Fabrice Lamarche *(He/Him)*

fabrice.lamarche@mila.quebec · linkedin.com/fabrice-lamarche · github.com/fbl10

Education

Ph.D. Candidate, Artificial Intelligence [2024 – present]
Mila — Polytechnique Montréal (LAMA-WeST Lab). Thesis topic: language models, nonparametric memories and synthetic data for medical text processing. Supervisor: Prof. Amal Zouaq.

M.Sc. in Artificial Intelligence — Université de Montréal, DIRO (RALI) [2021 – 2023]
Thesis: *Méthodes d'évaluation en extraction d'information ouverte*. Supervisor: Prof. Philippe Langlais. DIRO Writing Fellowship, 2023.

Bachelor's degree — Mathematics and Computer Science [2018 – 2021]
Data science speciality.

Teaching

Teaching Assistant — Natural Language Processing course, Polytechnique Montréal [Fall 2025]

Publications

[1] UMCR: Universal Medical Concept Recognition through Guideline Integration and Agentic Linking *(preprint).*
F. Lamarche, A. Zouaq. arXiv:2610.17363, 2026.

[2] MedHal: A Synthetic Dataset for Medical Hallucination Detection *(AACL-IJCNLP 2026).*
F. Lamarche*, G. Mehenni*, N. Elhami Fard, O. Rios-Ibacache, L. M. Wang, J. Kildea, A. Zouaq. *Proceedings of AACL-IJCNLP 2026*, Hengqin, China, pp. 4599–4619, 2026. arXiv:2504.08596.
* Equal contribution.

[3] BenchIE$^{\text{FL}}$: A Manually Re-Annotated Fact-Based Open Information Extraction Benchmark.
F. Lamarche, P. Langlais. *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 8372–8394, 2024.

Talks

- **UMCR** — presented at IVADO R3. [December 2026]
- **MedHal** (oral & poster) — AACL-IJCNLP 2026, Hengqin, China. [November 2026]
- **IVADO — AI awareness program**: *The ABCs of AI* and *AI, a risk for our democracy?* [2025 – 2026]
- **BenchIE$^{\text{FL}}$** (poster) — ACL 2024, Bangkok, Thailand. [August 2024]

Professional experience

Scientist in Residence — IVADO [2023 – 2025]
AI Consultant. AI specialist for an EdTech startup: fine-tuning and RAG setup for LLM tutoring.

Data Scientist — Intact Insurance [2024 – 2025]
Redefined the evaluation process for market value prediction; performance analysis and model maintenance.

Data Scientist — Santé Québec [2021 – 2023]
Developed an XGBoost model for hospital stay duration (< 1 RMSE); analysis and visualization of strategic data.

Awards and competitions

- DIRO Writing Fellowship (2023)
- **Datathon** (2022, 2023) — *PolyFinance*: AI and NLP competition applied to finance. Best performing solution for exchange rate prediction (time series) using FinBERT + CNN-LSTM, 0.0004 MSE.

Technical skills

Research methods: evaluation of ML models; LLM fine-tuning, RAG, in-context learning, agentic workflows; RL.
Software engineering: Python (PyTorch, Transformers, vLLM, Ollama).
Infrastructure: Azure ML, Snowflake, SQL. **Languages:** English, French.