Vaibhav **KULKARNI**

Engineering Lead: Data Engineering | Data Science | Bioinformatics | AI-Driven Therapeutic Design

🇨🇭 *Switzerland (C Permit)*

✉ vkulkarn@protonmail.com | 🏠 vaibhav90.github.io | in vaibhav-kulkarni-b94891166 | 🎓 Vaibhav KULKARNI

Work Experience

Debiopharm
Lausanne, Switzerland

Engineering Manager: Data & Bioinformatics
July 2025 - Present

- Leading cross-functional team of data engineers & bioinformaticians to architect & deliver next-generation data fabric & bioinformatics platforms
- Engineering bioinformatics pipelines for processing multi-omics & non-clinical data for biomarker discovery programs & harmonizing datasets
- Partnering with Translational Medicine department to provide data analysis & bioinformatics support for projects in small molecule discovery & antibody-drug conjugate (ADC) research, leveraging foundational platforms like the Metadata Repository to harmonize vast pre-clinical datasets

Data Engineering Team Lead
July 2022 - June 2025

- Architected & led the development from scratch of the company's most utilized data platform; Central Data Repository, automating ingestion & quality control of clinical & non-clinical trial data at enterprise-scale & positioning it as central data hub for all research & development programs
- Engineered the deployment of custom-configured scientific analytics platforms (SAS Viya, Posit, Genedata, Dotmatics) on company-controlled cloud infrastructure, creating a unified ecosystem interfaced with the central data repository to ensure stringent performance & data governance
- Engineered an entity-recognition & search platform by mining scientific publication databases -Automated internal trend-monitoring functions
- Productionised privacy-aware large language model-based applications to streamline document analysis and study protocol writing processes
- Supported compliance team via DevOps-based computerised system validation, infra qualification & security audits -Automated GxP workflows
- Responsible for technical due-diligence of healthcare-startups (technical-stack, software & ML strategy) to aid our VC funds investment decisions

SOPHiA Genetics
Lausanne, Switzerland

Senior Data Engineer
April 2021 - June 2022

- Implemented terabyte-scale ETL pipelines to transform raw genomics data in-to analytics-ready format -Significantly lowering processing time
- Developed API endpoints & CLI tools to expose database functionalities to internal teams- Facilitated data democratisation and access control
- Implemented regression testing mechanisms for variant detection bioinformatics pipelines & dashboards delivering real-time insights into KPIs
- Authored formal TLA+ specifications for ETL operations, enabling quantification and benchmarking of logical correctness across the data stack

GenLots
Lausanne, Switzerland

Software Engineer | Research & Development Lead
September 2019 - March 2021

- Developed a microservices architecture, streamlining ingestion, transformation, & delivery of client ERP data to production-grade ML services
- Deployed reinforcement learning-based solution, optimising end-to-end supply chain planning, warehouse management, and CO_2 emissions
- Engineered a suite of tools for reporting & monitoring integrated into dashboards delivering real-time business performance metrics to clients
- Directed research projects in collaboration with Swiss watch manufacturing firms to predict inbound material needs & stabilise supply chains

Distributed Object Programming Lab
Lausanne, Switzerland

Research Scientist
November 2015 - August 2019

- Partnered with public transport firms to architect & implement systems capturing high-frequency, real-time ridership data into a data warehouse
- Developed a Machine Learning-driven platform to proactively detect and reduce ticketless travel across the Lausanne public transport network
- Led platform development for a large-scale (>300 participants) spatiotemporal mobility-data collection project to facilitate ML-privacy research

ETH Zürich
Zürich, Switzerland

Project Engineer
January 2015 - August 2015

- Implemented a software-defined-radio-based platform for reliable IoT product testing, generating interference patterns of wireless appliances
- Deployed the platform in production at ETH Zurich & TU Berlin with toolsets allowing for remote configuration & logging performance metrics

BOLT IoT
India

Embedded Systems Engineer
September 2011 - January 2013

- Designed PCB layouts for embedded platforms used in robotics & IoT systems & collaborated with software teams to ensure firmware integration
- Implemented toolsets & applications to benchmark platform performance, validation tests, EMI simulations & industry compliance certifications

National Institute of Oceanography
India

Project Intern
June 2011 - August 2011

- Contributed to autonomous underwater vehicle localization project by developing triangulation algorithms & learning satellite communication & ocean current modeling

Education

University of Lausanne - Information Systems department
Lausanne, Switzerland

PhD, Computer Science (Thesis: information systems for privacy-aware machine learning)
November 2015 - August 2019

Technische Universiteit Eindhoven (Master Thesis at ETH Zürich)
Eindhoven, Netherlands

MS, Embedded Systems (graduated with cum laude)
August 2014 - August 2015

Technische Universität Berlin
Berlin, Germany

MS, Information & Communication Technology (graduated with honours)
February 2013 - August 2014

Core Competencies

- **Software Engineering:** Object-oriented programming, test-driven development, microservices architecture, API development, and software design patterns. Experience with Python, Java, Scala, and Rust
- **Data Engineering & System Architecture:** Enterprise data platform design, ETL pipeline development, data modeling, and quality assurance. Built enterprise-scale data repositories across multiple domains including healthcare, ERP systems, public transportation, and research labs, handling real-time and batch processing at varying scales
- **Cloud & Infrastructure:** Containerization, orchestration, and cloud-native application deployment. Experience with Azure, AWS, Docker, Kubernetes, and high-performance computing environments
- **Data Science & Machine Learning:** Statistical analysis, predictive modeling, and production ML systems. Developed and deployed reinforcement learning solutions, real-time analytics platforms, & large language model applications
- **Bioinformatics:** Genomics analysis, multi-omics data processing, and clinical data harmonization. Experience with RNA-seq & variant calling pipelines, scientific computing platforms, and bioinformatics workflow management
- **Team Leadership & Management:** Cross-functional team leadership, technical strategy development, and project management. Led data engineering and bioinformatics teams, managed agile development processes, and conducted technical due diligence for healthcare startups

Publications & Research Contributions

- ***Generative models for simulating mobility trajectories*** (collaboration with Statistics Department at Columbia University)
  V. Kulkarni, N. Tagasovska, T. Vatter, B. Garbinato (published at: *Neural Information Processing Systems*, Montreal, 2018)
  Impact: Advanced state-of-the-art results for synthetic data generation using ML (Ranked 1 Conference in AI)

- ***Breadcrumbs: A mobility dataset with point-of-interest annotations*** (collaboration with Information Security-Privacy Lab)
  A. Moro, V. Kulkarni, PA Ghiringhelli, B. Chapuis, K. Huguenin, B. Garbinato (published at: *SigSpatial*, Chicago, 2019)
  Impact: Large-scale geospatial dataset to facilitate ML research (Ranked 1 Conference in Geo Information Systems)

- ***Controlled interference generation for wireless coexistence research*** (ETH Zurich - Distributed Systems Research Group)
  A. Hithnawi, V. Kulkarni, S. Li, H. Shafagh (published at: *Implementation-Mobile Computing and Networking*, Paris, 2015)
  Impact: open source project for a complete software-defined-radio platform for controllable interference generation

- ***Information Disclosure in Location-based services: An Extended Privacy Calculus Model*** (collaboration with BISA Lab)
  D. Naous, V. Kulkarni, C. Legner, B. Garbinato (published at: *International Conference on Information Systems*, 2019)
  Impact: Integrate user information disclosure behaviours in Apps (Ranked 1 Conference in Information Systems)

- ***Privacy-Preserving Location-Based Services by using Intel Secure Guard Extensions*** V. Kulkarni, B. Chapuis, B. Garbinato
  Open-source software project, (published at: *International Conference on Embedded Networked Systems*, Delft, 2017)
  Impact: Demonstrated the ability of hardware-based execution environments to deploy privacy-aware ML applications

Continuous Learning & Certifications

Data Engineering with Rust	*Duke University, 2024*
Generative AI with Large Language Models	*DeepLearning.AI & AWS, 2023*
Data Engineering with Azure	*Udemy, 2023*
Applied Machine Learning & Data Science	*University of Michigan, 2021*
Amazon Web Services	*AWS, 2020*
Information Security & Privacy	*University College London, 2019*
Fundamentals of Machine Learning	*Stanford University, 2017*

Talks, Grants & Awards

Generalising Machine Learning Models for Unseen Chemical Spaces in Drug Discovery	*BioTechX Europe, 2023*
Data Repository: Making Informed Technology Choices	*BioTechX Europe, 2023*
Expert Viewpoint on AI's Role in Data Space	*Swiss Association of Actuaries, 2023*
Generative AI Report for Healthcare	*AI for Health, 2023*
HEC-research grant for independent doctoral research	*Swiss National Fund, 2019*
Zeno Karl Schindler award for master thesis	*ETH Zürich, 2015*
EIT-ICT Mobility award best business plan in IoT domain	*EIT-ICT, 2013*

Extracurricular Activities

- Launched a technology outreach initiative aimed at fostering a passion for science and technology among rural Indian students. Personally visited 25 schools, engaging over 400 students with hands-on demonstrations and presentations on real-world scientific applications
- Competed in multiple endurance races - Lausanne Marathon ('18, '19), Kilomètre Vertical de Fully, & 20KM de Lausanne ('16-'25)
- Played for Flyhigh Lausanne in Ultimate Frisbee (2018, 2019) and actively participated in Padel Tennis tournaments in Switzerland