Moncif El-mouden

NLP Researcher | PhD Candidate | Co-founder & Head of AI, HyperSec

Paris, France | moncif.el-mouden@u-pec.fr | moncif.ai
github.com/moncifem | huggingface.co/moncefem | linkedin.com/in/moncif-el-mouden

Profile

Second-year PhD researcher in NLP working on domain-specific intelligence where economics meets law: large-scale structured extraction with constrained decoding, domain ontologies and knowledge graphs, and post-training — RLVR and RLAIF — to push reasoning on verifiable domain tasks. In parallel I co-founded HyperSec, where I lead platform development and research on agentic offensive security. 17 open models and 4 datasets published on Hugging Face.

Education

PhD, AI applied to Economics and Law | Université Paris-Est Créteil (UPEC) Nov 2025 – Present

- **Structured extraction at scale.** Constrained-decoding pipeline over the full French Competition Authority archive: 1,900 decisions converted to clean Markdown and extracted against a ∼50-field typed schema, benchmarked across 18 LLMs on a manually annotated ground truth. Public database and publication in preparation.
- **Ontologies and knowledge graphs.** Legal ontology extraction mapping decisions, undertakings, practices and sectors into a graph, to surface sanction patterns and test graph-based prediction.
- **Reasoning and post-training.** Domain adaptation of open models for economics and law (continued pre-training, SFT, RLVR, RLAIF) for accuracy, auditability and on-device deployment.
- **Deep research agents** automating literature review and empirical workflows. Presented the extraction work at Econom'IA (May 2026).

MEng, Information Systems Engineering (double degree) | EPISEN – UPEC 2022 – 2024
Specialisation: Cloud & Security. Thesis on LLMs in educational systems and the limits of their reasoning.

Engineering degree, Computer Engineering | UIASS University 2021 – 2022

CPGE PCSI–PSI | Preparatory classes for the French *grandes écoles* 2019 – 2021

Additional training. ML Summer School (Cohere Labs, 2026) | TRAIL Research Camp, Université de Lausanne (2026) | PyTorch: Fundamentals (DeepLearning.AI).

Experience

Co-founder & Head of AI | HyperSec Sep 2025 – Present

- Lead platform development and the research on our agentic harness: specialised agents for reconnaissance, scenario and payload construction, and pentest planning, which attack a live application and report only what they actually exploited — with the proof.
- Designed multi-agent systems that reason through exploit paths and find logic flaws (IDOR, BOLA, business-logic abuse) that pattern-matching scanners cannot reach.
- Built real-browser reconnaissance orchestrated on Kubernetes to explore complex web applications at scale, and published a benchmark of 7 models on structured XSS challenges (accuracy, speed, cost).
- From terminal experiments to a top-6 hackathon prototype, a month at STATION F backed by Visionaries Club, and a public launch.

Software Engineer (permanent) | Amundi Technology Dec 2024 – Sep 2025

- Java engineer on a core module of the ALTO platform at one of Europe's largest asset managers: production services, data pipelines and platform infrastructure, with ownership of design and architecture.
- Built *Merlin*, an internal agent over 3M+ lines of compiled documentation to accelerate engineering support, then *Morgan*, a broader agent answering support questions across several teams and projects.

Apprentice Software Engineer | Amundi Technology 2023 – 2024 (12 months)
Java, JAX-RS, Spring Core, Hibernate, SQL, ActiveMQ, Centreon, Splunk, Jenkins, Keycloak (SSO), Maven, WildFly, EJB.

Intern Team Supervisor | GPSEA (Grand Paris Sud Est Avenir) Jan 2025 – Jun 2025
Supervised a team of interns building a RAG chatbot on Qdrant.

AI Platform Developer (IAG4UPEC) | Université Paris-Est Créteil 2024 (5 months)
Led a team building the university's generative-AI assistant, fully self-hosted on UPEC servers — Mistral inference with vLLM on CUDA GPUs, FastAPI, vector retrieval, plus data ingestion and practical evaluation of answer quality.

Software Engineering Intern | Amundi Technology 2023 (3 months)

Full-stack Developer Intern | Information systems 2022 (4 months)

Selected Research & Open-Source

Apertus Nash 8B — post-training for reasoning in economics & mathematics RLAIF, GRPO, vLLM

- Fine-tuned Apertus-8B through three SFT stages then GRPO with an AI judge grounded in code-verified references, on a single NVIDIA DGX Spark. Weights, data and the full recipe are public.
- Served the reward as a hot-reloadable service so the rubric could be fixed mid-run without losing optimiser state: truncated answers fell 44.7% → 4.4%, cost per update \$0.074 → \$0.022.
- Taught the model to think less when it can: 1,152 → 55 thinking tokens on simple requests, while reasoning longer on genuinely hard ones.

Légitus — a French legal assistant that knows when to abstain CPT, SFT, model merging

- Four-stage pipeline on Apertus-8B in 25 hours: continued pretraining on the LEGI corpus, 27k multi-teacher synthetic instructions built around grounded citation and abstention, assistant-only SFT, then a weight merge.
- Dataset and merged model released publicly; built for the first AI hackathon of the French National Assembly.

Memory-LoRA — a hypernetwork that writes repository-specific adapters meta-learning, LoRA

- Emits a LoRA adapter for any GitHub repository in one forward pass — no fine-tuning, no RAG, zero repository tokens at inference. Trained on 2,146 repositories and 55,700 QA pairs.
- −5.25 nats cross-entropy against the frozen model on unseen repositories, winning 9 of 9 benchmark combinations; diagnosed and fixed a conditioning collapse where 64% of every embedding was a shared constant.

LatticeDB — an agent-native database for software teams Rust, CRDTs, formal methods

- Open-source database keeping one permissioned history of a team's work items, docs, specs and deployments, with AI agents bound by the same permissions as people. MVCC, checksummed WAL, CRDTs, PostgreSQL wire protocol.
- Built evidence-first: a deterministic simulator replaying crashes and disk faults across millions of seeds, 53 deliberately planted defects proving the checks can fail, and power-cut recovery verified at 4,729 cut points.

Gemma Miner — autonomous dataset construction published on PyPI

- Agent that turns a one-sentence goal into a typed, research-grade dataset: crawls the source, designs a 20–60 variable codebook, extracts every row under strict type discipline, self-verifies, then exports Parquet and a codebook. Public datasets include CNIL sanctions 2011–2025 and AI clinical trials 2000–2025.

Earlier work

medannot — French clinical NER (GLiNER2 over mDeBERTa-v3), fully offline, with exact offsets. | **Orbit Studio / TRAIL** — interpretable orbital propagation with Kolmogorov–Arnold Networks. | **Smart Farms** (17 months) — fault-tolerant microservice platform, 8-person agile team. | **Smart City** — backend and API lead, intelligent building management.

Awards, Talks & Community

- **1st place** — AI Tinkerers & Anthropic, Paris: an agent reading Oura health data over MCP to coach daily activity.
- **3rd place** — Pioneers Hackathon, STATION F (260 builders): *Agent Honeypots*, prompt-injection detection for LangChain and LangGraph agents.
- **Top 6** — {Tech: Europe} Hackathon: the autonomous pentest agent that became HyperSec.
- **2 of 4 challenges** — Koyeb & Weights & Biases AI Hack Night: a private document RAG on self-hosted Qwen2.5 with vLLM.
- **Speaker** — AI Tinkerers Paris: domain-specific knowledge representation in graphs, and where GraphRAG outperforms classical RAG retrieval. Also built the first **Apple Health MCP server** at AgentHack (Hugging Face, Anthropic & Unaite).

Technical Skills

NLP / ML	Constrained decoding & structured extraction, continued pretraining, SFT, LoRA/QLoRA, RLVR, RLAIF, GRPO, model merging, knowledge graphs & ontologies, GraphRAG, NER, evaluation & reward design
LLM stack	PyTorch, Transformers, TRL, PEFT, vLLM, Outlines, LangChain, LangGraph, Hugging Face Hub, Neo4j, Qdrant, CUDA
Programming	Python, Java, Rust, TypeScript/JavaScript, SQL, Cypher, Bash, PHP
Security	Agentic pentesting, reconnaissance, exploit-path reasoning, IDOR/BOLA, XSS, injection, API security, prompt-injection defence
Backend	Spring (Core, Boot, Data, Security), JAX-RS, Hibernate, FastAPI, Flask, Next.js, React
Infrastructure	Kubernetes, Docker, Jenkins, Git/GitLab, Linux, Nginx, Kafka, RabbitMQ, ActiveMQ, Maven, vSphere, Proxmox
Spoken	Arabic (native) • French (fluent, working & academic) • English (fluent, professional & research)