

# Paul-Antoine Le Tolguenec

PHD · MACHINE LEARNING/REINFORCEMENT LEARNING

☎ (+33) 612245893 | ✉ letolguenecpa@gmail.com | 🏠 paul-antoineletolguenec.github.io | 🌐 Paul-antoineLeTolguenec

## Summary

---

PhD in Deep Reinforcement Learning (CIFRE Airbus/ISAE-SUPAERO), NeurIPS and ACM Distinguished Paper Award. Research on robust RL, exploration, and sim-to-real transfer, with hands-on robotics experience (ROS, Gazebo, UAV on Nvidia Jetson).

## Work Experience

---

### Mirakl (ADS)

APPLIED SCIENTIST

*Remote, France*

*Jul. 2025 - Present*

- Evaluated and improved a multi-armed bandit selecting among recommendation algorithms in production.
- Managed end-to-end training pipelines for multi-modal models (SigLIP, CLIP) on GPU clusters.
- Maintained and evolved the production ML stack (pipelines, monitoring, deployments).

### ISAE-SUPAERO

ASSOCIATE PROFESSOR / RESEARCH SCIENTIST

*Toulouse, France*

*Feb. 2025 - Jul. 2025*

- Designed and delivered a graduate hands-on course on deploying RL policies on real robots using the [Duckietown](#) platform: sim-to-real transfer and on-robot evaluation.
- Taught reinforcement learning ([course](#)), diffusion processes ([slides](#)), evolutionary algorithms, and machine learning.

### Airbus Commercial Aircraft France / ISAE-SUPAERO

PHD IN DEEP REINFORCEMENT LEARNING

*Toulouse, France*

*Dec. 2021 - Feb. 2025*

- Developed novel exploration and diversity-driven policy search algorithms benchmarked on MuJoCo locomotion and manipulation tasks (NeurIPS 2024).
- Applied adversarial and evolutionary search to identify worst-case scenarios and quantify policy robustness (ACM Distinguished Paper Award).
- Scaled distributed GPU training across MuJoCo and JSBSim simulators.

### MBDA France

RESEARCH INTERN IN MACHINE LEARNING

*Paris, France*

*Mar. 2021 - Oct. 2021*

- Implemented deep RL algorithms for autonomous decision-making in simulation.
- Designed multi-agent planning in adversarial POMDP settings.

### LIRMM

RESEARCH INTERN IN ROBOTICS

*Montpellier, France*

*May 2020 - Aug. 2020*

- Built a Gazebo simulation environment and ROS control/navigation stack for a mobile robot in karst terrains.
- Designed an interval-analysis framework for uncertainty quantification.

### Lab-STICC

RESEARCH INTERN IN ROBOTICS AND COMPUTER VISION

*Brest, France*

*Dec. 2018 - Feb. 2019*

- Implemented path planning, low-level hardware interfaces, and OpenCV-based perception on an autonomous robot.
- Used ROS and Gazebo for simulation and on-hardware deployment.

## Education

---

### ISAE-SUPAERO

PHD IN DEEP REINFORCEMENT LEARNING

*France*

*Feb. 2022 - Feb. 2025*

- PhD research in deep reinforcement learning (exploration and robustness evaluation).
- Published peer-reviewed papers in machine learning venues.
- Teaching assistantship in computer science and machine learning.

### ENSTA

MSC IN COMPUTER SCIENCE AND ENGINEERING

*France*

*Sep. 2018 - Sep. 2021*

- Core training in probability, statistics, and optimization, with a strong focus on machine learning and robotics.
- Software engineering, ML and robotics stack: Python, Gazebo, C++, Docker, PyTorch/TensorFlow, ROS.

### Jean Mermoz

PREPARATORY CLASSES FOR FRENCH ENGINEERING SCHOOLS (CPGE)

*France*

*Sep. 2016 - Aug. 2018*

- Intensive coursework in mathematics, physics, and engineering sciences.
- Research project on thermal insulation using silica aerogel.

# Publications & Pre-prints

---

## Exploration by Running Away From the Past

Pre-print

LE TOLGUENEC, PAUL-ANTOINE, WILSON, DENNIS G., BESSE, YANN, RACHELSON, EMMANUEL

2024

- Proposes RAMP, a method encouraging agents to explore by distancing themselves from past experiences.
- [Paper Link](#)

## Exploration by Learning Diverse Skills through Successor State Measures

NeurIPS 2024

LE TOLGUENEC, PAUL-ANTOINE, WILSON, DENNIS G., BESSE, YANN, RACHELSON, EMMANUEL

2024

- Introduces LEADS, creating diverse skills to enhance exploration without rewards or bonuses.
- [Paper Link](#)

## Exploration-Driven Reinforcement Learning for Avionic System Fault Detection

ISSTA 2024: **ACM SIGSOFT Distinguished Paper Award**

LE TOLGUENEC, PAUL-ANTOINE, RACHELSON, EMMANUEL, BESSE, YANN, WILSON, DENNIS G.

2023

- Uses heuristic-based testing to identify failure scenarios, enhancing flight system reliability.
- [Paper Link](#)

## Curiosity Creates Diversity in Policy Search

ACM Transactions on Evolutionary Learning

LE TOLGUENEC, PAUL-ANTOINE, RACHELSON, EMMANUEL, BESSE, YANN, WILSON, DENNIS G.

2023

- Introduces Curiosity-ES, using curiosity as a fitness metric to explore reward-sparse environments.
- [Paper Link](#)

# Projects

---

## genesis-robust-rl

Personal project

ROBUST RL LIBRARY FOR QUADROTOR CONTROL

2026 - Present

- Open-source Gymnasium-compatible library for robust RL on the [Genesis](#) simulator:  $\sim 70$  perturbations (physics, motor, sensor, wind, delays, payload) with Lipschitz-continuous dynamics.
- Adversarial API and training loop supporting domain randomization, RARL, and RAP, with a PPO reference implementation.
- [GitHub](#)

## trackmania-il

Personal project

IMITATION LEARNING + RL FINE-TUNING ON TRACKMANIA

2026

- Behavior cloning with flow matching on expert trajectories, followed by RL fine-tuning on TrackMania Nations Forever.
- End-to-end pipeline: data collection, offline pretraining, and online policy improvement.
- [GitHub](#)

## parlobot

Personal project

LANGUAGE-CONDITIONED MANIPULATION WITH VLM REWARDS

2026

- Text-conditioned RL policy for a Franka Panda arm in Genesis, trained via contrastive C-Learning.
- Frozen VLM encoders (ViCLIP, InternVideo2) provide a shared video-text embedding used for reward computation and goal conditioning on reach/push tasks.
- [GitHub](#)

## Forsea Robotics – ENSTA

Brest, France

UAV CONTROL USING DEEP REINFORCEMENT LEARNING

Sep. 2020 - Mar. 2021

- Implemented DDPG for autonomous UAV docking with a CNN+LSTM visual controller.
- Deployed on Nvidia Jetson with ROS; validated in Gazebo and on hardware.
- [GitHub](#)

## HuggingFace

Toulouse, France

HUGGINGFACE LE ROBOT HACKATHON

Apr. 2025

- Assembled and programmed a robotic arm; trained a manipulation policy with TD-MPC (model-based planning).