

MORITZ SCHALLER

London, UK +43 6608744235 moritz.schaller@hotmail.com moritzschaller.com [LinkedIn](#) [GitHub](#)

PROFILE

ML researcher building the **world's first foundation model for neurosurgical video**: **multi-GPU** distributed pretraining on **700+ hours** of surgical footage, tested on **segmentation, tracking, classification**, and **low-latency** inference. Prior work spans **computer vision** (mAP **60%→85%**, defect detection across 1000+ km of rail tracks) and **LLM systems (on-device personalisation with RLHF/GPO, multi-agent RAG)**. Hands-on with **PyTorch distributed training** and research-to-deployment pipelines. **Agentverse hackathon winner** (700+ participants). Seeking ML/AI/Research roles from **August 2026**.

EDUCATION

UNIVERSITY COLLEGE LONDON

Artificial Intelligence and Imaging MSc

London
2025-2026

- On track for passing with **Distinction**
- Relevant modules: Advanced Topics in Machine Learning; Applied Deep Learning; Inverse Problems in Imaging; Machine Learning, Open-Endedness and General Intelligence; AI for Surgery & Intervention; Information Processing, Open-Endedness and General Intelligence

UAS VIENNA

BSc Computer Science (Specialisation in Data Science & AI)

Vienna
2022-2025

- **1.09** (Austrian 1-5 scale; 1.0 best) **First-Class**
- Relevant modules: Data Science & Machine Learning; Computer Vision & Natural Language Processing; Probabilistic Programming; Big Data Analytics; DevOps & Cloud Computing; Software Engineering I-II (Lab); Scientific Writing & Research Methods

SELECTED EMPLOYMENT

ATOS

Machine Learning Engineer

Austria
October 2024 - September 2025

- Automated new hire support with a **multi-agent RAG** system built on **LangGraph**, instantly answering known questions and routing unknown queries to experts, reducing average onboarding time by **20%**
- Deployed a **LangChain**-orchestrated RAG pipeline with **knowledge graphs** and **VLMS** to multiple customers on **Azure**, engineering automated file-sync for persistent knowledge bases and achieving **90%** average QA accuracy
- Created a system automatically measuring wire length with a precision of ± 0.01 **mm** using the **Canny edge detector**
- Collaborated with manufacturing to deploy a real-time defect detector achieving **>95%** accuracy at line speed, using **UI attention maps** to explain predictions and establishing monitoring systems for operator workflows
- Leveraged NLP to find differences across **30-page** insurance documents using **LLM-based RAG** (chunking + vector search), **Named Entity Recognition**, and semantic similarity scoring to flag clause-level changes

SMART DIGITAL

Data Science Engineer

Austria
January 2023 - September 2024

- Raised detection mAP from **60%** → **85%** by switching from Detectron2 to **YOLOv8**, cleaning labels, and adding **Copy-Paste/MixUp** with targeted LR and confidence-threshold tuning
- Mapped georeferenced detections to latitude/longitude with **1 cm** precision
- Leveraged **LLMs** for object classification to cut down labelling time by **50%**
- Delivered an end-to-end railway defect detection system, analysing over **1000km** of track imagery to automate manual inspections and save clients **thousands of man-hours** annually

AWARDS

UNIVERSITY EXCELLENT SCHOLARSHIP

Awarded 3 consecutive years, top 4% of students

Austria
2022-2025

AGENTVERSE HACKATHON

Winning team, 700+ participants, £45k prize

London
2025

LEADERSHIP AND VOLUNTEERING

CARDBID (cardbid.eu)

Co-Founder

Austria
March 2025 - Present

- Founded auction-first Pokémon marketplace (www.cardbid.eu) with **computer vision grading** (centering, wear, defects), **ML pricing models**, and **LLM chatbot** serving **200k+** product templates
- Building automated condition assessment pipeline generating standardised grades to reduce disputes; live auctions

UCL AI SOCIETY

Machine Learning Officer

London

October 2025 - Present

- Leading weekly interactive workshops on machine learning topics for over **100** technical and non-technical members

AUSTRIAN RED CROSS

Paramedic

Austria

June 2021 - October 2024

- Responded to **144** emergency calls, stabilising critically ill patients using advanced equipment
- Coordinated with teams on **300+** callouts, providing continuous treatment during transit and delivering handovers to doctors

SELECTED PROJECTS

MASTER'S THESIS [🔗](#)

Development of a Foundational Model for Neurosurgical Videos

London

October 2025 - Present

- Developing **first foundational model** for neurosurgical applications via **multi-GPU** distributed training on a unique dataset of over **700 hours** of mostly unlabelled surgical video
- Evaluating model's ability to understand instruments through **detection, segmentation, pose estimation, and tracking**
- Assessing model's capacity to map operative environment via **depth estimation, tissue segmentation, tissue classification, and anatomical recognition**
- Targeting **low-latency** inference for intraoperative use, profiling and optimising model throughput under real-time constraints

BACHELOR THESIS [🔗](#)

Automatic LLM Personalisation with Human Feedback

Austria

October 2024 - June 2025

- Introduced a **user-embedding layer** parallel to word/positional embeddings in a single shared base model, enabling per-user personalisation with **minimal extra parameters**
- Trained user embeddings from scratch and fine-tuned base via **LoRA** with **RLHF** and **Group Policy Optimisation (GPO)**, achieving responses aligned with preferences **90%** more often after retraining
- Evaluated on **Natural Questions (NQ)** and **CoQA**; feedback-driven user embedding for an on-device LLM produced answers measurably

PERSONAL PROJECT [🔗](#)

Multi-Modal AI Trading Agent

Austria

September 2023 - February 2025

- Developed a **Double DQN** trading agent processing a unique **3-channel state** by combining candlestick charts, Google Trends data, and news sentiment analysis
- Engineered agent with a **CNN state encoder** and **prioritised experience replay**, accelerating training and improving trading performance by an average of **50%**
- Applied **Explainable AI (XAI)** techniques to interpret model behaviour, using attention maps to visualise candlestick patterns and to quantify the agent's reliance on each input channel
- Validated strategy through extensive **backtesting** and demonstrated its ability to maintain profitability during **live paper trading**

CERTIFICATIONS

COURSERA

Deep Learning

Online

March 2022 - May 2024

- Neural nets, backpropagation & optimisation; building **CNNs, RNNs/LSTMs**, and regularisation (dropout/BN) with practical training workflows (initialisation, LR schedules)
- Tokenization and vectorisation (**TF-IDF, word2vec/GloVe**), sequence models (**RNN/LSTM + attention**), sentiment, translation

HARVARD

CS50's Introduction to Artificial Intelligence

Online

September 2021 - January 2022

- Graph search, knowledge representation, optimisation, neural networks & NLP, and **reinforcement learning**

SKILLS

PROGRAMMING & FRAMEWORKS: Python, C, C++, C#, R, PyTorch, TensorFlow, Numpy, Pandas

SOFTWARE DEVELOPMENT: end-to-end system design, managing multiple components with Docker containers

AI DEVELOPMENT: Data processing pipelines, MLOps (model profiling, optimisation), neural networks, HPC, LangChain/LangGraph

SOFTWARE ENGINEERING PATTERNS: MVVM, modular design, scalable architecture for complex software projects

TOOLS: Git, Linux Shell Environment, Docker, Azure/AWS/GCP, MLflow/Weights & Biases, CI/CD, Databricks, Jira

LANGUAGES: English: C2 Proficient (TOEFL 110/120); German: Native