

JEAN MERCAT

AI Senior Research Scientist

I am looking to push the quickly evolving state-of-the-art in machine learning research and application to advance capabilities and understanding of large models, their evaluation, and their training data with impactful applications.

 [Scholar](#) @ jean.mercat@gmail.com 📞 +1 (669)-250-6406 📍 Mountain View, CA

EXPERIENCE

Machine Learning Senior Research Scientist

Toyota Research Institute

📅 Jul 2021 - Present 📍 Los Altos, CA

- **Technical lead** research in LLMs, VLMs, and VLAs
- Pre-training LLMs (up to 7B for 1T tokens), multi-modal models
- Manage interns and university collaborators

📖 **Publications** [ArXiv 2025](#), [ArXiv 2025](#), [ICLR 2025](#), [COLM 2024](#), [NeurIPS 2024](#), [NeurIPS 2024](#), [RSS 2024](#), [NeurIPS 2023](#), [CoRL 2022](#), [ICRA 2022](#), [ICLR 2022](#)

Ph.D. Road Scene Motion Forecasting

Renault & Paris Saclay University

📅 Sep 2017 – Mar 2021 📍 Paris

- Road scene **data analysis** and pre-processing
- **Neural networks** for time series (RNN, CNN, VAE, Attention)
- Research and applications

🏆 **First place** [Argoverse Motion Forecasting Competition](#)
Won twice: [NeurIPS 2019](#) and [CVPR 2020](#)

📖 **Publications** [ArXiv](#), [IQPC 2019](#), [NeurIPS 2019](#), [ICRA 2020](#)

Data Engineer

Inria

📅 2017 (9 months) 📍 Bordeaux

- **C++/PostgreSQL** interface OOP
- **Medical data** processing (DICOM)
- Interdisciplinarity between **computer science** and **medicine**

Research Engineer

Bordeaux University

📅 2016 (12 months) 📍 Bordeaux

- **Image processing** in C++, Qt, Cimg
- Anisotropic annealing **simulation**
- Interdisciplinarity between **computer science** and **chemistry**

📖 **Teaching** 22h Fortran tutorial at ENSEIRB-MatMeca

📖 **Publication** [Macromolecular rapid communications](#)

Research Intern

M2P2

📅 2015 (6 months) 📍 Marseilles

- MPI Fortran HPC
- Lattice Boltzman **simulation**, two-phase fluid and solid interaction
- Interdisciplinarity between **computer science**, **physics** and **biology**

📖 **Publications** [APS 2015](#), [ICTAM 2016](#)

SKILLS

LLM/VLM/VLA

Multi-node training

Machine Learning

Scientific Computing

Technical Lead

Interdisciplinarity

FEAT. PUBLICATIONS

[A Careful Examination of Large Behavior Models for Multitask Dexterous Manipulation](#)
[OpenThoughts: Data Recipes for Reasoning Models](#)

[Should VLMs be Pre-trained with Image Data?](#)

First author, ICLR 2025

[Linearizing Large Language Models](#) First author, COLM 2024

[Language models scale reliably with over-training and on downstream tasks](#)
NeurIPS 2024

[DataComp-LM: In search of the next generation of training sets for language models](#)
NeurIPS 2024

OPEN SOURCE



GitHub

Contributions: [OpenLM](#), [Evalchemy](#)
Pytorch: [Linearizing LLMs](#), [Risk-Aware Prediction](#), [Kalman Forecast](#)
C/C++: [MMG](#), [LBM](#)



[LLM evaluation analysis](#)
[Mamba-7b Mistral-SUPRA](#)
[Risk Aware Prediction](#)

REFERENCES

TRI Managers: [Adrien Gaidon](#), [Blake Wulfe](#)

@ {adrien.gaidon.ctr@tri.global,
blake.w.wulfe@gmail.com}

Ph.D. advisors:

[Guillermo Pita Gil](#), [Guillaume Sandou](#)

@ {guillermo@burro.ai,
Guillaume.Sandou@centralesupelec.fr}

HOBBIES

Paragliding

Audio books

Rock climbing

Science

DIY

Outdoors