

Lukas J. Hondrich

Berlin, Germany

[website](#) 🏠
lukashondrich@gmail.com

Profile

AI Engineer with 7 years building ML/AI systems, specialized in LLM applications, RAG optimization, and context engineering. Experienced in fine-tuning models (LoRA/PEFT) and deploying scalable production systems serving 25k+ users with comprehensive observability (LiteLLM, LangFuse). Track record across education technology, e-commerce, and healthcare, with deep expertise in AI safety and data protection.

Professional Experience

Independent AI Engineer

09, 2025 – present
Berlin, Germany

- Building **Open Working Hours**: privacy-preserving data donation platform using differential privacy and k-anonymity to aggregate unreported overtime in German healthcare [React Native, FastAPI, PostgreSQL, Docker](#) [Link]
- Building **real-time voice-to-voice** conversational language tutor with agentic memory, RAG, and interactive vocabulary visualization [Python, OpenAI, Elasticsearch, Haystack](#) [Try!]

AI Engineer - Quality & Safety

02, 2024 – 08, 2025
Berlin, Germany

Cornelsen - educational technology

- **Led development** of an LLM-based (Llama3.3) tutoring app in a highly regulated educational setting for 25k students, on market within 10 weeks [Python, FastAPI](#)
- **Optimized system performance** through function calling and rate limiting, reducing inference costs by 50% and increasing concurrent capacity to 520 users (2.4x) [Python, FastAPI](#)
- Designed and implemented a **Quality and Safety platform** for running automated evaluations, data-driven analyses, resulting in significant improvements in error rates [Python, Streamlit](#)
- **Student space modelling**, using psychological, demographic, didactic dimensions to generate coverage optimized synthetic data on demand [Python](#)
- Conducted **statistical modelling** using linear mixed-effects models to analyze temporal trends in conversation flows helping domain experts improve AI quality [Python, SciPy](#)
- **Optimized AI/RAG** for three customer-facing educational applications leading to sign. improvements in retrieval quality and customer satisfaction [Python, FastAPI, ChromaDB](#)
- Engineered **PII-removal service**, including finetuning of Llama3.1, [FastAPI, LoRA, Presidio, Docker, K8s](#) PEFT/LoRA, integrated with Presidio

Machine Learning Engineer

03, 2022 – 01, 2024
Berlin, Germany

Institute of Medical and Human Genetics, Charité

- Developed cVAE, U-net for **3D computer vision** and expert refinement [Python, TensorFlow, OpenCV](#)
- Built Infrastructure for **deploying, logging, analyzing** experiments [Bash, Slurm, Neptune.AI](#)
- **Created Researcher productivity tool**: iOS app for processing of 3D images [Python, PyQt, PySide](#)

AI Consultant

03, 2021 – 03, 2022
Berlin, Germany

AlgorithmWatch

- **Consulted** on measuring and improving fairness in embedding-based search systems
- **Advised Bosch GmbH** on implementing AI fairness measures in production systems

Machine Learning Engineer <i>Thomann - e-commerce</i>	10, 2018 – 09, 2020 <i>Berlin, Germany</i>
<ul style="list-style-type: none"> Built collective filtering based recommendation engine <i>GCP, Tensorflow, PySpark, Python, SciPy</i> Built data analysis tool for search engine usage for >1 million daily customers, significantly decreasing drop-off rate <i>GCP, Docker, Python, SciPy</i> Engineered forecasting system for global parcel delivery <i>GCP, Docker, Python, TF</i> 	

Selected Technical Skills

LLM & Alignment:

- Prompting, PEFT/LoRA finetuning, RLHF, LangChain, LLamaIndex, LiteLLM, LangFuse

Machine Learning & Data

- Python, PyTorch, TensorFlow, scikit-learn, NumPy, SciPy, mongoDB, chromaDB, Elasticsearch, Haystack*

APIs & Infrastructure

- OpenAI API, Huggingface Transformers, Accelerate, FastAPI, Docker, Kubernetes, Terraform, GCP, Hetzner*

Tools & Collaboration

- Git, GitHub, GitLab, CI/CD pipelines, Playwright*

Education

M.Sc. Cognitive-Affective Neuroscience <i>Technical University Dresden</i>	2015 – 2018 <i>grade: 1.6</i>
B.Sc. Psychology <i>Johannes Gutenberg-Universität Mainz</i>	2011 – 2015 <i>grade: 1.7</i>

Other Research Experience

Internship, M.Sc. Thesis <i>Bernstein Center for Computational Neuroscience & Princeton University</i>	04, 2017 – 04, 2018 <i>Berlin, Germany</i>
<ul style="list-style-type: none"> Thesis (grade 1.1): attention & learning in multidimensional environments <i>1.1</i> Modelled behavioural data with reinforcement-learning & Bayesian learning <i>Matlab, Slurm</i> Performed Bayesian model comparison on high-dimensional fMRI data 	

Workshops, Summer-Schools, Courses

AI-Officer Certification <i>aiofficer.de, according to Art. 4, EU-AI-Act</i>	12, 2024 <i>Berlin, Germany</i>
<ul style="list-style-type: none"> Training: Building AI-systems compliant to the EU-AI-Act 	
The Race to Regulate AI: Global Comparative Perspectives <i>Oxford University</i>	06, 2022 <i>Oxford, UK</i>
<ul style="list-style-type: none"> Workshop: EU-AI-act: High-risk systems and protection of fundamental rights 	
Data Engineering Training <i>Pipeline Data Engineering Academy</i>	10, 2021 – 12, 2021 <i>Berlin, Germany</i>
<ul style="list-style-type: none"> Data acquisition, data modelling, orchestration, CI/CD <i>Bash, SQL, GCP, Github Actions</i> 	

Relevant *Publications*

Implementing employee interest along the machine learning pipeline

Hondrich, Mollen, Ponce del Castillo (ed.), ETUI, *Artificial intelligence, labour and society* (2024)

Automation Bias in public administration – an interdisciplinary perspective from law and psychology

Ruscheimer, **Hondrich** (2024). *Government Information Quarterly*.

Three-Dimensional Histological Characterization of the Placental Vasculature Using Light Sheet Microscopy

Freise, **Hondrich**, Hägerling et al. (2023). *Biomolecules*

Addressing Automation Bias through Verifiability

Hondrich, Ruschemeier (2023). *European Conference on Alg. Fairness*.

Working paper: From risk mitigation to employee action along the machine learning pipeline

Mollen, **Hondrich** (2023). *Böckler Impuls*.

Public Speaking

Evaluating Multi-turn Conversational AI

Merantix AI Campus

EvalOps unfiltered: Evaluating LLM-based applications

2025

Shaping a new social contract for AI

European Trade Union Institute

Panel: Book launch of "Artificial intelligence, labour and society"

2024

EWAF 2023 - Lightning Round

Europ. Workshop on Alg. Fairness

Verifiability as a Minimal Requirement for Human Oversight

2023

AI Talks at ETUI

European Trade Union Institute

The Machine Learning Pipeline: Space for Action

2022

Successful Applications I Declined

Oxford University - DPhil in Computer Science

2023

- proposal: *Participatory Governance for Reinforcement Learning Systems in Work-Management Systems*

Languages

German - native

English - C1

Spanish - B1