

Johann Leidenfrost

johann@leidenfrost.pro | +49 152 07772824 | [Personal Website](#) | [GitHub](#)

EDUCATION

Imperial College London: MSc in Computing (Artificial Intelligence and Machine Learning) 2024 – 2025
Distinction (74.87%)

MSc Thesis: Vision-based interpretable trajectory prediction

- Developed an explainable AI model to predict the future paths of moving objects using video data. Accepted at the Workshop on Safe Physical AI, IJCAI 2026. See submission [here](#).

Modules: Deep Learning, Information Theory, Mathematics for ML, Software Engineering for ML Systems, Computational Optimization, Reinforcement Learning, Deep Graph-Based Learning, Quantum Computing, Robotics.

University College London: Bachelor of Arts and Sciences with Major in Sciences and Engineering 2020 – 2024
First Class Honours (74.94%); Placed on the Dean's List awarded to the top 5% of the graduating cohort.

BSc Thesis (in collaboration with Farming Revolution GmbH): Generation of a realistic plant image dataset

- Built and trained a CycleGAN pipeline to generate synthetic datasets of crops and weeds, evaluating its viability for augmenting training data in commercial plant detection and classification models.

Modules: Probability Theory, Numerical Methods, Analysis, Algorithms and Data Structures, Logic, Neuroscience.

Year Abroad: University of Toronto (GPA: 4.0). Modules: Multivariable Calculus, Linear Algebra, Symbolic AI.

EXPERIENCE

Research Engineering Intern (BioAI): InstaDeep Ltd October 2025 - April 2026
Project: Advancing interpretability of RNA foundation models *LoRA, Git, Docker, CI/CD, PyTorch, NumPy*

Research Objective: Identifying RNA motifs that drive model predictions.

- For the first time, introduced attribution-based interpretability into BioNTech's RNA foundation models.
- Established a standardised pipeline for integrating interpretability into InstaDeep's task-specific RNA model development using integrated/expected Gradients and attention maps.
- Confirmed the importance of hypothesised sequence motifs in state-of-the-art biological models.
- Presented my results to senior stakeholders at BioNTech.
- A detailed blog article about my work adapted to public data and models is available [here](#).

Software Engineering Project: MSc Coursework at Imperial College London January 2025 - March 2025
Project: Acute kidney injury (AKI) detection service *Distributed systems, Kubernetes, Docker, CI/CD, Git*

- Built a Kubernetes-orchestrated microservice to continuously process real-time hospital blood test data.
- Containerised the application using Docker and guaranteed high availability and zero data loss during pod failures. Integrated a predictive AKI detection model.

Working Student: Farming Revolution GmbH October 2023 - January 2025
Project I: Annotation and evaluation feature for plant detection model *React, TypeScript, SQL*

Problem: For densely growing plants, plant annotation required labelling plant stems one by one. This is time-intensive.

- Developed a new plant annotation feature allowing the annotation of many stems at once.

Project II: Adaptive evaluation system *Model evaluation, TensorFlow, NumPy, HDF5, gzip*

Problem: The model's raw predictions on the training and validation data were not saved. It was not possible to evaluate the model using different evaluation parameters without re-running inference on the model.

- Built a pipeline saving the raw model predictions quickly and using minimal memory due to lossless data compression. This reduced re-evaluation time of models from 1 h to 5 min.

Research Internship: University of Tokyo Graduate School of Frontier Sciences June 2023 - August 2023
Project: Synthesis of earthquake ground motion time histories using a variational autoencoder *TensorFlow, NumPy*

- Evaluated the validity of the generated time histories when sampled from the feature space.
- Optimised generated time histories for increasing the nonlinear response of the target structure.

TECHNICAL SKILLS AND LANGUAGES

Programming Languages/Frameworks/Libraries: Python, PyTorch, TensorFlow, NumPy, Pandas, Git, Docker, Kubernetes, TypeScript, Java, SQL, OpenCV, React, Captum, Scikit-learn, Slurm, HuggingFace

Languages: English, German, French, Spanish, Bulgarian