

Oliver Dippel

Research Engineer · Reinforcement Learning & Foundation Models

St. Gallen, Switzerland · o.dippel91@gmail.com · +49 176 31407346
oliverdippel.github.io · github.com/oliverdippel · linkedin.com/in/oliver-dippel-phd-875707192 · Google Scholar

SUMMARY

PhD research engineer in reinforcement learning and foundation models, with peer-reviewed work at AAMAS 2026 and SGAI 2024/2023. Specialises in why training breaks: reward hacking, posterior and entropy collapse, distribution shift. Diagnoses these through systematic ablation and statistical root-cause analysis rather than hyperparameter search, builds the distributed training stack the work runs on (JAX, PyTorch, Ray, SLURM), and ships the result into production. Deployed systems include a decision-support tool in sustained operational use and agentic systems in safety-critical and B2B settings.

EXPERIENCE

Technical Advisor, Agentic Systems · CuratelQ Jun 2026 – Present

Remote, from Switzerland · independent, part-time

- Co-designed the intent-routing and RAG architecture behind LushLife.ai and Birch (birchstore.com), fusing catalog, product-safety and category-expertise sources so multi-skill hand-off resolves to the correct specialist agent.
- Specified the retrieval pipeline over live product data, keeping every generated answer strictly grounded in the merchant catalog.
- Leading the extension of a single-merchant agent stack into multi-tenant infrastructure other retailers can deploy on.

Research Engineer, RL & Foundation Models (PhD) · University of Liverpool Dec 2021 – Jun 2026

CDT in Distributed Algorithms · industry-funded

- Designed the Heuristic Transformer, a belief-augmented in-context RL architecture that improves zero-shot policy generalization over standard RL baselines (AAMAS 2026).
- Diagnosed and resolved recurring training-time pathologies — reward hacking, posterior collapse, entropy collapse — through systematic ablation studies and statistical root-cause analysis, not hyperparameter patching.
- Deployed a model fine-tuned on structured expert feedback as a real-time recommender; it remains in sustained operational use as a decision-support tool among domain specialists.
- Cut experiment turnaround across HPC job cycles by architecting fault-tolerant, reproducible distributed training and MLOps pipelines in JAX and PyTorch on Ray and SLURM.

Research Engineer, LLM Agents & Agentic Systems · ZeroAI Sep 2023 – Dec 2024

Liverpool, UK · part-time, concurrent with PhD

- Built a drift-detection and regression-testing harness that surfaced distinct agent failure modes under distribution shift before deployment, via a systematic adversarial evaluation suite.
- Integrated diffusion models into a production perception pipeline, improving multi-step trajectory forecast stability in high-dimensional sensor state-spaces.
- Engineered tool-use and structured-output validation on AWS for a safety-critical autonomous-driving agent system, with CI/CD-integrated monitoring and observability.

Statistical Consulting & ML Teaching · statistiknachhilfe.ch Jan 2021 – Present

Freelance

- Designed and taught a curriculum in Bayesian and probabilistic ML for technical and non-technical audiences.

Data Science Intern, Applied ML · Novartis Pharma AG Jun – Dec 2020

Basel, Switzerland

- Validated ML pipelines against GxP data-quality standards in a regulated healthcare environment.

SELECTED PUBLICATIONS

Heuristic Transformer: Belief Augmented In-Context Reinforcement Learning. **AAMAS 2026**. With A. Lisitsa, B. Peng.

Contextual Transformers for Goal-Oriented Reinforcement Learning. **SGAI 2024**. With A. Lisitsa, B. Peng.

Deep Reinforcement Learning for Continuous Control of Material Thickness. **SGAI 2023**. With A. Lisitsa, B. Peng.

SKILLS

Research: Reinforcement learning (policy gradients, Bayesian RL, exploration, credit assignment), RLHF and reward design, LLM fine-tuning (SFT, LoRA/PEFT), diffusion models, distribution-shift and drift diagnosis, applied statistics

Agentic systems: Tool-use agents, multi-step reasoning, RAG, structured output, multi-agent orchestration, evaluation and regression harnesses

Engineering: Python (advanced), JAX (Haiku/Flax), PyTorch, Ray, SLURM/HPC, AWS, Kubernetes, Terraform, Docker, Weights & Biases, Git, CI/CD

Languages: German (native) · English (fluent) · French (intermediate)

EDUCATION

Ph.D. Electrical & Control Engineering — University of Liverpool, CDT in Distributed Algorithms 2021 – 2026

M.Sc. Applied Statistics — University of Göttingen 2017 – 2020

B.A. Economics — University of Göttingen 2014 – 2017