

Tobias von Arx

tobiasvonarx@proton.me | +41 79 903 67 74

avanis.xyz | github.com/tobiasvonarx | linkedin.com/in/tobias-vonarx

Work Experience

Callosum

Incoming, Oct 2026

Member of Technical Staff, Capabilities

London, UK

- Working on new ways to map heterogeneous AI workloads to silicon
- Building inference capabilities across heterogeneous GPUs and multi-node systems

UBS

Jan 2024 – Jul 2024

Cyber Security Risk Intern, Third Party Risk

Zurich, CH

- Built LLM-based pipeline to assist human review and reduce effort in grading cyber security questionnaires
- Developed and supported internal cyber security risk assessment and incident management processes

ETH Zurich

Feb 2023 – Jun 2023

Teaching Assistant

Feb 2022 – Mar 2022

Zurich, CH

- Data Modelling and Databases, Prof. [Ce Zhang](#), and [ABZ Institute](#) at ETH
- Prepared course material and conducted exercise sessions for the entire student cohort (200+ students)
- Taught multiple classes of middle school students about cryptography and self-verifying codes

Education

MSc Data Science, [ETH Zurich](#), GPA: 5.6/6

Sep 2024 – Sep 2026

Selected coursework: Reliable & Trustworthy AI, LLMs, Probabilistic AI, Network Security, Optimization for Data Science, Computational Statistics, Big Data, Algorithms Lab

Visiting Student, [University of Cambridge](#), Fitzwilliam College

Mar 2026 – Sep 2026

BSc Computer Science, [ETH Zurich](#), GPA: 5.2/6

Sep 2020 – Sep 2023

Selected coursework: Formal Methods and Functional Programming, Compiler Design, Quantum Physics, Approximation Algorithms, Machine Learning, Theoretical Computer Science, Discrete Mathematics

Research Experience

AI Societal Impact Lab

Incoming, Oct 2026

Fellow

Remote, UK

- Reading and writing about how AI could be safely democratized and diffused across society and the economy, and how wider access shapes security, risk, and trust in open- versus closed-weight models

University of Cambridge, EEG Lab*Master Thesis, advised by Prof. [Srinivasan Keshav](#)*

Mar 2026 – Sep 2026

Cambridge, UK

- Augmenting 3D Roof Reconstruction for Photovoltaic Placement and Simulation
- Optimization and Vision Transformers for rooftop solar potential estimation from visual/geospatial data
- Produced two preprints (under review) and a web tool for 3D roof reconstruction and shading-aware PV relocation, design, and assessment

ETH Zurich, SRI Lab*Research in Data Science, advised by Prof. [Martin Vechev](#)*

Mar 2025 – Dec 2025

Zurich, CH

- Developed [AutoBaxBuilder](#), an LLM-agent system for generating web-backend security benchmarks
- Implemented the generation pipeline, execution-based refinement, sandboxing, and guardrails

ETH Zurich, with SwissAI / ETH AI Center*Data Science Lab, advised by Dr. [Imanol Schlag](#)*

Sep 2025 – Dec 2025

Zurich, CH

- Researched [memorization dynamics](#) in 100B+ token pretraining experiments on GH200s with Megatron
- Analyzed left-to-right vs. Fill-in-the-Middle (FIM) memorization via extraction and attention patterns

ETH Zurich, Boeva Lab*Bachelor Thesis, advised by Prof. [Valentina Boeva](#)*

Feb 2023 – Jun 2023

Zurich, CH

- Neural Networks for Signal Deconvolution in Cancer Transcriptomics Data, [report](#)
- Developed autoencoders and variational autoencoders for cell-type deconvolution in bulk RNA-seq data
- Demonstrated deconvolution on simulated tumors and identified limitations in resolving malignant states in clinical glioblastoma tumor transcriptome

Publications and Preprints

[Shading-Aware Rooftop PV Placement](#)*under review***Tobias von Arx**, Srinivasan Keshav**[Emboss: Image- and LiDAR-based Roof Detail Augmentation for 3D City Models](#)***under review***Tobias von Arx**, Srinivasan Keshav**[AutoBaxBuilder: Bootstrapping Code Security Benchmarking](#)**

ICML 2026

Tobias von Arx, Niels Mündler, Mark Vero, Maximilian Baader, Martin Vechev**[Memorization Dynamics of Fill-in-the-Middle Pretraining](#)**

MemFM @ ICML 2026

Tobias von Arx, Tanguy Dieudonné (equal contribution)**References**

Prof. Dr. [Srinivasan Keshav](#) FRSProf. Dr. [Martin Vechev](#)**Skills**

Languages English (fluent), German (native), French (intermediate), Chinese (intermediate)**Programming** Python, C++, Bash, JavaScript, Haskell, Elixir, Go, SQL, Linux, Git**Machine Learning** PyTorch, torchrun, Megatron-LM, Accelerate, vLLM, distributed pretraining, LLM evaluation, agent benchmarking, inference, Slurm, Docker