STEFFEN BACKMANN

Zürich, CH steffenbackmann.com me@steffenbackmann.com /steffen-backmann /sbackmann

Education

ETH Zürich Zürich, Switzerland

Computer Science, M.Sc.: Machine Intelligence (Major), Software Systems (Minor) – GPA: 5.6/6 *Sep 2022 – Aug 2025*

- **Courses**: Probabilistic AI, Machine Perception, Design of Parallel and High-Performance Computing, Interactive Machine Learning: Visualization & Explainability, Reliable and Trustworthy AI
- **Thesis**: Evaluating LLM Agents Alignment in Ethically Charged Social Dilemmas

University of Münster Münster, Germany

Information Systems, B.Sc.– GPA: 1.0 (Germany)/4.0 (US); Rank 1 out of 234 *Oct 2019 – Aug 2022*

- **Courses**: Deep Learning, Operations Research and Optimization, Data Analysis & Simulation
- **Thesis**: Evaluation and implementation of Deep Learning based Instance Segmentation of trees in satellite data
- **Activities**: Exchange term at UFV in Abbotsford, Canada – Dean's List standing; teaching assistant for Linear Algebra & Analysis; teaching assistant for Operations Research; member of Accounting Talents excellence program; recipient of the Deutschlandstipendium scholarship

Experience

Oracle Labs Zürich, Switzerland

Senior Member of Technical Staff – Python, LLM, Langchain, OCI, Oracle DB *Oct 2026 - Current*

Member of Technical Staff *Sep 2025 – Oct 2026*

Research Assistant *Jul 2024 – Dec 2024*

- Research team on integrating AI functionalities into Oracle DB
- Designing and implementing multi-agent LLM application for automating financial investigations – patent submitted
- Contributing to LLM agent execution and specification framework

Porsche AG, Financial Strategy and Data Science Stuttgart, Germany

Data Science Intern – Python, JavaScript, Dataiku *Jan 2022 – Apr 2022*

- Built and deployed a model for automated transaction confirmation PDF processing
- Successfully prototyped an entity resolution solution of bolted joints in different production manuals

Silke & Christian Dohrmann GbR Steyerberg, Germany

Completed Agricultural Vocational Education *Aug 2017 – Mar 2019*

- Acted independently in all forms of agricultural work including plant production, machine operating and livestock
- Involved in project planning and implementation of renewable energies (wind, solar, biogas)

Projects

- **Evaluating LLM Agents Alignment in Ethically Charged Social Dilemmas**: Designed and implemented a simulation framework to evaluate LLM behavior in game-theoretic social dilemmas. Benchmarked alignment of frontier models' decision-making under conflicting ethical and reward structures. Supervised by Prof. Bernhard Schölkopf and Prof. Zhijing Jin.
- **Instance Segmentation of Trees in Satellite Data**: Designed and implemented a pipeline for instance segmentation of Rwandan satellite imagery, leveraging Mask R-CNN and Cascade Mask R-CNN models using PyTorch and the MMDetection framework.
- **Benchmarking CUDA in Julia vs. CUDA in C**: Implemented and evaluated highly optimized kernels in CUDA C and CUDA Julia for a set of benchmarks in collaboration with the ETH SPC Lab.
- **Soft Masking Techniques for LLM Explainability**: Proposed and implemented a soft masking strategy using textual replacement explanations for a more accurate comparison of LLM explainability methods. This included generating text embeddings with PyTorch and Google's Gemma model and vector similarity search with faiss.
- **Interactive Visualizations of GNN Explainability Methods**: Developed an interactive visualization for comparing different Graph Neural Network explainability methods on a molecule toxicology dataset using ReactJS and Spark. This enables domain experts to interact with the graph representation of the molecule, reassign edge explanation weights and observe on-the-fly recomputed metrics.