Xavier Theimer-Lienhard

Machine Learning Research Engineer – Meditron Core Team - [[website](#)] · [[Scholar](#)] · [[GitHub](#)] · [[HuggingFace](#)] · [[email](#)]

SUMMARY

AI Researcher & Engineer specializing in Safe & Reliable Medical LLMs. Core developer of Meditron (open-weight suite). Research focus: Alignment, Reasoning (System 2), and Uncertainty Quantification. 2+ years' experience on end-to-end scaling: pre/post-training on 100+ GPU clusters, RLHF/DPO/GRPO, and clinical validation in resource-constrained settings.

EXPERIENCE

EPFL & Ariadne Labs (Harvard T.H. Chan School of Public Health) - Research Engineer (LiGHT Lab)

Apr 2025-present

- Led pre-training and post-training of Meditron-3 8B/70B ([HF](#)) with SFT, DPO; launched distributed runs up to 128 nodes/512 GPUs on the [SwissAI](#) cluster.
- Finetuned Apertus into Apertus-Meditron for medicine, +**10%** on medical benchmark (USMLE), data curation.
- Designed evaluation pipelines for clinical tasks: optimized inference with vLLM.
- LLMs **validated by clinicians** in 22+ organizations (Moove project, [jointhemoove.org](#)); feedback used in post-training (DPO).
- **Supervised** 4 MSc projects (distillation, SFT, GRPO, evaluation) on reinforcement learning and reasoning.

Yale University - Postgraduate Research Associate

Sep 2024-Apr 2025

- Finetuned Meditron into Meditron-Reasoning; +**30%** on general reasoning and +**9%** on medical accuracy.
- Built multimodal synthetic data pipeline generating **10k**+ clinical notes. Prompt engineering with direct feedback from clinicians.
- Coordinated EPFL-Yale ops; **supervised** 2 MSc projects (synthetic data, clinician evaluation, clinical notes).

EPFL - Research Assistant, Meditron Core Team (MLO Lab)

Sep 2023-Sep 2024

- Created **Meditree** medical **multi-agent** framework; +**5%** on medical exams; [paper](#), prompt engineering.
- Finetuned Llama-3 into Meditron reaching 81% on medical exams ([Llama story shoutout](#), [paper](#)).

Irbis Consulting SA - ML & Software Engineer

Sep 2023-Sep 2024

- Built app automating bidding document creation and Pytorch captcha solver and scraper.

EDUCATION

- **MSc, Data Science**, Swiss Federal Institute of Technology (EPFL) - Sep 2022 - Apr 2025 - **GPA: 5.35/6.0**
- **BSc, Computer Science**, Swiss Federal Institute of Technology (EPFL) - Sep 2018 - Sep 2022

PUBLICATIONS & TALKS

- Sallinen, A., Solergibert, A. J., Zhang, M., Boyé, G., Dupont-Roc, M., ***Theimer-Lienhard, X.**, Boisson, E., Bernath, B., Hadhri, H., Tran, A., Rabbani, T., Brokowski, T., Meditron Medical Doctor Working Group, **Rudner, Tim G. J.**, and **Hartley, Mary-Anne**. 2025. *[Llama-3-Meditron: An Open-Weight Suite of Medical LLMs](#)*, AAAI 2025 Workshop on Generative AI for Health (GenAI4Health)
- Popescu-Belis, A., Atrio, À. R., Bernath, B., Boisson, E., Ferrari, T., ***Theimer-Lienhard, X**., and Vernikos, G. 2023. *[GPoeT: A language Model for Rhyme Generation on Synthetic Data](#)*, Proceedings of the 7th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (ACL)
- **Invited Speaker, [Meta Open-Source AI Summit 2024](#)**.

SKILLS

ML: PyTorch, RL (DPO/GRPO), Transformers, Finetuning, Quantization, Prompt Engineering, Experimental Design.

Engineering: HPC/Slurm, Distributed Training (FSDP/DeepSpeed), Docker/Containers, Linux, Git, vLLM/Inference Optimization.

Languages: Python, C++, Java, SQL.