LEONARDO SCALCO

DATA ENGINEER

CONTACT

+55 (45) 99933-1455

LEONARDO.SCALCO@HOTMAIL.COM

HTTPS://WWW.LINKEDIN.COM/IN/LEONARDO-FERNANDES-SCALCO/

PROFILE

Bachelor of Science in Computer Science from the State University of Western Paraná (Unioeste). Master degree in information visualization at University of Sao Paulo (USP). Worked in the development of web and mobile applications for regional advances in the west of Paraná, operated in the area of cyber security through research and development in the Itaipu Technological Park. After developing more than one hundred crawlers and scrapers, he realized the importance of data pipelines well built and organized, starting his carrer as data engineer on Serasa Experian. Nowadays, he works developing data pipelines, both batch or real-time, and with distributed and scalable processing.

SKILLS

- Kafka (managed and confluent)
- GCP Infrastructure: Functions, Big Query, DataProc, GKE, CloudSQL, Storage
- AWS Infrastructure: Lambda, RDS, EC2, Fargate, ECR, S3, SQS, ECS ...
- Hadoop Ecosystem: HDFS, Hbase, Hive, Spark
- Programming Languages: Python, Javascript, Scala, C, C++, Java
- Dataform and SQL
- Data Serialization: Avro, Parquet, Protobuf

DATA ENGINEER EXPERIENCE

SENIOR DATA ENGINEER
ADEVINTA (BARCELONA, SPAIN) | MAY 2022 - CURRENTLY

- Developed batch pipelines for different orchestration tools:
  - Airflow, Argo, Luigi
  - ETL
- Scala batch pipelines as service

SENIOR DATA ENGINEER
ISAAC (SAO PAULO, BRAZIL) | JAN 2021 - MAY 2022

- Developed different types of data pipelines:
  - Real-time pipelines using debezium (CDC), kafka and faust library to get data directly from databases
  - Batch pipelines through rest apis
  - Batch and real time pipelines using SaaS data integrations (Stitch and Fivetran)
- Managed ELT workflows using Big Query and Dataform
- Developed data products using Python, GKE, Rancher to empower Brazilian schools

SENIOR DATA ENGINEER
QUANTO (SAO PAULO, BRAZIL) | JUL 2020 - JAN 2021

- Realtime data pipelines with kafka and spark streaming for open banking events
- Data lake and data warehouse on the GCP infrastructure, using technologies such as Storage, Big Query and DataProc
- Developed open banking data products using technologies such as Protobuf, Python, Scala

DATA ENGINEER
SERASA EXPERIAN (SAO PAULO, BRAZIL) | SEP 2018 - JUL 2020

- Built several different crawlers and scrapers to collect unstructured web data by enriching Serasa Experian's databases
- Managed more than 2000 EC2 instances running in parallel
- Worked with Python, Scala, Hadoop, Big Data and Data Analysis to build solutions for getting data faster, more efficiently and up to date

EDUCATION

M.SC. IN COMPUTER SCIENCE
UNIVERSIDADE DE SAO PAULO | 2017 - 2020
INFORMATION VISUALIZATION:

B.SC. IN COMPUTER SCIENCE
UNIVERSIDADE ESTADUAL DO OESTE DO PARANA | 2013 - 2016

AWARDS

SERASA EXPERIAN
ONE EXPERIAN RECOGNITION - DATA STRATEGY - INDIVIDUAL CATEGORY
NOV 2019

SERASA EXPERIAN
ONE EXPERIAN RECOGNITION - DATA STRATEGY - GROUP CATEGORY
APR 2019