Luca Morlok

PhD Student | AI Alignment, Continual Learning & Distillation

lmorlok.github.io
luca.morlok@icloud.com

EDUCATION

International Max Planck Research School for Intelligent Systems, PhD Student **2025 on**
AI alignment, continual learning and distillation, advised by Prof. Dr. Rediet Abebe

Stanford University, Visiting Student Researcher **2024**
Stanford Trustworthy AI Research (STAIR) Lab
- MS Thesis: *Towards Active Learning for Large Language Model (LLM) Alignment* - Aligning LLMs to Human Preferences using Reinforcement Learning, Bandit theory and Active Learning techniques known from Bayesian Optimization (grade: 1.3)

Technical University of Munich, M.Sc. **2021 - 2025**
School of Computation, Information and Technology
Major: Robotics, Cognition, Intelligence, grade: 1.8

Baden-Wuerttemberg Cooperative State University (DHBW), B. Eng. **2016 - 2019**
Major: Mechatronics, Automotive System Engineering, grade: 1.5
- BEng Thesis: Investigations into the highly dynamic control of multiphase machines (grade: 1.1)

Heinrich Schickhardt School, Freudenstadt **2013 - 2016**
High-school diploma (“Abitur”), grade: 1.0

WORK EXPERIENCE

Mercedes-Benz AG, Electric Drive Unit Testing **2019 - 2025**
R&D Engineer
- EQXX Project - see PROJECT EXPERIENCE
- Responsible for Test specification, operation, troubleshooting and evaluation of complete electrical drivetrains on road-like test benches for functional as well as endurance testing

Levitum e.V., Munich **2023 - 2024**
Flight Control Engineer
- Building the worlds longest range hydrogen-powered eVTOL drone
- Responsible for flight-control software and flight-testing

DHBW Engineering e.V., Stuttgart **2018 - 2019**
Vehicle Dynamics Engineer
- Responsible for Inverter application as well as simulation and testing of the electric powertrain on the test bench and on track.
- Implemented a learning algorithm for power-estimation to approach operation near the regulated power limit (MATLAB & Simulink).

TEACHING EXPERIENCE

Teaching Assistant - University of Tübingen **Summer 2026**
ML-4509 AI Safety, with Jonas Geiping, Sahar Abdelnabi and Maksym Andriushchenko
- Gave tutorials, created homework and exams, and was responsible for grading

Teaching Assistant - University of Tübingen **Winter 2025/26**
ML-4331 The Science of Machine Learning Benchmarks, with Moritz Hardt
- Gave tutorials, created homework and exams, and was responsible for grading

Tutor - Baden-Wuerttemberg Cooperative State University (DHBW) **2020 - 2021**
- Supervision and evaluation of student research projects and seminar papers about a vehicle dynamics simulation of a complete Formula-Student prototype.

Trainer - Mercedes-Benz AG **2020 - 2025**
- Trainer and examination board committee member for training of employees to Qualified Electricians according to DIN VDE 1000-10

Supervisor - Mercedes-Benz AG **2022**
- Supervision and evaluation of a bachelor thesis “Investigation of methods for measuring the efficiency of electric drivetrains on system test benches” at Mercedes-Benz R&D.

RESEARCH EXPERIENCE

Situation Awareness of Users of AI Coding Agents - MPI-IS, Tübingen **2026**
Adapted the human-factors method SAGAT (Situation Awareness Global Assessment Technique) to measure how aware users are of what a coding agent did.

Rubric-Contrastive Advantage for GRPO - MPI-IS, Tübingen **2026**
Developed a rubric-contrastive advantage (RCA) for GRPO that compares a group's rollouts criterion by criterion instead of collapsing each rollout's rubric scores into a single scalar reward.

Rich Feedback for Fine-Tuning (RFFT) - MPI-IS, Tübingen **2025 - 2026**
Fine-tuned LLMs on natural-language feedback instead of scalar rewards or pairwise preferences, arriving at an approach closely related to concurrent work on self-distillation.

LLM Alignment - Stanford University **2024**
Human-Centered Artificial Intelligence, Stanford Trustworthy AI Research Group (STAIR, Prof. Koyejo)
- Modified, developed and evaluated code base for Large-Scale LLM alignment, including fine-tuning, active learning via Thompson-sampling methods and RL via PPO (Python)
- Contributed sections to and revised *Reinforcement Learning from Human Feedback* book draft of Prof. Sanmi Koyejo

Visual-Inertial SLAM for mobile robots - TUM **2023**
Implementation of an autonomous navigation software for a quadrocopter, including features such as localization, state estimation, motion planning and control (C++, ROS).

Using Reinforcement Learning to Control an eVTOL Drone - TUM **2022**
Expanded simulation and implemented reinforcement learning to train a neural network as a controller of a tail-sitter UAV (Julia & ReinforcementLearning.jl)

Control of multiphase machines - Mercedes Benz Inverter Research **2019**
Researched, simulated and compared results of different approaches to control dynamics of multiphase electric machines (Matlab & Simulink).

PROJECT EXPERIENCE

Vision EQXX - Mercedes-Benz AG, Stuttgart **2020 - 2021**
- Developed and managed the only powertrain test-bench for the Vision EQXX
- Accomplished application and optimization of efficiency-increasing and energy-saving measures and operating strategies of the prototype drivetrain of the Vision EQXX on the efficiency test bench.

The EQXX is the world's longest-range electric car to date. It achieved first 1008 km (Sindelfingen - Cassis, 5th April 2022) and then 1202 km (Stuttgart - Silverstone, 24th June 2022) range with a single charge at about 8.3 kWh/100km.

Building the world longest range eVTOL drone < 25 kg - LEVITUM e.V. **2023**
Developing a hydrogen powered eVTOL drone capable of flying over 300 km (C++, PX4, Gazebo).
- Investigate and adapt the PX4-flight control software utilizing flight tests and simulation.

Inverter application - DHBW Engineering e.V. **2018 - 2019**
Responsible for Inverter application as well as simulation and testing of the electric powertrain on the test bench and on track. Implemented a learning algorithm for power-estimation to approach operation near the regulated power limit (MATLAB & Simulink).

AWARDS & ACHIEVEMENTS

PROMOS Stipend **2024**
Programme to increase the Mobility of Students from German Universities by the German Academic Exchange Service ("Deutscher Akademischer Austauschdienst" – DAAD)

B.Eng. - ECTS Classification A **2019**
Graduated amongst the best 10% of the department

SOCIAL COMMITMENT

New Apostolic Church, Conductor - Youth choir **2019 on**
Organise and conduct rehearsals, devine services, weekends and concerts