

Ilias Tsingenopoulos

Postdoctoral Researcher at KU Leuven
Leuven, Belgium, 3000

itsiggen@gmail.com
ilias.tsingenopoulos@cs.kuleuven.be
LinkedIn: [linkedin.com/in/itsiggen](https://www.linkedin.com/in/itsiggen)
+32 486 319 227

ABOUT ME

I am a postdoctoral researcher specializing in the intersection of Artificial Intelligence and Computer Security, with a proven track record in (Adversarial) Machine Learning and Reinforcement Learning. My research spans the theoretical and practical aspects of adversarial attacks and defenses across a broad range of AI systems and modalities: from circumventing web bot detection systems such as Google reCaptcha, to hardening commercial antivirus against adversarial malware, and more generally on adaptively optimizing attacks and defenses against each other under the competitive game they form.

I investigate the fundamentals of robust learning and its adversarial and counterfactual aspects, essential components towards achieving trustworthy and safe AI. As human decision-making increasingly transitions to automated, I am actively developing new techniques and methodologies for training models resilient to evasion and other failure modes. Currently, I focus on safeguarding LLM generation through the integration of alignment, guardrails, and other safety mechanisms to ensure ethical and reliable outputs in real-world deployments.

PROFESSIONAL & RESEARCH EXPERIENCE

DistriNet, KU Leuven

Postdoctoral Researcher in AI Safety and Security

- Exploring principled approaches for safe LLM generation.

Leuven, Belgium

October 2024 – Present

DistriNet, KU Leuven

Doctoral Researcher in Adversarial Machine Learning (AML)

- Adversarial attacks and defenses on AI models.

Leuven, Belgium

May 2019 – October 2024

S2Lab, University College London

Visiting Researcher

- Rendering a commercial antivirus robust to evasive malware.

London, UK

January, 2023 – April 2023

DistriNet, KU Leuven

Research Assistant

- Preventing abusive DNS registrations in the .eu domain.
- General black-box attack optimization.

Leuven, Belgium

May, 2018 – April 2019

Information Technologies Institute, CERTH

Research Assistant

- Formulating and writing research proposals.
- Research and implementation in several Horizon 2020 projects.

Thessaloniki, Greece

January, 2017 – April 2018

EDUCATION

Aristotle University of Thessaloniki

Diploma in Electrical and Computer Engineering - 7.54/10

Dissertation: Fuzzy Clustering Algorithms in Feature Subspaces

Thessaloniki, Greece

Sep. 2006 – Jul. 2016

Technical University of Berlin

Erasmus in Electrical Engineering and Computer Science Department

Berlin, Germany

Feb. 2012 – Aug. 2012

High School

Arsakeio Thessalonikis – Final GPA: 19.3/20

Thessaloniki, Greece

Sep. 2002 – Jun 2005

SELECTED PUBLICATIONS

- **I. Tsingenopoulos**, J. Cortellazzi, B. Bosansky, S. Aonzo, D. Preuveneers, W. Joosen, F. Pierazzi, and L. Cavallaro. “How to Train your Antivirus: RL-based Hardening through the Problem-Space”. In: 27th International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2024).
- **I. Tsingenopoulos**, V. Rimmer, D. Preuveneers, F. Pierazzi, L. Cavallaro, and W. Joosen. “On Adaptive Decision-Based Attacks and Defenses”. In: 7th Deep Learning Security and Privacy Workshop (DLSP 2024).
- **I. Tsingenopoulos**, D. Preuveneers, L. Desmet, and W. Joosen. “Captcha me if you can: Imitation Games with Reinforcement Learning”. In: 2022 IEEE 7th European Symposium on Security and Privacy (EuroS&P 2022).
- **I. Tsingenopoulos**, A. M. Shafiei, L. Desmet, D. Preuveneers, and W. Joosen. “Adaptive Malware Control: Decision-Based Attacks in the Problem Space of Dynamic Analysis”. In: 1st Workshop on Robust Malware Analysis (WoRMA 2022).
- C. J. Hernández-Castro, Z. Liu, A. Serban, **I. Tsingenopoulos**, and W. Joosen. “Adversarial Machine Learning”. In: Security and Artificial Intelligence: A Crossdisciplinary Approach. Springer, 2022.

KNOWLEDGE & TECHNICAL SKILLS

Domain Expertise:

- **Mathematical:** Linear Algebra, Non-convex/Derivative-free Optimization
- **Machine Learning:** Clustering, SVMs, Decision Trees/Random Forests, XGBoost
- **Deep Learning:** Convolutional/Graph Neural Networks, GANs
- **Reinforcement Learning:** Q-Learning, Policy Optimization, Multi-agent
- **Adversarial ML:** White/Black-box attacks, Adversarial Training

Development and Tools:

- **Programming Languages:** Python, C, C++, Java, Matlab
- **Deep Learning:** PyTorch, Tensorflow, Keras
- **Reinforcement Learning:** Gym/Gymnasium, Stable Baselines, RLLib
- **Scientific Computing Libraries:** Numpy, Scipy, Pandas, Scikit-learn, openCV
- **Others:** Git, LaTeX

SUMMER SCHOOLS & ONLINE COURSES

- **Deep Reinforcement Learning:** CS 285 Fall 2021 (UC Berkeley)
- **M2L 2023:** Mediterranean Machine Learning Summer School
- **Security and Privacy in the Age of AI:** Participated/organized editions 2022 & 2023

LANGUAGES

Greek: Native | English: Excellent - C2 | German: Very Good - B2/C1 | Dutch: Good - B1

OTHER INTERESTS

A curious and restless spirit, in my free time I enjoy reading and (computer) games; I am also a competitive capoeirista and water polo player.