

Vaibhav Kulkarni

DATA ENGINEERING TEAM LEAD (INFRASTRUCTURE | FULL-STACK DEV | MACHINE LEARNING) · PHD - COMPUTER SCIENCE

Lausanne - Switzerland

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



Work Experience

Debiopharm

Lausanne, Switzerland

DATA ENGINEERING LEAD

July 2022 - Present

- Lead a technical team to design, architect & implement enterprise-data-warehouse to collect-harmonize-process clinical-genomics data
- Build an entity-extraction & search platform to scrape public & private datasets to construct knowledge-graphs & perform trend-analytics
- Design system architecture, implement & maintain large-scale statistical & analytics-applications for biostatistics & bioinformatics teams
- Support quality team to perform software validation, infrastructure qualification & computerised system validation under GxP regulation
- Conduct due-diligence of technical stacks, ML-Ops & AI best practices of technology startups to aid our digital health investment decision

SOPHiA Genetics

Lausanne, Switzerland

SENIOR DATA ENGINEER

April 2021 - June 2022

- Designed & implemented high performance & scalable extraction-transformation-load (ETL) pipelines for large-complex-geonmics data
- Developed & deployed secure API & CLI tools to expose large-scale-database functionalities to internal teams maintaining access control
- Architected & implemented analytical & debugging tools & dashboards to provide actionable insights into business performance metrics
- Delivered formal-specifications (TLA+) for data loading, transformation & analytical operations to quantify-benchmark logical correctness

GenLots

Lausanne, Switzerland

SOFTWARE ENGINEER | RESEARCH & DEVELOPMENT LEAD

Sept. 2019 - Mar. 2021

- Designed, implemented & scaled microservices-based architecture to ingest client ERP data, transform & serve to production ML services
- Implemented reinforcement-learning services for optimising end-to-end supply chain planning - warehousing, transactions, CO2 emissions
- Developed reporting, monitoring, logging tools integrated with dashboards to automate client-sharing on business performance indicators
- Led industrial research projects for forecasting inbound material requirements to ensure supply chain stability for large manufacturing firms

Distributed Object Programming Lab

Lausanne, Switzerland

RESEARCH SCIENTIST

Nov. 2015 - Aug. 2019

- Implemented an infrastructure to ingest high-frequency-realtime-streaming public-transport-ridership-data to a data-warehouse
- Implemented a machine-learning-based platform to monitor & minimise ticket-free riding on Lausanne public transport network
- Led a large-scale geospatial data collection campaign; Worked with the team to implement backend system & mobile application

ETH Zürich

Zürich, Switzerland

PROJECT INTERN

Jan. 2015 - Aug. 2015

- Implemented a software-defined-radio-based platform to generate controllable interference patterns for testing IoT products
- Deployed the platform on open IoT testbeds (Berlin, Zurich); Implemented toolsets & APIs to control & perform benchmarking

BOLT IoT

Goa, India

EMBEDDED SYSTEMS ENGINEER

Aug. 2011 - Aug. 2013

- Designed & optimized PCB layouts of embedded platforms for robotics & IoT applications; Set-up testing strategies
- Implemented tools & applications to benchmark platform performance; Led client support & troubleshooting teams

Education

University of Lausanne

Lausanne, Switzerland

PHD, COMPUTER SCIENCE

Nov. 2015 - Aug. 2019

Thesis: Information Systems for Privacy-Aware Machine Learning

Technische Universiteit Eindhoven

Eindhoven, Netherlands

MS, EMBEDDED SYSTEMS (GRADUATED WITH CUM LAUDE)

Aug. 2014 - Aug. 2015

Technische Universität Berlin

Berlin, Germany

MS, INFORMATION & COMMUNICATION TECHNOLOGY (GRADUATED WITH HONOURS)

Feb. 2013 - Aug. 2014

Skills & Competencies

- **OO-Programming, Scripting, Test-driven Development:** Java, Python, C/C++, Bash (6 years); Scala, Javascript (3 years); JUnit, pytest
- **Database Development:** SQL, PostgreSQL, MySQL (5 years); SQLite, MongoDB, Redis, Elasticsearch (2 years), pgAdmin (debugging tools)
- **Microservices-driven Architecture:** Docker-compose, Kubernetes, RabbitMQ, Kafka, Flask, Spring, Jersey, Airflow, AWS, Azure (4 years)
- **Observability, Logging, CI/CD:** Elasticsearch-Kibana-Logstash stack (2 years); Setting Gitlab/Jenkins/Bitbucket CI/CD pipelines (3 years)
- **ML & Data Science Stack:** NumPy, SciPy, pandas, Dask, scikit-learn, TensorFlow, Keras, R (4 years), Apache Spark/Spark streaming (2 years)
- **Version Control & Data Visualisation:** Git, SVN, JetBrains Toolbox (6 years); matplotlib, seaborn, Tableau, D3, deck.gl, Folium (4 years)
- **Security Audit Tools & Project Management:** npm-audit, kubeaudit, docker-bench-security, Bandit, Git-Secrets (2 years); Atlassian suite
- **Languages:** English (Fluent), French (working proficiency), German (elementary proficiency), Hindi (Fluent), Marathi & Konkani (Native)