anankei
we support IT
staffing • sourcing • solutions

Data Engineer

| **Function type:** | Freelance - Vast | **Location:** | Brussels |
| **Duration:** | | **Reference:** | 202108482 |

Description:

For our customer we're looking for a Data Engineer.

In the first place we're looking for candidates with the "Anankei DNA".
Our DNA embodies a positive and joyful attitude.
Dream IT , Trust IT, Go for IT !

As a Data Engineer, you will play a key role preparing the infrastructure and data that will be used to deliver high quality products.
You will help the client design, develop and maintain data pipelines that will deliver insights.
By using a DevOps approach, you will make sure the overall system is running on all times by automating tasks so you can spend time on creating and not deploying.
You will also make sure the system is appropriately tested and monitored by using adapted methods and tools.

Here is a list of your resposibilities :

- Conceive and build data architectures
- Participate in the short/mid/long term vision of the overall system
- Execute ETL (extract/transform/load) processes from complex and/or large data sets
- Ensure data are easily accessible and that their exploitation is performing as requested, even in highly scalable circumstances
- Participate to the architecture and planning of the big data platform to optimize the ecosystem's performances
- Create large data warehouses fit for further reporting or advanced analytics
- Collaborate with machine learning engineers for the implementation and deployment of different solutions
- Ensure robust CI/CD processes are in place

Requirements:

What is important to have as background and skills ?

We are looking for strong candidates with the following academic and professional experiences :

A Master in Informatics, Engineering, Mathematics, or related field
Demonstrable experience with big data platforms (Hadoop, Cloudera, EMR, Databricks, ...)

Technical knowledge in:

- Data pipeline management
- Cluster management
- Workflow management (Oozie, Airflow)

- Database management of SQL and noSQL databases
- Large file storage (HDFS, Data Lake, S3, Blob storage,..)

Strong knowledge of :

- Hadoop ecosystem: Hortonworks/Cloudera/EMR
- Strong knowledge of Java/Scala and Python
- Strong knowledge of Spark (Scala and Pyspark)
- Strong knowledge of CI/CD concepts
- Stream processing such as Kafka, Kinesis, Elisticsearch
- Good knowledge of a cloud environment
- High level understanding of data science concepts
- Knowledge of Data Visualisation framework like QlikSense is a plus

Your Profile :

You're open minded , collaborative, team player, ready to adapt to the changing needs
You are multi-disciplined ability to work with divers APIs and understand multiple languages well enough to work with them
You're quality oriented
You are an excellent problem analyser and solver
Curiosity about new techniques and tools, eagerness to always keep learning
You're committed to deliver, pragmatic and solution oriented.
Experience with an agile way of working is a plus
Languages : English (very good in reading, writing, speaking) is a must!

Start: 2021-01-30 23:00:00