mETal

Task 5: Abrosov, Anisin, Lanin, Malysh, Pan

Product description

Ensuring high data quality and availability is crucial for applications operating within our framework. Our extension for VS Code utilizes "data modules," which are stored either on Amazon S3 or in a relational database management system (RDBMS). The ETL (Extract, Transform, Load) service is designed to pull data from internal corporate systems, perform extensive cleansing, merging, anonymization, and transformation of the data, and subsequently upload it to create a new version of the data module.
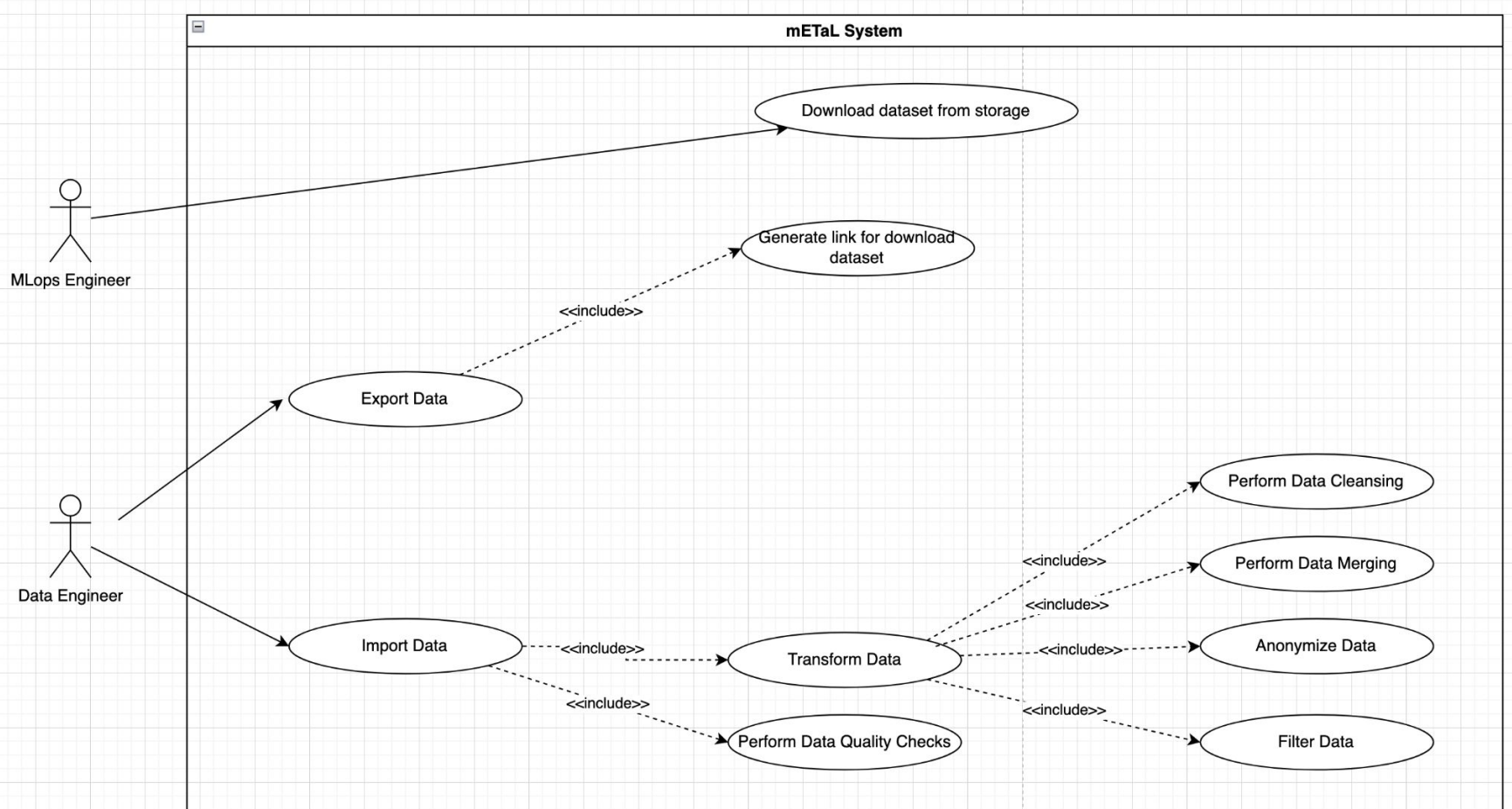
Team: Micro SD

Members:

- Abrosov Sergey
- Anisin Aleksandr
- Lanin Georg
- Malysh Igor
- Pan Zhengwu

Project repo:
https://github.com/abrosov-sergey/Micro-SD.git

This report: https://clck.ru/3Dbtvs

Use case diagram

Class candidates (problem domain)

Candidate	Criteria	Stored information	Operations
MLOps Engineer	SIAOT	Model data, training parameters, performance metrics	Load data
Data Engineer	SIAO(U?)T	Raw data, intermediate data, final data sets	Extract, transform, load data
Data	SIAOUT	Tables, records, data attributes	CRUD
Dataset	SIAOUT	Collections of data, metadata, data structures	CRUD

? == Optional

Domain model diagram