

# Edoardo Debenedetti

PHD STUDENT IN CS @ ETH ZÜRICH

☎ (+41) 76 699 43 27 | ✉ [edebenedetti@inf.ethz.ch](mailto:edebenedetti@inf.ethz.ch) | 🌐 <https://edoardo.science> | 📺 [dedeswim](#) | 📞 0000-0003-3343-9477 | 📄 Edoardo Debenedetti

## Education

### ETH Zürich - Federal Institute of Technology Zürich

*Zürich, Switzerland*

PHD IN COMPUTER SCIENCE

08/2022 - 12/2026 (exp.)

- Focus: **Real-world machine learning security and privacy**, advised by **Prof. Florian Tramèr** in the **SPY Lab**.
- **IT Coordinator** for the group: managing the GPU servers and hardware resources.
- Fully funded by the **CYD Doctoral Fellowship**, awarded by the Armasuisse Cyber-Defense Campus.

### EPFL - Federal Institute of Technology Lausanne

*Lausanne, Switzerland*

MSC IN COMPUTER SCIENCE

09/2019 - 04/2022

- GPA 5.63/6, focus on **Machine Learning** ∩ **Security** ∩ **Privacy**.
- Master's Thesis about the **adversarial robustness of Vision Transformers** supervised by **Princeton University's Prof. Mittal**.

### Politecnico di Torino

*Turin, Italy*

BSC IN COMPUTER ENGINEERING

09/2016 - 07/2019

- GPA 28.4/30, graduation mark 110/110, **top 9%**.
- **Exchange year at 同济大学** (Tongji University), in Shanghai (China), supported by a **full scholarship** granted to the top 31% applicants.

## Experience

### Bloomberg LP

*London, United Kingdom*

SOFTWARE ENGINEERING INTERN

07/2021 - 09/2021

- Worked in the **Multi Asset Risk System** team, on the re-design and implementation of the configuration of a distributed logging library.
- Move the configuration of a **distributed logging library** from an internal technology to a **centralized SQL DB**, using a **cache** and a **C++ service**.
- The configuration is checked **~1M times per minute**, and the usage of the cache gave a **~23x speed improvement** w.r.t. querying the DB.

### Armasuisse Cyber-Defence Campus

*Lausanne, Switzerland*

RESEARCH INTERN

08/2020 - 01/2021

- Worked on **Machine Unlearning** and **Membership Inference Attacks** against Generative Models, supervised by **Prof. Mathias Humbert**.
- Adapt the **MIA** technique proposed by the *GAN-Leaks* work (by Chen et al.), to work after the removal some datapoints from the training set.
- The technique achieved **promising results** when attacking DCGAN trained on the CelebA dataset

## Conference papers

- **Debenedetti, E.**, Carlini, N., Tramèr, F., "*Evading Black-box Classifiers Without Breaking Eggs*", 2nd IEEE Conference on Secure and Trustworthy Machine Learning, 2024.
- **Debenedetti, E.**, Sehwag, V., Mittal, P., "*A Light Recipe to Train Robust Vision Transformers*", 1st IEEE Conference on Secure and Trustworthy Machine Learning, 2023.
- Croce\*, F., Andriushchenko\*, M., Sehwag\*, V., **Debenedetti\*, E.**, Flammarion, N., Chiang, M., Mittal, P., Hein, M., "*RobustBench: a standardized adversarial robustness benchmark*", Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track, 2021. (\* equal contribution).

## Manuscript

- **Debenedetti, E.**, Severi, G., Carlini, N., Choquette-Choo, C. A., Jagielski, M., Nasr, M., Wallace, E., Tramèr, F., "*Privacy Side Channels in Machine Learning Systems*", arXiv ePrint 2309.05610.

## Honors and Awards

- 2023 **Oral presentation - AdvML Frontiers Workshop**, Top 10% accepted papers.
- 2023 **CYD Doctoral Fellowship**, full PhD funding for 4 years, worth **USD 516'000** (CHF 461'000), from Armasuisse CYD Campus and EPFL.
- 2021 **Google TPU Research Cloud Program**, extensive **hardware support for 8 months** to work on the Master's Thesis.
- 2021 **Best Paper Honorable Mention - ICLR Workshop on Security and Safety in ML Systems**, top 2 out of 50 accepted papers.

## Teaching

- **Information Security Lab** – ETH Zürich: 2022, 2023 (Teaching Assistant)
- **Large Language Models** – ETH Zürich: 2023 (Teaching Assistant)

## Service

### Reviewer

- **NeurIPS Datasets and Benchmarks Track**: 2022, 2023
- **CCS AI Sec workshop**: 2023

### Open Source Maintainer

- **RobustBench**: adversarial robustness benchmarking library and model zoo.
    - More than 150 models spanning 3 datasets and 3 threat models.
    - **517 stars**, with 379 unique cloners in 2 weeks (measured in September 2023).
    - Refactored the code to improve the extensibility of the library.
- Repository at <https://github.com/RobustBench/robustbench>.

### Conference service

- **Competition organizer at SaTML 2024:** co-organizing the [Large Language Models Capture-the-Flag](#). More than **70 teams and 100 users** signed up.
- **Volunteer at NeurIPS 2021:** helped with monitoring the website and technical issues.

### Invited talks

---

- **ACL SIGSEC** – *Privacy Side-channels in Machine Learning Systems*, 2023.
- **TU Graz EfficientML Reading Group** – *Privacy Side-channels in Machine Learning Systems*, 2023.