

Jurgen Limonard - Data Engineer

Profiel

Ik ben een ervaren Data Engineer met een verleden in Business Intelligence als Datawarehouse designer/developer. Ik werk sinds 1997 in de IT. Ik heb gewerkt voor o.a Booking.com, Rabo, ING en VodafoneZiggo, meestal in scrumteams. Goede analytische en communicatieve vaardigheden. Heb kennisdelingssessies gegeven en IT-gerelateerde artikelen gepubliceerd. Continu blijven leren is één van de leuke kanten van het big data-ecosysteem.

Skills

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

Opleiding

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

Projecten

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

- Tools: Spark, PySpark, Airflow, Astronomer, Kubernetes, Snowflake, DBT, Linux, docker, Python, Hadoop, Hive, Oozie
- ETL-pipelines gemaakt, presentaties gegeven, ondersteuning bij migratie van hadoop onPrem naar de cloud
- Optimaliseren spark pipelines: significante verbetering doorlooptijd; complexiteit gereduceerd
- PoC voor DBT - Snowflake; definiëren nieuwe (ontwikkel) standaarden en starten nieuwe DBT flows
- Ondersteuning bij compliance (regulatory) programma's
- Data kwaliteit - en validatie checks: demo's en presentaties (pyspark, DBT)

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

- Tools: Hadoop, MapR, Spark, PySpark, scala, Hive, oozie, Airflow, Snowflake, Linux, docker, Python, AWS, Databricks
- Spark applicaties op Aws Emr, via PySpark en Scala
- Bronnen: dump files, Kafka streams en api-calls
- Opzetten Snowflake - Data Lake ingestie via een PoC en standaarden definiëren (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
- ETL pipelines gemaakt (batch en streaming). Formats als: parquet, avro, orc, json
- hdfs, HBase en Hive als opslag
- Data preparatie via Google BigQuery en Databricks
- Kennisdeel sessies gegeven over HBase
- Opzetten standaarden voor nieuwe Rabo Data Lake

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

- Maken CI/CD voor Oracle en PowerCenter sources met Python. Enorme time saver!
- Maken ETL-flows met 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)

Cursussen

- 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