of Science in Engineering, Dept. of Artificial Intelligence & Data Science	2023 – 2026 <i>Alès, France</i>
Federal University of Rio de Janeiro (UFRJ) M.Sc. Academic Exchange in Computer Science and Data Science	Mar. – Aug. 2025 <i>Rio de Janeiro, Brazil</i>
Lycée François Malherbe Classes Préparatoires aux Grandes Écoles (CPGE), MPSI-MP	2020 – 2023 <i>Caen, France</i>

EXPERIENCE

Data, AI & Cybersecurity Engineering Intern (Master's Thesis) Orange France – <i>LLM Fine-Tuning for SAST Code Auditing</i>	Mar. – Sept. 2026 <i>Mougins, France</i>
- Knowledge distillation (Sonnet 4.6 via API) and Chain-of-Thought (CoT) reasoning generation - Built a dataset of 20,000+ vulnerability records (CVEs) using agentic pipelines (LangGraph, MCP) - Fine-tuned Qwen 3.6 27B using LoRA in BF16 on NVIDIA H100 GPUs (GCP) - Deployed optimized inference pipelines and managed LoRA adapters using vLLM on a GCP Virtual Machine Recommendation: Letter of recommendation from internship supervisor (Engagement Jeunes)	
IT Project Management Assistant Intern NTN-SNR – <i>Digitalization of Quality Wall Processes</i>	Nov. – Dec. 2024 <i>Alès, France</i>
- Conducted a full user needs assessment and data flow audit for paper-based quality forms - Created functional wireframes (Figma) and detailed functional specifications to validate user requirements - Wrote specifications and use cases on Azure DevOps to streamline handoff to software developers	

PROJECTS

Secure RAG Chatbot for Portfolio <i>LangChain, FastAPI, Pinecone, Groq, Render, GitHub Actions</i> <i>Deployed and live on my portfolio: romainsebire.fr</i>	2026
- Designed a 2-pass ReAct architecture with targeted vector search to eliminate irrelevant context - Optimized memory usage by executing embeddings via Pinecone API (multilingual-e5) - Implemented UUIDv4 session isolation (anti-IDOR), IP rate-limiting, and Markdown sanitization (anti-XSS)	
Dynamic Hospital Queue Optimization via RL <i>Stable-Baselines3, Maskable PPO, Gymnasium</i> <i>R&D Project – Reinforcement Learning applied to Operations Research for resource allocation</i>	2025
- Built a custom simulation environment and evaluation metrics (wait times, punctuality, service rate) - Optimized action space, reward function design, and observation state representation - Trained in a Docker container (NVIDIA T4 GPU) and evaluated: 92.5% performance score vs. 45% baseline	
Agricultural Disease Detection via Self-Supervised Learning (SSL) <i>Vision Transformers</i> <i>AgVision (ICPR 2026): Beyond Visible Spectrum: AI for Agriculture</i>	2025
- Processed 125 GB of satellite imagery (Sentinel-2) for unsupervised crop disease classification - Benchmarked SSL architectures: SimCLR/ResNet (59.2% accuracy) vs. MAE/ViT-Small (71%) - Fine-tuned Sat-MAE with backbone freezing and Focal Loss (handling class imbalance): reached 87.5% accuracy	

SKILLS & LANGUAGES

Programming Languages: Python, Java, SQL, HTML/CSS
ML Frameworks: PyTorch, Scikit-learn, Transformers (Hugging Face), Stable-Baselines3, Pandas, NumPy
Gen AI & Agentic Systems: LangChain, vLLM, Ollama, Pinecone, QLoRA, MCP, Copilot / Antigravity CLI
MLOps & Cloud: Git, Docker, GCP, Azure DevOps, GitHub Actions, Linux
Languages: French (Native), English (C1), Spanish (B1), Portuguese (A2)