

# Ha-Quang LE

📍 Paris, France | ✉️ [winlp4ever@gmail.com](mailto:winlp4ever@gmail.com) | 🔗 [linkedin.com/in/ha-quang-le](https://linkedin.com/in/ha-quang-le) | 💻 [github.com/winlp4ever](https://github.com/winlp4ever)

## PROFESSIONAL SUMMARY

AI systems engineer and architect with 5+ years of experience designing, deploying, and leading large-scale production AI systems. Currently **Chief AI Officer at Stellia.ai**, a French startup building next-generation educational AI assistants. Proven ability to bridge applied research and scalable infrastructure: retrieval-augmented generation (RAG), agentic reasoning, multi-cloud orchestration, and model serving at scale.

## PROFESSIONAL EXPERIENCE

### Stellia.ai — Paris, France

#### Chief AI Officer (formerly CTO)

*Jun 2019 – Present*

##### Company Overview:

Stellia.ai (previously **ProfessorBob.ai**) builds AI-powered assistants for education, combining advanced document understanding, retrieval, and reasoning capabilities. The company successfully **raised €4 M in funding** in partnership with Innovacom and INCO Ventures — [see announcement](#).

##### Key Achievements:

- **Architecture & Scale:** Designed and implemented the company’s **Graph-RAG question-answering platform**, processing **~2 K concurrent LLM requests** with **1–3 s TTFT**, deployed across multi-cloud (AWS + GCP + Scaleway).
- **Performance & Quality:** Achieved **+5–10% accuracy improvement** over baseline RAG pipelines (LangChain/OpenAI), with significantly better context retention and reasoning depth.
- **Infrastructure & Reliability:** Built distributed AI infrastructure combining **vector and hybrid search** (Elasticsearch, Qdrant, Milvus) with robust orchestration using **Prefect** and **Kubernetes**.
- **Core Systems Development:**
  - Engineered a complete document ingestion and parsing stack preserving layout, tables, and hierarchy using Tesseract, MinerU, and Marker.
  - Developed an agentic AI assistant integrating retrieval, reasoning, and tool-usage (web search, code execution, API calls) with [OpenAI Agents SDK](#) and SmolAgents.
- **Leadership:** Managed and mentored **7 AI engineers and data scientists** from top French programs (École Polytechnique, ENS, Paris-Saclay).
- **Collaboration:** Co-developed a **math reasoning solver** for the University of Arizona, integrated into their e-learning platform.
- **Tech Stack:** Python, Go, TypeScript, React, Elasticsearch, Qdrant, Milvus, PyTorch, TensorFlow, Docker, Kubernetes, AWS / GCP / Scaleway.

## Technicolor R&D — Rennes, France

### Deep Learning Research Intern

Apr 2018 – Sep 2018

Researched neural architectures for **image and audio style-transfer** (Gatys, Fader Networks, Adversarial Autoencoders).

Technologies: PyTorch, TensorFlow, WaveNet, NSynth.

---

## BioSerenity — Paris, France

### Java Backend Developer Intern

Jun 2017 – Sep 2017

Developed **real-time signal processing pipelines** for medical IoT systems, including data filtering and spectrogram generation.

Technologies: Java, AWS, Maven, Spring Boot.

---

## EDUCATION

**Master of Science, Data Science** — *Université Paris-Saclay, France* (2018–2019)

**Engineering Diploma, Cycle Ingénieur Polytechnicien (Machine Learning & Computer Vision)** — *École Polytechnique, France* (2015–2018)

**Bachelor in Mathematics** — *University of Natural Sciences, Hanoi, Vietnam* (2012–2014)

---

## TECHNICAL SKILLS

**AI & ML Systems:** Graph-RAG, Retrieval-Augmented Generation, Agentic AI, LLM fine-tuning, Text Embeddings, NLP, Information Retrieval

**Frameworks:** PyTorch, TensorFlow, Hugging Face, OpenAI Agents SDK, SmolAgents

**Infrastructure & Orchestration:** Docker, Kubernetes, Prefect, Airflow

**Databases & Search Engines:** PostgreSQL, Elasticsearch, Qdrant, Milvus

**Cloud & DevOps:** AWS (EC2, RDS, S3), GCP, Scaleway, Azure, Terraform

**MLOps & Observability:** model registry, eval pipelines, feature store, drift detection, canary deploys, Prometheus, Grafana

**Languages:** Python, Go, Java, TypeScript, Node.js

**Frontend:** React, Zustand, TanStack React Query

**Others:** Linux, Git, CI/CD, REST & GraphQL APIs

---

## HONORS & AWARDS

- 🥇 **1st Prize** — National Mathematics Olympiad for University Students, Vietnam (2014)
  - 🥈 **2nd Prize** — National Mathematics Olympiad for High School Students, Vietnam (2012)
- 

## LANGUAGES

- Vietnamese:** Native
- French:** Fluent
- English:** Fluent