

# Javier Abad Martínez

PhD Student in Machine Learning, ETH AI Center

javier.abadmartinez@ai.ethz.ch jaabmar.github.io Google Scholar GitHub LinkedIn

## RESEARCH INTERESTS

---

Language model safety and alignment; privacy-preserving machine learning; causal inference.

## EDUCATION

---

### ETH Zurich

*PhD in Machine Learning, advised by Prof. Fanny Yang*

Dec. 2022 – Jul. 2026

*Zurich, Switzerland*

ETH AI Center Doctoral Fellow; research visit at Simons Institute (UC Berkeley), Fall 2024; GPAI Code of Practice expert group.

### University of Cambridge

*MPhil in Machine Learning and Machine Intelligence; Distinction*

Oct. 2020 – Aug. 2021

*Cambridge, UK*

### Universidad Politécnica de Valencia

*BSc Telecommunications Engineering; BSc Business Administration*

Sep. 2015 – Jul. 2020

*Valencia, Spain*

GPA: Telecommunications 96%, Business 95%; Georgia Tech exchange GPA 4.0/4.0.

## RESEARCH EXPERIENCE

---

### Google DeepMind

*PhD Student Researcher, Gemma Post-training Team*

Oct. 2025 – Dec. 2025

*Zurich, Switzerland*

### Apple

*Machine Learning Researcher, Visual Generation & AI Safety*

Mar. 2025 – Oct. 2025

*Cupertino, CA / Zurich*

### Featurespace

*Research Scientist*

Mar. 2022 – Nov. 2022

*London, UK*

### University of Cambridge

*Research Assistant, Computational & Biological Learning Lab with Adrian Weller MBE*

Oct. 2021 – Feb. 2022

*Cambridge, UK*

### University of Cambridge

*Research Assistant, Van der Schaar Lab with Prof. Mihaela van der Schaar*

Apr. 2021 – Aug. 2021

*Cambridge, UK*

### Universidad Politécnica de Valencia

*Research Assistant, Computer Vision & Behavior Analysis Lab*

Nov. 2019 – Jul. 2020

*Valencia, Spain*

## SELECTED PUBLICATIONS

---

Efficient Randomized Experiments Using Foundation Models

*NeurIPS 2025*

Copyright-Protected Language Generation via Adaptive Model Fusion

*ICLR 2025, Oral*

Privacy-Preserving Data Release via Optimal Transport and Particle Gradient Descent

*ICML 2024*

Detecting Critical Treatment Effect Bias in Small Subgroups

*UAI 2024*

Hidden Yet Quantifiable: Lower Bounds for Confounding Strength

*AISTATS 2024*

Approximating Full Conformal Prediction at Scale via Influence Functions

*AAAI 2023, Oral*

## SELECTED TALKS

---

### Google DeepMind

Inference-time compute for student-teacher distillation

*Zurich, 2025*

### Apple

Pre-training visual generation models

*Zurich, 2025*

### ICLR 2025

Adaptive model fusion for copyright protection

*Singapore, 2025*

### Hasler Foundation

Reliable inference from observational data

*Bern, 2024*

### ETH AI Center

Lower bounds for hidden confounding strength

*Zurich, 2024*

### AAAI 2023

Scaling conformal prediction with influence functions

*Washington DC, 2023*

## SELECTED AWARDS

---

Excellence Fellowship in the Valencian Community

2023

ETH AI Center Doctoral Fellowship

2022

“La Caixa” Excellence Fellowship (declined)

2021

Best Student, Faculty of Telecommunications and Faculty of Business

2021

First place, Spanish national ranking of Telecommunications Engineering students

2021

## SKILLS AND LANGUAGES

---

**Technical:** Python, PyTorch, JAX, TensorFlow, Triton, MATLAB.

**Languages:** Spanish and Catalan (native); English (fluent, C2); German (B1).