

CONTACT INFORMATION

LiGHT Lab, EPFL
Lausanne, Switzerland
e-mail: xavier.theimerlienhard@gmail.com
web: xaviertheimerlienhard.com · [Scholar](#) · [GitHub](#) · [HuggingFace](#)

RESEARCH INTERESTS

My research focuses on making clinical large language models safer, more accurate, and more transparent. As lead of the Meditron project at EPFL's LiGHT lab, I develop fully-open medical LLMs whose data, training pipelines, and weights are entirely auditable; most recently MeditronFO, a medical adaptation of the Apertus 8B/70B models trained on Switzerland's Alps supercomputer at CSCS, matching MedGemma on open-ended clinical benchmarks. I work on post-training (SFT, RLVR), safety and evaluation, and large-scale clinical validation with physicians through the MOOVE evaluation platform.

CURRENT

Fellowship PhD Student, EDIC, EPFL (LiGHT Lab) since 08/2026

- Leading the Meditron project: open, safe, and transparent LLMs for clinical decision support
- Building AutoMOOVE, automating large-scale clinical validation of medical LLM outputs

EDUCATION

École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
MSc, Data Science, April 2025

- GPA 5.4/6.0
- Thesis: Enhancing Meditron Capabilities with Synthetic and Reasoning Datasets
- Carried out as visiting researcher at Yale University

École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
BSc, Communication Systems, September 2022

PROFESSIONAL EXPERIENCE

EPFL, LiGHT Lab, Lausanne, Switzerland 04/2025 – 08/2026
AI Research Engineer, Meditron

- Built MeditronFO end-to-end, the first fully-open medical LLMs (post-training of Apertus 8B/70B) matching MedGemma on open-ended clinical benchmarks
- Scaled and debugged training on Alps cluster at CSCS, yielding +10 pp on medical benchmarks
- Released every artifact, weights, datasets, code, and training recipes, making the full pipeline auditable and reproducible

Yale University, LiGHT Lab, New Haven, USA 09/2024 – 04/2025
Visiting Researcher, Meditron

- Implemented GRPO-based RL and reasoning-SFT distillation, driving +30 pp on general benchmarks (MMLU, GPQA) and +9 pp on medical benchmarks
- Shipped a synthetic-data pipeline (DSPy, RAG) producing 10,000+ clinical notes

EPFL, LiGHT Lab, Lausanne, Switzerland 09/2023 – 09/2024
Graduate Researcher, Meditron

- Introduced Meditree, a Tree-of-Thought inference framework lifting diagnostic reasoning +5 pp to GPT-4 level, core contribution to Llama-3-Meditron (AAAI 2025, GenAI4Health)

Irbis Consulting SA, Morges, Switzerland 06/2023 – 09/2023
ML & Software Engineering Intern

- Shipped an Electron app automating tender-document drafting; built a PyTorch captcha solver and retrieval-based web scraper

PUBLICATIONS

Fully Open Meditron: An Auditable Pipeline for Clinical LLMs, preprint, 2026

- First fully-open medical LLM, reaching MedGemma performance on open-ended medical benchmarks

Llama-3-Meditron: An Open-Weight Suite of Medical LLMs, AAAI GenAI4Health, 2025

- Open-weight medical LLMs reaching GPT-4 performance via the Meditree method

GPoeT: A Language Model for Rhyme Generation on Synthetic Data, ACL SIGHUM, 2023

- GPT-2 fine-tuned on poems, producing more rhymes than the baseline

**INVITED
TALKS**

- Meta Open-Source AI Summit, 2024, Menlo Park, California
- Invited speaker on open medical LLMs (Llama-3-Meditron)