

# Javier Carnerero Cano

AI Security PhD Researcher

London, UK  
✉ [j \(dot\) cano \(at\) imperial \(dot\) ac \(dot\) uk](mailto:j(dot)cano(at)imperial(dot)ac(dot)uk)  
📄 [javiccano.github.io](https://javiccano.github.io)  
in [linkedin.com/in/ccano-javi](https://linkedin.com/in/ccano-javi)  
🔗 [scholar.google.com/citations?user=Pk2TMyEAAAAJ](https://scholar.google.com/citations?user=Pk2TMyEAAAAJ)  
🐙 [github.com/javiccano](https://github.com/javiccano)

## About Me

AI Security PhD Researcher at [Imperial College London](#). My current interests are [ML security](#), [GANs](#), and [federated learning](#). I focus on [data poisoning attacks](#), where attackers can manipulate training data collected from untrusted sources to degrade the ML algorithm's performance. I have extensive experience in prototyping [ML algorithms](#) in [Python](#) and [PyTorch](#). I have worked as a [teaching assistant](#) in several courses in [ML](#), [deep learning](#), and [probabilistic methods](#) at [Imperial College London](#). I did a research internship in summer 2022 at [IBM Research Europe - Ireland](#) on [ML security](#). I was included in the [Santander-CIDOB 35 under 35 List](#) in 2021. My background is also in [Telecom Engineering](#). If you want to know fun facts about me, you can have a look at this [video](#).

## Work Experience

- 2018 – pres. **PhD Researcher**, Machine Learning Security, [Imperial College London](#).
- 2019 – 2022 **Teaching Assistant**, Dept. of Computing, [Imperial College London](#). Courses: [Introduction to ML](#) (22/23 and 19/20), [Mathematics for ML](#) (21/22 and 19/20), [Probabilistic Inference](#) (20/21), [Reinforcement Learning](#) (20/21), and [Deep Learning](#) (19/20).
- 2022 **Research Intern**, AI Security and Privacy, [IBM Research Europe - Ireland](#).
- 2017 – 2018 **Intern**, Data Engineering, [Santander Digital Services Spain](#).
- 2016 – 2017 **Research Assistant**, RF, Antennas, and Sensors, [Universidad Carlos III de Madrid](#).

## Education

- [exp.] 2023 **PhD in Machine Learning Security**, [Imperial College London](#).
- 2017 **MRes in Multimedia and Communications**, [Universidad Carlos III de Madrid](#).
- 2017 **MSc in Telecommunications Engineering**, [Universidad Carlos III de Madrid](#).
- 2015 **BEng in Telecommunications Engineering**, [Universidad Carlos III de Madrid](#).

## R&D Interests

- [ML](#), [Deep Learning](#), and [Adversarial ML](#).
- [Data Poisoning](#), [Bilevel Optimization](#), and [GANs](#).
- [Federated Learning](#).

## Computer Skills

- **Prog. lang.:** [Python](#), [MATLAB](#), [Java](#), and [C](#).
- **Python ML Frameworks:** [PyTorch](#), [NumPy](#), [Scikit-learn](#), and [TensorFlow](#).
- **Databases:** [SQL](#).

## Languages

English **full professional proficiency**  
Spanish **native**

## Awards and Grants

- 2022 **Top Talent**, [Nova](#).
- 2022 **Alumni Excellence Award**, [Universidad Carlos III de Madrid](#).
- 2021 **35 under 35 List**, [Santander-CIDOB](#): brings together 35 potential minds of 35 or less years of age which are experts on the global digital order, algorithmic governance and AI.
- 2020 **Best Poster Award**, [Machine Learning Summer School Indonesia](#).
- 2018 **PhD Scholarship**, [Defence Science and Technology Laboratory \(Dstl\)](#).
- 2016 **MSc Research Scholarship**, [Universidad Carlos III de Madrid](#).
- 2014 – 2016 **Tuition-fee Scholarships**, [Spanish Ministry of Education](#).

---

## Selected R&D Projects

- 2018 – pres. **Evaluating the Robustness of Machine Learning Algorithms in Adversarial Settings**, funded by [Dstl](#), in collaboration with [Imperial College London](#).
- 2022 **Machine Unlearning under Data Poisoning**, in collaboration with [IBM Research](#).
- 2017 **Development of a Multiband Feeder with Autotracking Capability**, funded by [Prodetel](#), in collaboration with [Universidad Carlos III de Madrid](#).

---

## Community Service

### Public Engagement

- 2023 “Defense Against the Dark Arts and Potions: AI Models Can Be Easily Poisoned”, [T3chFest](#). [\[Link\]](#).
- 2022 **DoC Clock**: video series which features some of the work and an insights into the personality of PhD students in the Dept. of Computing, [Imperial College London](#). [\[Link\]](#).

### Invited Talks

- 2023 “Machine Learning Models Can Be Easily Poisoned (But Not All Is Lost)”, [Universidad Carlos III de Madrid](#).
- 2022 “Machine Learning Models Can Be Easily Poisoned (But Not All Is Lost)”, [Universidad Pontificia Comillas](#).

### Mentoring Assistance

- 2022 – pres. **PhD Buddy**, [Imperial College London](#).
- 2022 – pres. **Alumni Mentor**, [Universidad Carlos III de Madrid](#).
- 2018 – 2022 Assisted in the supervision of 2 MSc (one of them awarded “Distinguished” status), 1 MEng, and 1 Undergraduate Research Opportunities Programme (UROP) student research projects, and 1 group project (5 students) [\[Link\]](#) on data poisoning attacks against machine learning, [Imperial College London](#).

### Peer Review of Conference Papers

*AISTATS, NeurIPS, CPSIoTSec at CCS, AIsSec at CCS, and MLCS at ECML PKDD.*

### Peer Review of Journal Papers

*IEEE OJSP, IEEE TIFS, and EURASIP JIS.*

---

## Selected Publications (Full List [\[Here\]](#))

### Journal Papers

- 2023 **J. Carnerero-Cano, et al.**, “Hyperparameter Learning under Data Poisoning: Analysis of the Influence of Regularization via Multiobjective Bilevel Optimization”, under review in *IEEE Transactions on Neural Networks and Learning Systems*. [\[Link\]](#).
- 2018 **J. Carnerero-Cano, et al.**, “A Contactless Dielectric Constant Sensing System Based on a Split-Ring Resonator-Loaded Monopole”, *IEEE Sensors Journal*, vol. 18, no. 11, pp. 4491–4502. [\[Link\]](#).

### Conference and Workshop Papers

- 2021 **J. Carnerero-Cano, et al.**, “Regularization Can Help Mitigate Poisoning Attacks... with the Right Hyperparameters”, in *ICLR Workshop on Security and Safety in Machine Learning Systems*. [\[Link\]](#).

### Papers in Preparation

- [in prep.] L. Muñoz-González, B. Pfitzner, M. Russo, **J. Carnerero-Cano**, and E. C. Lupu, “Poisoning Attacks with Generative Adversarial Nets”, in *arXiv preprint arXiv:1906.07773*. [\[Link\]](#).