

# VLADIMIR IVANOV

## EDUCATION

---

### ENS de Paris

Graduate Computer Science

2023-2024

Undergraduate Computer Science & Mathematics

2022-2023

### Lycée Louis-le-Grand

Undergraduate Mathematics & Computer Science & Physics  
(Classe Préparatoire MP2I)

2021-2022

## INTERNSHIPS

---

### Aspiration Learning with Deep Reinforcement Learning

Feb-Jun 2024

- Empirical and experimental study of non optimizing RL algorithms for AI safety.

### Realizability Interpretation of the Countable Axiom Of Choice

Jun-Jul 2023

- Theoretical computer science research.
- Formalization in Coq.

## TRAINING

---

### Silver and bronze IMO (International Mathematical Olympiad) medals

- Other olympiads and many camps.

### ML4Good camp by EffiSciences

- One week long hands on technical AI safety camp.

### 2022 Xena Project Undergraduate Workshop

Imperial College, London

- One week long hands on Lean theorem prover workshop.

## SELECT PROJECTS

---

### Replication of Towards Monosemanticity with Dictionary Learning

- Replication of Anthropic's sparse autoencoder paper on the first layer of TinyStories-1M and reimplementation of the transformer architecture. [GitHub](#)

### Investigation of the Influence of Dropout on Xor Probes in Neural Networks

- Small original interpretability research project, somewhat conclusive results. [GitHub](#)

### PetitC Compiler

- Compiler of a subset of C.
- Semantics formalized and proven to be deterministic in Coq.
- Work in progress towards fully certifying the compiler.
- [GitHub](#)

### Ahlsweide-Zhang Formalized in Lean

- Collaboration.
- GitHub

## **TEACHING**

---

**Math olympiad classes at Animath and Lyon Discrete Math Club.**

**Assistance with AGI Safety Fundamentals at ENS Paris**

## **COMMUNITY ENGAGEMENT**

---

**Member of EffiSciences**

- EffiSciences is a nonprofit effective altruism organisation funded by Open Philanthropy.