

Vincent Le Goualher  
1426 Corcelles-Concise  
Born on 30.04.1990  
French citizen with B permit  
Category B driving licence

# Data Engineer

M. Sc. Data Science, M. Sc. Software Engineering

+41 77 471 07 76  
vlg.engineer@gmail.com  
blog.vlgdata.io  
github.com/datatrigger  
linkedin.com/in/vlgdata

## SKILLS

<b>Tools</b>	SQL, Python, Unix shell, Git, GitHub, Tableau, Docker, Kubernetes
<b>Data Engineering</b>	Data modeling, Data pipelines, ETL, Data Warehousing, Machine Learning, Data Viz
<b>Big Data</b>	Distributed computing frameworks Apache Spark, Apache Kafka
<b>DevOps - MLOps</b>	CI/CD pipelines, Container deployment, Microservices, Agile (SCRUM), Unit testing, APIs
<b>Cloud</b>	AWS e.g. IAM, EC2, Glue, S3, ECR, ECS, EventBridge   GCP e.g. GKE, BigQuery
<b>Soft skills</b>	Genuine team player, resolute problem solver, skilled vulgarizer

## EXPERIENCE

**Data Engineer** **08.2021 — 09.2021**  
*Layoff, Personal projects* *Concise, Switzerland*

- Developed and deployed a multi container ML application on Kubernetes at [translate.vlgdata.io](https://translate.vlgdata.io). See also [blog.vlgdata.io](https://blog.vlgdata.io)
- Built a Swiss air quality data stream with Apache Kafka, see [blog post](#)

**Data Engineer** **04.2021 — 07.2022**  
*Logitech* *Lausanne, Switzerland*

- Voice of Customer* project: Built an ETL workflow (vendor API → Snowflake) to integrate Logitech's online store customer reviews in the data warehouse, re-deployed project with Docker and Kubernetes in DEV/UAT/PROD environments for safety and maintainability, increased pool of analysed reviews by implementing language identification, updated Tableau dashboard and Snowflake data sources with sustainability data
- Implemented a CI/CD deployment pipeline for Logitech's internal data science models to manage Docker container runs on AWS
- Built time series data pipelines for CX forecasting models by leveraging data sources with different granularities
- Designed and built a COVID dataset to develop an incidence rate model predicting when to safely reopen Logitech's offices
- Participated in technical recruitment interviews, hosted knowledge-sharing sessions on technologies like Docker, GitHub Actions

**Data Engineer** **06.2019 — 03.2021**  
*Crédit Mutuel Leasing* *Paris, France*

- Implemented a Spark ETL on top of a Vertica database to perform anomaly detection and identify compliance issues
- Developed a data pipeline by merging two separate data sources (Access/SAS) and built a classification model for business introducers, improving accuracy by ~15 p.p with a random forest algorithm
- Co-led the internal data community to vulgarize data engineering and share best practices

**Math Teacher** **01.2014 — 03.2019**

**Software Engineering Intern** **02.2013 — 08.2013**  
*Hardis* *Nantes, France*

- Designed and developed the Android blood sugar tracker app <sup>©</sup>[MobiDiab](#) using a SCRUM agile framework

## EDUCATION

**Master in Statistics & Data Science**, *Paris Institute of Statistics, Sorbonne University* 2019 — 2020

**Master in Software Engineering**, *Institut Mines-Télécom, France* 2010 — 2013

**Bachelor of Mathematics**, *Rennes University, France* 2008 — 2010

## LANGUAGES

<b>French</b>	Native speaker
<b>English</b>	Full professional proficiency — TOEIC: 990/990

## REFERENCES

<b>Denis Zuba</b>	Head of Data Science at Logitech
<b>Paul Coudret</b>	Head of Data Science at Crédit Mutuel Leasing