Etienne Boisson

etienne.boiss@yahoo.fr; Linkedin Github

EDUCATION

EPFL & **Yale**
MSc Thesis — *New Haven,USA, 2024-2025*
MSc of Data Science, 5.5/6.0 GPA — *Lausanne, Switzerland, 2022-2024*
EPFL
BSc of Computer Science / Communication Systems, — *Lausanne, Switzerland, 2018-2022*

EXPERIENCES

Research Engineer, Vijil.ai, Jun 2026 – Present
Research and engineering for the safety, security, and trust of AI agents. Working on formal guarantees methods, and adversarial red-teaming.

Research Associate, LiGHT Lab, **EPFL** May 2024 – Nov 2025
Worked on Apertus-Meditron-FO, a medical finetuning of the first Fully Open Swiss LLM Apertus, and Mediset, a curated medical instruction-tuning dataset. Handling of clusters of A100 GPUs for the training.

MSc Thesis & Postgraduate Associate, LiGHT Lab, EPFL & Yale Oct 2024 – Apr 2025
Master thesis on enhancing safety, instruction-following, and medical capabilities of Meditron LLMs, through supervised-finetuning and preference optimisation techniques. Built comprehensive evaluation pipeline and supervised 4 students. Expansion of the Meditron Protocol to various families of models on a complex infrastructure.

Intern at the International Commitee of the Red Cross, ICRC Jul 2024 - Aug 2024
Co-designed multi-agent architecture for LLM chat capabilities in low-resource settings. Machine learning methods for budget prediction and resource allocation.

Research Engineer Student, MLO Lab, EPFL Sep 2023 – Jul 2024
Core contributor to Meditron-70B medical LLM. Designed instruction-tuning datasets (Scraping, Optical Character Recognition, Synthetic Data Generation), developed multi-agent pipeline, adapter fine-tuning (QLoRA), and Reinforcement Learning by Human Feedback.

Startup: Co-Founder & Founding Engineer @ Deal'icious (**EPFL**) Sep 2023 - Aug 2024
App empowering healthier food choices with affordability. Led vision, architecture, technical decisions and project management.

AI Engineer Intern @ Synature Jul 2023 – Aug 2023
Data processing and Machine Learning and Deep Learning methods on audio data for wildlife monitoring.

PUBLICATIONS

Llama-3-Meditron: An Open-Weight Suite of Medical LLMs Based on Llama-3.1 at GenAI for Health, AAAI 2025
High-performing open-weight medical LLM suite built on Llama-3.1, pre-trained on curated texts, fine-tuned with diverse instructions, and enhanced by MediTree for state-of-the-art medical QA and benchmarks.

GPoet: a Language Model Trained for Rhyme Generation on Synthetic Data at LaTeCH-ClfL 2023
Fine-tuning a GPT-2 model to address text constraints in the context of poem generation.