

100 % remote - Data Engineer/Analyst (m/w/d)

Freelance-ID: OCR-FL-1554

Task:

- Technical analysis of 4 data sources provided by BI (Boehringer Ingelheim) to identify how to obtain the relationships between KEE's
- Programming (R/Java languages) of 4 ETL's to load the KEE relationships into the database
- Design, creation and modification of the KEE database
- Manual testing of the ETL's
- Creation of the technical Documentation for the ETL's
- Tasks for the contractor (Developer role as part of the development team) will be documented and assigned in a Jira project. The Product Owner will be responsible of the creation and clarification of these tasks before they can be started. All data needed to perform the assigned tasks will be there
- As mentioned, the product is being developed in SCRUM, with a dedicated Scrum Master (Boehringer Ingelheim employee) and a Product Owner (Boehringer Ingelheim employee). Participation in the planned SCRUM (sprints of 2 weeks) ceremonies is required:

Responsibilities:

- Understanding of business goals and transformation
- Query and extraction of raw data from different data sources
- Building a data pipeline
- Building of an identification system for predicting the matching likelihood of HCPs based on metadata
- Definition of business rules and algorithms to ensure the unique identification of HCPs

Skills:

- Experience in NLP pipelines and semantic models
- Experience in the development of ETLs
- Experience in querying databases
- Experience handling semi-structured data (JSON, XML...)
- Skills in R and/or Python
- At least 3 years working as data engineer

Additional Information:

Location: remote

Start: 22.11.2021, latest during November 2021

Duration: February 2022

Utilization: 4-5 Tage / Woche

On-Site: -

Off-Site: 100 %

Your OCR contact

Laura Colon

phone: 07472-95176-17

E-mail: lc@oc-recruitment.de