

ALI RAMLAOUI

PhD student, machine learning for atomistic systems

✉ ramlaoui ali [at] gmail [dot] com | 🌐 ali.ramlaoui.org | 🏠 Ramlaoui | 📍 Paris, France

RESEARCH INTERESTS

Transferable representations for atomistic machine learning: self-supervised pretraining of geometric GNNs, machine-learned interatomic potentials, and the open datasets and data infrastructure needed to train and evaluate them at scale.

EDUCATION

CentraleSupélec (CVN lab), Université Paris-Saclay & Entalpic <i>PhD in Machine Learning (CIFRE industrial PhD)</i> • Advisors: Fragkiskos D. Malliaros (thesis director, CentraleSupélec), Joseph Musielewicz (Entalpic).	Paris, France Jun. 2025 – present
École normale supérieure Paris-Saclay <i>Master's Degree – Mathematics, Vision and Learning (MVA)</i> • Coursework: Advanced Deep Learning, Graph Machine Learning, Optimal Transport, Geometric Data Analysis, Probabilistic Graphical Models, Bayesian Machine Learning.	Gif-sur-Yvette, France Sep. 2023 – Sep. 2024
CentraleSupélec <i>Engineering Degree (MSc) – Mathematics and Data Science</i>	Gif-sur-Yvette, France Sep. 2020 – Sep. 2024
Lycée du Parc <i>Classes Préparatoires (MPSI–MP*) – Mathematics, Physics and Computer Science</i>	Lyon, France Sep. 2018 – Jul. 2020

PUBLICATIONS

- [1] **Atompack: A Storage and Distribution Layer for Read-Heavy Atomistic ML Training Datasets**
Ali Ramlaoui, Daniel T. Speckhard, Sagar Pal, Fragkiskos D. Malliaros, Alexandre Duval, Victor Schmidt
arXiv preprint, 2026 · [arXiv:2606.29975](#)
- [2] **TriForces: Augmenting Atomistic GNNs for Transferable Representations**
Ali Ramlaoui, Alexandre Duval, Hannah Bull, Victor Schmidt, Hugues Talbot, Fragkiskos D. Malliaros, Joseph Musielewicz
International Conference on Machine Learning (ICML 2026) · [arXiv:2605.20581](#)
- [3] **Efficient Cross-Functional Learning for Atomistic Modeling of Materials**
Raphaël Hirsch, Walid Bouainouche, Jibril El hassani, Fragkiskos D. Malliaros, Alexandre Duval, Ali Ramlaoui
ICML 2026 Workshop on AI for Science (AI4Science) · [OpenReview](#)
- [4] **LeMat-Rho: High-Fidelity Charge Density Dataset for Atomistic Materials Modeling**
Mathilde L. D. Franckel, Richard Tran, Martin Siron, Daniel Speckhard, Georgia Channing, Guilherme Penedo, Ali Ramlaoui, Alexandre Duval, Jonathan Schmidt
LeMaterial, Hugging Face, 2026 · [Report](#)
- [5] **LeMat-GenBench: A Unified Evaluation Framework for Crystal Generative Models**
Siddharth Betala, Samuel P. Gleason, Ali Ramlaoui, Andy Xu, Georgia Channing, Daniel Levy, Clémentine Fourier, Nikita Kazeev, Chaitanya K. Joshi, Sékou-Oumar Kaba, Félix Therrien, Alex Hernandez-Garcia, Rocío Mercado, N. M. Anoop Krishnan, Alexandre Duval
NeurIPS 2025 Workshop on AI for Accelerated Materials Design (AI4Mat) · [arXiv:2512.04562](#)
- [6] **Catalyst GFlowNet for Electrocatalyst Design: A Hydrogen Evolution Reaction Case Study**
Lena Podina, Christina Humer, Alexandre Duval, Victor Schmidt, Ali Ramlaoui, Shahana Chatterjee, Joshua Bengio, Alex Hernandez-Garcia, David Rolnick, Félix Therrien
NeurIPS 2025 Workshop on AI for Accelerated Materials Design (AI4Mat) · [arXiv:2510.02142](#)
- [7] **LeMat-Traj: A Scalable and Unified Dataset of Materials Trajectories for Atomistic Modeling**
Ali Ramlaoui, Martin Siron, Inel Djafar, Joseph Musielewicz, Amandine Rossello, Victor Schmidt, Alexandre Duval
NeurIPS 2025 Workshop on AI for Accelerated Materials Design (AI4Mat) · [arXiv:2508.20875](#)
- [8] **LeMat-Bulk: Aggregating, and De-Duplicating Quantum Chemistry Materials Databases**
Martin Siron, Inel Djafar, Ali Ramlaoui, Etienne du Fayet, Amandine Rossello, Edvin Fako, Matthew McDermott, Félix Therrien, Luis Barroso-Luque, Flaviu Cipcigan, Philippe Schwaller, Thomas Wolf, Alexandre Duval
ICLR 2025 Workshop on AI for Accelerated Materials Design (AI4Mat) · [arXiv:2511.05178](#)
- [9] **Improving Molecular Modeling with Geometric GNNs: an Empirical Study**
Ali Ramlaoui, Théo Saulus, Basile Terver, Victor Schmidt, David Rolnick, Fragkiskos D. Malliaros, Alexandre Duval
ICML 2024 Workshop on ML for Life and Material Science (ML4LMS) · [arXiv:2407.08313](#)

RESEARCH EXPERIENCE

Entalpic

Paris, France

PhD Researcher – Machine Learning for Materials (CIFRE)

Jun. 2025 – present

- Designed **TriForces**, which augments geometric GNNs (eSEN, Orb, EquiformerV2) with composition and structure streams and multi-objective self-supervised pretraining for transferable representations [2].
- Built **Atompact**, an append-only, memory-mapped storage layer for atomistic training data: 96× faster shuffled reads and 79% smaller files than ASE LMDB [1].

Machine Learning and Chemistry Scientist

Dec. 2024 – Jun. 2025

- Led **LeMat-Traj**, 120M+ DFT configurations harmonized from Materials Project, Alexandria and OQMD, released with the **LeMaterial-Fetcher** pipeline as part of the LeMaterial open-science collaboration [7].
- Trained and evaluated interatomic potentials (UMA, eSEN, Orb, MACE) on Slurm clusters for Entalpic's internal platform, including teacher–student distillation.

Mila & CentraleSupélec

Gif-sur-Yvette, France

Student Researcher – Geometric GNNs for 3D atomic systems

Sep. 2023 – Jul. 2024

- Benchmarked design choices of geometric GNNs for molecular property prediction; first-author paper at the ICML 2024 ML4LMS workshop [9].

Réseau de Transport d'Électricité (RTE)

Paris, France

Research Intern – Research and Development

Apr. 2024 – Oct. 2024

- Designed a GNN architecture to learn power-flow models from large-scale real grid data, and contributed to an open-source package for training and deploying GNNs on power systems.

CentraleSupélec

Gif-sur-Yvette, France

Research Project – Guaranteed control of RL-piloted systems

Jan. 2022 – Jun. 2022

- Developed an interval-based approach to certify the safety of neural-network controllers and compared it with classical control.

INDUSTRY EXPERIENCE

Veolia Asia

Hong Kong

Machine Learning Intern – Research and Innovation

Jan. 2023 – Jul. 2023

- Modeled a hazardous-waste treatment process and implemented explainable time-series forecasting with process and operations engineers.

Réseau de Transport d'Électricité (RTE)

Paris, France

Data Science Intern – Energy Economics

Jun. 2022 – Dec. 2022

- Developed a model simulating grid-scale batteries to relieve congestion on the electrical network.

OPEN-SOURCE SOFTWARE & DATASETS

Atompact · storage and distribution layer for atomistic ML datasets

github.com/LeMaterial/atompact

LeMat-Traj, **LeMat-Bulk**, **LeMat-Rho** · open materials datasets (LeMaterial)

huggingface.co/LeMaterial

LeMaterial-Fetcher · harmonizing public materials databases

github.com/LeMaterial/lematerial-fetcher

TriForces · code for [2]

github.com/Ramlaoui/triforces

ssync · running and following Slurm experiments from a local workspace

ssync.ramlaoui.org

SKILLS

Machine learning: PyTorch, PyTorch Geometric, JAX, Hydra; geometric GNNs and interatomic potentials (eSEN, Orb, MACE, UMA)

Scientific computing: ASE, pymatgen, HPC and Slurm clusters, Docker, Git, CI/CD, GCP

Languages: French (native), Arabic (Lebanese dialect, native), English (fluent), Wolof (spoken)

HONORS & TALKS

French Mathematical Olympiad – Ranked 1st among all candidates in West Africa.

Veolia – Invited speaker at the group-wide Data Science seminar of Veolia's international data community.