

YOANN POUPART

Trainee Researcher & Tech-savvy

Paris, France @ yoann.poupart@ens-lyon.org



RESEARCH EXPERIENCE

PhD Student LIP6, Sorbonne University

Sep. 2024 - Ongoing Paris

- ★ Interpretability of multi-agent systems
- 🔧 Mechanistic Interpretability applied to MARL

Visiting Researcher LIP6, Sorbonne University

Mar. 2024 - Aug. 2024 Paris

- ★ Interpretability of chess agents
- 🔧 Sparse Autoencoders applied to RL

Visiting Researcher Fraunhofer HHI

Oct. 2023 - Mar. 2024 Berlin

- ★ Mechanistic interpretability of chess RL models
- 🔧 Probing, saliency maps, activation patching, etc.

Machine Learning Engineer Esker

Jan. 2023 - Sep. 2023 Lyon

- ★ In Context Learning and Semantic Search
- 🔧 LLM inference, chaining and fine-tuning at scale

ENTREPRENEURSHIP

CTO (Co-Founder) Blockloads

Jun. 2023 - Ongoing Paris

- ★ Customer-driven development of tailored solutions
- 🔧 Cloud infrastructure and MLOps integration
- ➔ Leading software development (soon hiring interns)

PROJECTS

Apart Lab Fellowship Apart Lab

Jan. 2024 - May. 2024 Remote

- ★ Feature analysis of adversarial attacks
- 🔧 Linear probing and contrast activation vectors

Supervised Program for Alignment Research SPAR

Sep. 2023 - Jan. 2024 Lyon

- ★ Failure modes of predictive agents
- 🔧 Bayesian formulation and goal-directedness training

ML for Good EffiSciences

Oct. 2022 Lyon

- ★ Study of the AI Safety main issues
- 🔧 Advanced Deep Learning models

PHILOSOPHY

"Practice makes perfect"

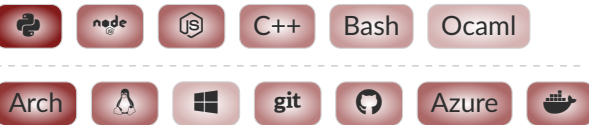
GOALS

- 🗣️ Speak 7 languages fluently
- 🧗 Climb 8A boulders

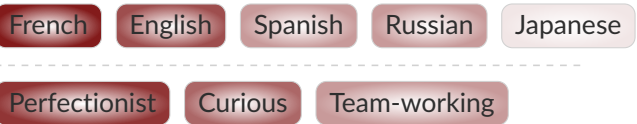
PUBLICATIONS

- ✳ Perspectives for Direct Interpretability in MADRL, AAMAS ALA Workshop, 2025.
- ✳ Contrastive Sparse Autoencoders, ICML MI Workshop, 2024.
- ✳ Topology-Driven Ordering of Flocking Matter, Physical Review X, 2021.

ROBOTS



HUMANS



EDUCATION

Long Research Project

ENS de Lyon Aug. 2024 Lyon

Master in Complex Systems

ENS de Lyon Aug. 2022 Lyon

REFERENCES

Nicolas Maudet
LIP6, Sorbonne University supervisor (2024-2025)
 LIP6 @ nicolas.maudet@lip6.fr

Ekkehard Schnoor
Fraunhofer HHI supervisor (2023-2024)
 HHI @ ekkehard.schnoor@hhi.fraunhofer.de