Jurgen Limonard - Data Engineer

Profile

I am an experienced Data Engineer with a past in Business Intelligence as Data warehouse designer/developer. Working in IT since 1997. Worked for companies like Booking.com, Rabo, ING and VodafoneZiggo, mostly in scrum teams. Good analytical and communication skills. Presented knowledge sharing sessions and published IT related articles. As continuous learner I enjoy the big data ecosystem.

Skills

- Main: Spark, Python, PySpark, Scala, DBT, Linux, SQL
- Data Modelling: Dimensional modelling (Kimball), Data vault
- Cloud
    - AWS (Emr, Athena, Dms, Ecs, Fargate, Lambda)
    - GCP (BigQuery, Dataflow, Dataproc)
    - Databricks

Education

- MA (drs), Maatschappijgeschiedenis, Erasmus universiteit Rotterdam, [1986 - 1993] Thesis: ISBN13: 9789065501165. Nominated 'Hiskus-scriptieprijs' 1993/1994
- Oracle systeemontwerp & -bouw, Fast Lane, Maarssen, Bussum [aug. 1996 - feb. 1997]. SQL*Plus, PL/SQL, SQL*Forms 3.0, Developer/2000 (Oracle Forms 4.5, Reports 2.5), NIAM, Oracle 7 DBA, Designer/2000, Linux, ERM

Projects

Data Engineer Booking.com (2023-10 - 2025-10)

- Tools: Spark, PySpark, Airflow, Astronomer, Kubernetes, Snowflake, DBT, Linux, docker, Python, Hadoop, Hive, Oozie
- Built ETL-pipelines, gave presentations, support migration from on premise hadoop to cloud, Snowflake/DBT
- Optimized (Spark) workflows to significantly improve performance and stability. Extended monitoring and logging
- PoC to optimize Spark pipelines to significantly improve runtimes and reduce complexity.
- PoC for DBT on Snowflake; define standards, promote usage of DBT, presentations about DBT
- Support compliance (regulatory) team
- Data quality and validation checks: promoted via demo's and presentations (pyspark and DBT)

Data Engineer VodafoneZiggo (2021-01 - 2023-09)

- Tools: Hadoop, MapR, Spark, PySpark, scala, Hive, oozie, Airflow, Snowflake, Linux, docker, Python, AWS, Databricks
- Worked on Spark applications on Aws Emr, using PySpark and Scala
- Sources: daily or hourly dump files, Kafka streams and api-calls
- Set up Snowflake to Data Lake ingestion via a PoC and define standards and made application (Python, Aws Fargate)

Data Engineer Rabo Nederland (2017-07 - 2020-12)

- Tools: Hadoop, Spark, PySpark, Scala, Hive, Kafka, Airflow, Linux, Python, Nifi, Hortonworks, Cloudera, sqoop
- Built ETL pipelines, both batch and streaming. Worked with parquet, avro, orc, json
- Used hdfs, HBase and Hive as sink
- Data preparation via Google BigQuery. Used Databricks
- Gave knowledge sharing sessions about HBase
- Set up standards for the new Rabo Data Lake

Data Warehouse Developer Rabo Nederland (2015-11 - 2017-06)

- Built deployment automation CI/CD of Oracle and PowerCenter sources using Python. Huge time saver
- Built ETL-flows using Oracle PL/SQL, PowerCenter, Python

Data Warehouse Developer, Rabo International (2013-07 - 2015-10)
Data Warehouse Developer, Liberty Global (LGI, UPC) (2011-10 - 2013-07)
Data Warehouse Developer, Tele2 (BBNED) (2010-11 - 2011-09)
Data Warehouse Designer/Developer, ABN AMRO (Fortis) (2010-01 - 2010-10)
Data Warehouse Developer, Royal Bank of Scotland (2008-06 – 2009-12)
Data Warehouse Designer/Developer, UWV (2006-04 - 2008-06)
Data Warehouse Designer, ING (2005-05 - 2006-03)
Data Warehouse Designer/Developer, ABN AMRO (2005-02 - 2005-05)

Courses

- Terraform Associate 003, Udemy
- Neo4j Fundamentals (Neo4j, Cypher, Graph data modeling), Neo4j GraphAcademy
- Data Engineering on Google Cloud Platform (Dataproc, Dataflow, Pub/Sub), Coursera
- Functional Programming Principles in Scala, Coursera
- Data Vault (modelling), VanderLek Advies
- Dimensional modelling, VanderLek Advies

Contact

(6) 12 40 17 89
jlimonard@protonmail.com