

Gabriele Prato

Email: gabriele.prato@mila.quebec

Website: gabprato.github.io

Google Scholar: scholar.google.com/citations?user=9D_y8ssAAAAJ

LinkedIn: linkedin.com/in/gabprato

GitHub: github.com/gabprato

RESEARCH INTERESTS

My research focuses on the fundamental capabilities and limitations of LLMs, particularly around knowledge consolidation and reasoning. I am also increasingly interested in how these systems can support multimodal understanding, continual learning, and long-term memory and planning.

EDUCATION

PhD in Artificial Intelligence — *Mila, Université de Montréal*

Sep 2019 - Dec 2025

Advisors: Sarath Chandar, Alain Tapp

GPA: 4.15 / 4.3

MSc in Artificial Intelligence — *Mila, Université de Montréal*

2019

BSc in Computer Science — *Université de Montréal*

2016

Received Excellence Scholarship

FIRST-AUTHOR PUBLICATIONS

Generating Large-Scale Multilingual Educational Instruction-Tuning Data from Scratch — *Under review*

[Effect of Document Packing on the Latent Multi-Hop Reasoning Capabilities of Large Language Models](#) — *Under review*

[Do Large Language Models Know How Much They Know?](#) — *EMNLP 2024*

[EpiK-Eval: Evaluation for Language Models as Epistemic Models](#) — *EMNLP 2023 (Oral)*

[PatchBlender: A Motion Prior for Video Transformers](#) — *NeurIPS 2022 Workshop*

[Scaling Laws for the Few-Shot Adaptation of Pretrained Image Classifiers](#) — *ICML 2021 Workshop*

[Fully Quantized Transformer for Machine Translation](#) — *Findings of EMNLP 2020*

[Towards Lossless Encoding of Sentences](#) — *ACL 2019*

CO-AUTHORED-PUBLICATIONS

Abbes, I., **Prato, G.**, Fournier Q., Rodriguez F., Boukhary A., Elwood A., Chandar S. [Small Encoders Can Rival Large Decoders in Detecting Groundedness](#) — *Findings of ACL 2025*

RESEARCH EXPERIENCE

Research Intern — *IBM Japan*

August 2025 - Nov 2025

Developed a scalable method for generating instruction-tuning data entirely from scratch for any language–location pair, producing substantially more data than prior methods while preserving diversity, quality, and cultural grounding.

Research Intern — *Microsoft Research Montreal*

June 2024 - Sep 2024

Led a research project on the impact of data segmentation on LLM capabilities, resulting in a paper published at EMNLP.

Associate Researcher — *Huawei Montreal*

Jan 2019 - Dec 2019

Tasked with quantizing to 8 bits both the weights and activations of the Transformer for neural machine translation, without compromising performance. We were the first to achieve this, resulting in a publication at EMNLP.

TEACHING & MENTORING

Teaching Assistant — *Polytechnique Montreal*

Fall 2021, Fall 2023

Designed and graded homework assignments, provided student assistance, and evaluated exams.

Research Mentor — *Mila*

Spring 2024 - Summer 2025

Mentored master's students on their research projects by generating ideas, answering questions, planning research schedules, and guiding them in conducting quality research and writing papers.

TECHNICAL SKILLS

ML Frameworks: PyTorch, Huggingface, DeepSpeed, Numpy, Accelerate (distributed training).

Languages & Tools: Python, C, C++, Java, Git, Docker.

PROFESSIONAL SERVICE

Conference Reviewer — *ICLR, ACL, EMNLP*

2023, 2025, 2026

Reviewed submitted papers, provided constructive feedback, and contributed to the peer-review process for top-tier NLP conferences.

LANGUAGES

Fluent in both English and French.