Lead Data Engineer - DataOps Fundamentalist

Simon Thelin

Flat 33, Canada Wharf, 255, Rotherhithe Street, London, SE16 5ES, UK
✆ (+44) 07544 38 23 09
✉ simon.thelin90@gmail.com
🖥 www.linkedin.com/in/simon-thelin-3ba96986/

▬▬ Work Experience

Current - May 2022
Lead Data Engineer, *7bridges*, London, UK, **full time**.
- Leading efforts in end to end implementation of 7Bridges data platform, pushing it to the next level for analytics, AI and beyond.
- Reducing compute and storage costs over 100%.
- Reducing query latency up to 100%.
- Introducing new ways of learning and pushing efforts in hackathons to push innovation.
- Documented improvement of output of team members delivering projects 1-2 quarters faster than normal.
- Introducing DataOps workflows.
- Main focus on Lakehouse architecture for batch, streaming. Inspired but not exclusive to datamesh.
- Huge focus on self serve around data. Built on top of open source technologies such as delta, mlflow, spark on k8s, kafka, trino, dbt, airbyte.

Dec 2021 - May 2022
Lead Data Engineer, *Kainos*, London, UK, **Full Time Consultant**.
- Working on architecture and guidance to current data platforms on consultant basis.

Aug 2020 - Dec 2021
Data Engineer, *IMG Arena*, London, **full time**.
- Designed Lakehouse architecture based on Apache Druid, data processing pipelines both streaming and batch.
- Managed and maintained standards for Data Analysis tools used across the organisation.
- Engaging with sports data providers about source data quality.
- Contributing to future design/architecture solutions that directly impact the business.

June 2019 - Dec 2019
Data Engineer, *GameAnalytics*, London, **full time**.
- Improved current data pipeline built in PySpark and Hive to insert data into BigQuery.
- Built up the company data science platform. Including being responsible for best practices for data engineering and implementing data science features.
- Introduced Apache Airflow from scratch. Responsible for maintaining, teaching others (workshops) in how our schedule and workflow platform works in the data team, and how it can be used elsewhere.
- Worked closely together with data scientist to put statistical AB testing models into production. This included both data engineering and data science aspects.

Dec 2019 - June 2019
Data Engineer, *YoYo Wallet*, London, **full time**.
- Engineered large scale data modeling. Data processing in real time and batch with analytics.
- Managed the transitioning of the data warehouse from Postgres to utilise Redshift.

Aug 2018 - June 2019
Data Engineer, *Yieldify*, London, **full time**.
- Maintaining maintenance of Redshift with Alembic.
- Working closely with Data Analysts to deliver business needs both ad-hoc and modifying existing architecture.

August 2018 - Jan 2018
Developer Connected Vehicle Cloud, *Ericsson AB*, Gothenburg, Sweden, **full time**.
- Worked closely towards OEMs and OEMs delivery partners. Developing connectivity for the modern automotive ecosystem.
- Engineered micro-service infrastructure in Ericcsons cloud environment based on Apcera (similar to k8s).

Dec 2017- June 2017
Software Engineer, *Drivec AB*, Helsingborg, Sweden, **full time**.
- Performed analytical work against fleet management systems. Integration of GPRS/GPS and CAN hardware. Preparing data for machine learning solutions, such as using Hellinger Distance.

2011 June - 2014 Sep
Support & Field Test Engineer, *Drivec AB*, Helsingborg, Sweden, **full time**.
- Analyze CAN-bus data from vehicles.

▬▬ Education

2014 - 2017
BA of Science in Engineering, Computer Science & Engineering, *Halmstad University*, Sweden.
- ***Machine Learning thesis project where we delivered a cloud module called ITMS - MLA***[2]. Which was honored as the best thesis in the categories of Computer Science, Electronic, Mechatronic Engineering.
- Delivered a prototype for an ingredient check up app for allergy at the hackathon Hack4Health in April 2016.
- Attended education and preparations for 6 months to be able to be responsible for the team building process for new students at the introduction week at the IT section fall 2015.
- Representative during the open day in March 2017 for the IT section at Halmstad University.

2016 - 2017 **Exchange Student BA Computer Science & Engineering**, *Korea University*, KOR, 87.2/100 GPA.
- Attended courses in Computer networks, Information security, Software Engineering, Internet Programming together with international and Korean students.
- Developer for the anonymous location based messages application **www.klotter.co** together with international students, A+ received.

2012 - 2013 **Computer Science**, *Malmö University*, Sweden.

Leadership skills

- Leading small teams on large projects through mentor ship and guidance to solve complex business problems.
- Aligning with other squads/teams teaching new technologies.
- Taking on product management duties to deliver end to end data products.
- Pushing team members and companies to learn new areas of work such as infrastructure as code, end to end architecture sessions, understanding the importance of having skills across the board to deliver successful data projects.

Technical skills

- Programming:~**6 years: Python, ~1 year: Scala**
- SQL Syntax:***Various engines from Postgres to Trino to Druid, whatever it is, hands on it.***
- Infrastructure As Code:~**3 years: Terraform**
- Job Scheduler Platform:~**4 years: Apache Airflow, ~3 years: Native K8S, ~1 years: Argo Workflow, ~2 years: AWS Serverless**
- BI Tools:~**2 years: Apache Superset, ~2 years: Metabase, ~1 years: Tableau**
- Distributed Computing:~**1 year: Scala Spark, ~4 years: PySpark, ~1 years: Hive**
- Parallel Computing: ~**3 years: Pandas**
- Stream Processing: **(~3 years: Apache Kafka)**
- CI: ~**3 years: Jenkins , Travis, CircleCI, GitLab**
- CD: **(>4 years: Docker, Docker-Compose), >3 years: ArgoCD)**
- Main DB Technologies:**(Columnar: Trino, Druid, Redshift, BigQuery, Athena), (NoSQL: DynamoDB), (RDBMS: Postgres, MySQL)**
- Main Cloud Technology:**(>5 years: AWS), (<1 year: Google Cloud)**
- Certifications:**[Coursera - Machine Learning]**
- Publications ITMS - MLA[2]: **[Thesis - Diva Portal][IOP Science Publication]**

References

Conference Speech Of Lakehouse Architecture:**[DataNova 2023]**
Portfolio: **[GitHub]**
Honors & Awards: **[Scholarship - Best Thesis] [Scholarship - Sparbanken Kronan]**

---

[1]CAN - Controller Area Network
[2]ITMS MLA - Indirect Tire Monitoring System - Machine Learning Approach