GILLES ESSOKI

Lead Data Engineer

@ robertondame@gmail.com   📞 0753060149   in www.linkedin.com/in/gillesessoki   📍 13 rue du bois Joly, 92000 Nanterre

EXPERIENCES

Tech Lead Data

Onepoint

📅 10 Oct 2023 – Present   📍 Paris, France

- Implement the construction of a retail search catalog with a BigQuery database as well as the set up of the server side tagging Cloud Run
- Gather knowledge on a legacy system with Azure Synapse and maintain the solution with DBT

Lead Data Engineer

Health Data Hub

📅 01 Janv 2022 – 06 Oct 2023 📍 Paris, France

- Work closely with cross-functional teams (product, technical management, etc.) and Data engineers to respond to various engineering issues
- Participate in the design and implementation of new solutions to meet the Hub's Data needs
- Support and help team members in their project

Senior Data Engineer

Health Data Hub

📅 Juillet 2020 – 31 Dec 2021   📍 Paris, France

- Ensure the proper reception of health databases (Documentation, compliance and consistency control, data preparation and availability) and training of juniors
- HUB technical expert for the omopisation of the SNDS (National Health Data System) in the OMOP-CDM (Observational medical outcomes partnership - Common Data Model) format.
- Implementation of applications and construction of libraries allowing to work on the technological platform of the HUB (hdhactions, siviciris, spark)

Data Engineer

Affini-Tech

📅 Dec 2016 – Avril 2020   📍 Issy-Les-Moulineaux, France

- Development of applications and APIs, implementation of structures to respond to infrastructure, code and science issues - Data Engineer Consultant
- POC on new elements resulting from technological monitoring with the aim of innovating (Ambassador, etc.)
- Development of the "Datatask" platform (Minikube, GKE, AKS)

SKILLS

- **Programming:** Python, Java, Golang, SQL, Bash, Rust, Elixir, Scala, JavaScript

PUBLICATIONS

Stackoverflow reputation: (457)   Activities
- Unable to run a basic GraphFrames example
- spark sql streaming cassandra
- spark structured stream: boontadata

CERTIFICATIONS

- Implementing Microsoft Azure Infrastructure Solutions
- Data Engineering, Big Data, and Machine Learning on GCP

LEAD

- Ensure that the team's performance meets technical standards, build a technical development culture to ensure the maintenance and continuity of the projects developed
- Supervise and promote the work of the team (poc, moc, etc...), build a culture of documentation
- Inspire, drive, develop team talent and foster a culture of growth and collaboration

SOME PROJECTS

CASINO

- Construction of hadoop clusters (dev, preprod and prod) with Mapr, secured by an openvpn and an openldap (double authentication: password and google authentication) and user access controls. The presence of a jupyterhub for the analytical part; all in the outscale cloud

Camaieu

- Porting of a brand discount application in the Google cloud with the particularity of event triggering (file in google storage => pubsub notifications => processing by a cloud function)

Solutions HDH

- Construction of a library allowing to perform data processing operations on the platform
- Construction of a package that automatically handles all the operation that are required for the provision of the data as well as extraction, pairing and targeting
- Study and configuration of spark on the platform (spark cluster, spark on K8S)

OMOP

- Implementation of the technical architecture (DBT, Postgres, Azure Storage)
- Configuration of validation tools (Achilles, Atlas, DataQuality Dashboard).

- **Tools/technology:** Kubernetes, Apache Kafka/Arrow/Airflow/Spark, Bases NoSQL, GCP/AWS/AZURE, Data Processing, Machine/Deep Learning, Gitlab CI/CD, gRPC

- **Interest:** Streaming, Open Source, Distributed Systems, Technical Innovation, Solution Improvement

---

EDUCATION

Computer and Network engineer

ESAIP

📅 Septembre 2013 – Mai 2016 📍 Angers, France