Daniel Marta

ETH Zurich
Zurich, Switzerland
Postdoctoral Researcher | dsimoes@ethz.ch | daniellsm.github.io | daniel-marta | daniellsm | Daniel Marta

Research Statement

I am a postdoctoral researcher at ETH Zurich working on vision-language-action (VLA) models, uncertainty-aware reward and policy learning with large vision-language model (VLM) backbones, reinforcement learning, and large language models. I envision a world where robots and artificial agents interact safely and adapt seamlessly to humans through diverse feedback while knowing when their actions may be unreliable. My research focuses on aligning autonomous systems with human and environmental feedback: preference-based reinforcement learning for safe and sample-efficient robot adaptation [J1, C1–C5, C7], dense credit assignment from rich textual feedback for efficient LLM post-training [C6], and uncertainty quantification for failure-aware and sample-efficient VLA adaptation [S1].
Keywords: Vision-Language-Action (VLA) Models, Uncertainty Quantification, Reward Models, Vision-Language Models (VLMs), Reinforcement Learning (RL), Large Language Models (LLMs), RL from human feedback (RLHF)

Selected Publications

C=Conference, J=Journal, S=In Submission, T=Thesis, *=equal contribution

[J1] **Marta, D.**, Pek, C., Melsión, G. I., Tumova, J., & Leite, I. (2021). Human-feedback shield synthesis for perceived safety in deep reinforcement learning. IEEE Robotics and Automation Letters, 7(1), 406-413.

[C1] **Marta, D.**, Holk, S., Pek, C., Tumova, J., & Leite, I. (2023, May). Aligning human preferences with baseline objectives in reinforcement learning. In 2023 IEEE International Conference on Robotics and Automation (ICRA) (pp. 7562-7568). IEEE.

[C2] Holk, S., **Marta, D.**, Pek, C., & Leite, I. (2024, May). POLITE: Preferences combined with highlights in reinforcement learning. In 2024 IEEE International Conference on Robotics and Automation (ICRA) (pp. 2288–2295). IEEE.

[C3] Holk, S.*, **Marta, D.***, & Leite, I. (2024, March). PREDILECT: Preferences Delineated with Zero-Shot Language-based Reasoning in Reinforcement Learning. In Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (pp. 259-268).

[C4] **Marta, D.***, Holk, S.*, Pek, C., & Leite, I. (2024, May). SEQUEL: Semi-Supervised Preference-based RL with Query Synthesis via Latent Interpolation. In 2024 IEEE International Conference on Robotics and Automation (ICRA) (pp. 9585-9592). IEEE.

[C5] **Marta, D.***, Holk, S.*, Vasco, M., Lundell, J., Homberger, T., Busch, F., Andersson, O., Kragic, D., & Leite, I. (2025). FLoRA: Sample-Efficient Preference-based RL via Low-Rank Style Adaptation of Reward Functions. In 2025 IEEE International Conference on Robotics and Automation (ICRA).

[C6] Hübotter, J., Lübeck, F., Behric, L. D., Baumann, A., Bagatella, M., **Marta, D.**, Hakimi, I., Shenfeld, I., Kleine Buening, T., Guestrin, C., & Krause, A. (2026). Reinforcement Learning via Self-Distillation. In International Conference on Machine Learning (ICML). Also accepted to the ICLR Workshop on Scaling Post-training for LLMs (SPOT).

[C7] **Marta, D.***, Holk, S.*, & Leite, I. (2026). MOSAIC: Multi-objective optimization from zero-shot language reasoning in preference-based RL. In 2026 IEEE International Conference on Robotics and Automation (ICRA).

[S1] Römer, R., Seeliger, M., Liu, S., Sturgis, B., Bagatella, M., **Marta, D.**, Krause, A., & Schoellig, A. P. (2026). Uncertainty Quantification for Flow-Based Vision-Language-Action Models. arXiv preprint arXiv:2606.18043.

Experience

- **Postdoctoral Researcher** *2025 - Present*
  ETH Zurich Zurich, Switzerland
  - Conducting research on vision-language-action (VLA) models and reward models with large vision-language model (VLM) backbones.
  - Developing reinforcement learning and large language model methods for robotic and agentic systems.
- **Visiting Researcher** *2024 - 2024*
  Interactive Machines Group, Computer Science Department, Yale University United States
  - Explored enhancing interpretability of human feedback when modeling reward functions in the context of RLHF.
- **Research Engineer** *2018 - 2018*
  RL Team, Waymo Research United Kingdom
  - Implemented distributed RL agent learning in Docker containers for self-driving car simulations.
- **Research Intern** *2017 - 2018*
  RL Team, Waymo Research United Kingdom
  - Implemented RL algorithms such as DDPG and DQN.
- **Graduate Research Fellow** *2016 - 2017*
  GAIPS, INESC-ID Portugal
  - Researched unsupervised learning methods within RL contexts.

Education

- **PhD in Computer Science** *2020 - 2025*
  RPL (Robotics, Perception and Learning) Department, KTH Royal Institute of Technology Stockholm, Sweden
  - Completed PhD in 2025.
  - Enhanced Preference-based RL methods for robot control tasks which can adapt to human feedback.
  - Devised RLHF methods to enhance query information for human-robot interaction.
  - Supervisors: Prof. Iolanda Leite, Prof. Jana Tumova.
- **Doctoral Researcher** *2019 - 2020*
  ATO (Air Transport and Operations) Department Delft, Netherlands
  - Investigated combinatorial optimization in RL to predict aircraft scheduling and passenger preferences for air transport operations.
- **B.Sc. & M.Sc. in Aerospace Engineering, Avionics (Grade: 16/20)** *2009 - 2016*
  Technical University of Lisbon Lisbon, Portugal
  - Thesis: *Deep Learning Methods for Reinforcement Learning*.
  - Supervisor: Prof. Rodrigo Ventura.

Honors and Awards

- **Best Paper Awards**
  - Best student paper award for an outstanding contribution to the WACV 2024 LLVM-AD workshop.
- **Nominations**
  - Paper C2 was a finalist for the Best HRI Paper Award, Best Student Paper Award, and Best Paper Award at the 2024 IEEE International Conference on Robotics and Automation (ICRA).
- **Grants**
  - Awarded a PhD Grant for the WASP doctoral program from the Wallenberg foundation, which is the largest individual research program in Sweden.
  - Awarded a Research Exchange Grant, to visit Yale University in the spring of 2024.

Teaching and Supervision

- **Teaching** *2020 - 2024*
  DD2380 AI course
  - Teaching assistant and leader of the RL assignment; lectured RL tutorials and coordinated teams of teaching assistants to deliver lectures and in-person examinations.
  - Developed and maintained the RL assignment and the dockerized evaluation process.
- **Supervision**
  - Oskar Stigland, Year:2024, Thesis: Explainability and Hierarchical Reinforcement. Learning
  - Mohammed Akif, Year:2023, Thesis: Safe Reinforcement Learning for Social Human-Robot Interaction

Skills

- **Programming & Scripting Languages:** Proficient in Python (socket networking, multi-processing/threading), Moderate experience with C++, Javascript, SQL, XML/XSL, LaTeX.
- **Machine Learning Libraries:** Proficient in PyTorch (current research focus) and TensorFlow (distributed learning in previous research engineer role).
- **Simulation Engines:** Implemented several RL environments in Unity, familiar with socket APIs for Python, moderate experience in Unreal Engine.
- **Containerization & Deployment:** Proficient in Docker, moderate experience in Kubernetes.
- **Web Development Frameworks:** Moderate experience with Angular.
- **Other Tools & Technologies:** Moderate experience in SolidWorks, minor experience with Blender and Photoshop.

Academic Service

- **IEEE International Conference on Robotics and Automation (ICRA)** *2020 - 2024*
  Reviewer
- **IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)** *2022, 2023*
  Reviewer
- **IEEE Robotics and Automation Letters** *2022, 2023*
  Reviewer
- **ACM/IEEE International Conference on Human-Robot Interaction (HRI)** *2023, 2024*
  Reviewer