

Simon Holk

Research Scientist | Machine Learning & General-Purpose Robot Policies

Tokyo, Japan | simon.holk00@gmail.com | [linkedin.com/in/simonholk](https://www.linkedin.com/in/simonholk)

RESEARCH PROFILE

Machine-learning and robotics researcher combining conceptual work in preference-based reinforcement learning with implementation-heavy development of large VLA policies for real bimanual manipulation. Published seven peer-reviewed papers across ICRA, IROS, HRI, and RA-L, then moved into full-stack model post-training, robot-data curation, deployment, and physical failure analysis. Interested in simple, general learning principles that survive contact with real hardware.

CORE EXPERTISE

Machine learning fundamentals • Robot learning • VLA and multimodal policies • Preference-based RL • System building • Deformable-object manipulation

RESEARCH & ENGINEERING EXPERIENCE

AIROA — Tokyo, Japan | Dec 2025–Present

Robotics / VLA Research Engineer

- Own the full development loop for vision-language-action (VLA) policies on a bimanual AgiBot G2/YUBI robot: data acquisition and quality, post-training, deployment, physical evaluation, and failure-driven iteration.
- Post-train large VLA models, including OpenPI/AiroPi pi0.5-family policies, on multi-task real-robot demonstrations using JAX/PyTorch workflows and multi-GPU HPC infrastructure.
- Analyze and curate heterogeneous robot datasets by aligning RGB, kinematic, gripper, task, and outcome signals; distinguish corruption, harmful supervision, valid recovery, and task-specific multimodality.
- Diagnose failures across model inference, camera transforms, coordinate frames, action timing, controller integration, payload calibration, and robot hardware, then carry fixes back into data, training, or deployment.
- Build reproducible training and evaluation workflows for large multimodal policies, with explicit dataset variants, checkpoint lineage, controlled comparisons, and robot-level acceptance criteria.

MATSUO-IWASAWA LABORATORY, THE UNIVERSITY OF TOKYO — Tokyo, Japan | May 2025–Present

Robotics Researcher

- Research generalizable robot learning for long-horizon bimanual manipulation of deformable objects, emphasizing data efficiency, corrective demonstrations, and deployment-grounded evaluation.
- Develop research directions spanning VLA post-training, preference and reward feedback, continual adaptation, predictive objectives, and policy-relative data selection.
- Design experiments that connect dataset or learning interventions to controlled physical-robot outcomes rather than relying on offline proxies alone.

KTH ROYAL INSTITUTE OF TECHNOLOGY — Stockholm, Sweden | Jan 2021–May 2025

PhD Candidate — Reinforcement Learning & Human-Robot Interaction

- Developed preference-based reinforcement learning methods that use language reasoning, temporal highlights, semi-supervised query synthesis, and active selection to make human feedback more informative.
- Published peer-reviewed papers at ICRA, IROS, and HRI; POLITE received three best-paper nominations at ICRA 2024.
- Built and evaluated learning pipelines at the intersection of reinforcement learning, human-robot interaction, language feedback, and sample-efficient preference acquisition.

KLARNA BANK AB — Stockholm, Sweden | Aug 2019–Dec 2020

Backend Engineer

- Developed production backend services, data pipelines, and machine-learning integrations in a large-scale financial technology environment.

LUND UNIVERSITY — Lund, Sweden | Jan 2018–Jun 2019

Teaching & Research Assistant

- Built a Unity VR visualization for ESA Gaia satellite data and taught VR development, databases, and software-engineering laboratory courses.

SELECTED RESEARCH & TECHNICAL WORK

PREFERENCE-BASED REINFORCEMENT LEARNING

- Developed PREDILECT and POLITE to represent richer human preferences through zero-shot language reasoning and temporally localized highlight feedback.
- Developed semi-supervised query synthesis and active query-selection methods through SEQUEL and VARIQuery, targeting more informative feedback acquisition.

VLA DATA QUALITY & CURATION

- Build task-, hand-, and phase-conditioned analyses of multimodal robot demonstrations instead of reducing data quality to a single episode-level score.
- Design non-destructive filtering, weighting, masking, and matched-control workflows that remove corrupt supervision while preserving valid recovery and rare task modes.

PHYSICAL DEPLOYMENT FIDELITY

- Trace policy failures across camera preprocessing, action semantics, coordinate frames, temporal execution rate, controller behavior, payload calibration, and hardware.
- Use synchronized robot trials, video, and telemetry to separate learning failures from deployment-stack failures and choose the next data, model, or systems intervention.

EDUCATION

MSc in Engineering — Information and Communication Technology | Apr 2020

Faculty of Engineering, Lund University — Lund, Sweden

SELECTED PUBLICATIONS & PREPRINTS

- **Holk, S.**, Takanami, R., Matsushima, T., Iwasawa, Y., Matsuo, Y., Wu, Y.-H., & Ota, K. (2026). Auditing Instruction–Trajectory Mismatches in Multimodal Robot Demonstrations. *IEEE Robotics and Automation Letters (RA-L)*.
- Odonchimed, S., Matsushima, T., **Holk, S.**, Iwasawa, Y., & Matsuo, Y. (2025). Retrieve-Augmented Generation for Speeding up Diffusion Policy without Additional Training. *arXiv preprint arXiv:2507.21452*.
- 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. *IEEE International Conference on Robotics and Automation (ICRA)*.
- **Holk, S.***, Marta, D.*, & Leite, I. (2024). PREDILECT: Preferences Delineated with Zero-Shot Language-based Reasoning in Reinforcement Learning. *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.
- **Holk, S.**, Marta, D., & Leite, I. (2024). POLITE: Preferences Combined with Highlights in Reinforcement Learning. *IEEE International Conference on Robotics and Automation (ICRA)*. Nominated for Best Conference Paper, Best Student Paper, and Best HRI Paper.
- Marta, D.*, **Holk, S.***, Pek, C., & Leite, I. (2024). SEQUEL: Semi-Supervised Preference-based RL with Query Synthesis via Latent Interpolation. *IEEE International Conference on Robotics and Automation (ICRA)*.

* *Equal contribution.*

TECHNICAL SKILLS

Research: Vision-language-action models; multimodal foundation models; robot learning; reinforcement learning; imitation learning; preference-based RL and RLHF; diffusion/flow policies; continual adaptation; human-robot interaction

Robotics: Real-robot deployment; bimanual manipulation; deformable objects; robot evaluation; data acquisition and curation; failure analysis; controller and sensor integration; ROS/ROS2

Engineering: Python; JAX; PyTorch; C/C++; SQL/NoSQL; Linux; Docker; Git; CI/CD; distributed and multi-GPU training; HPC/PBS; AWS; Azure

Languages: Swedish (native); English (professional proficiency)