Component: Data Elasticity Management Processes Generator

1. Objective

This document describes the interfaces and inputs for the component Data Elasticity Management Processes Generator (DEP generator). The DEP generator is put here https://github.com/tuwiendsg/EPICS/blob/master/depic/depic-dep-process-generator/src/main/java/at/ac/tuwien/dsg/depic/process/generator/DataElasticityManagementProcessesGenerator.java

2. Interfaces

The DEP generator requires 3 inputs including a data analytics function (daf), a quality of result (qor) and a primitive action metadata (pam). While the daf and the qor are inputs of the constructor, the pam is loaded from its database.

```
Public DataElasticityManagementProcessesGenerator(
DataAnalyticsFunction daf, QoRModel qorModel) {
        this.daf = daf;
        this.qorModel = qorModel;
        config();
    }
```

3. Inputs

a. Quality of Results (QoR)

The presentation of the QoR model includes a set of metrics, a set of QoR elements (qElement) and a form of data asset. Fig. 1 shows the class diagram of QoR model. The class `QoRModel` has:

- A `listOfMetrics` measures quality of data assets. Each metric can be accessed/adjusted by primitive actions.
- A `listOfQElements`: A qElement is defined as a set of conditions established based on these metrics in specific `ranges` and a specific `price` for a data asset.
- A `dataAssetForm`: The form of data assets indicates the format of the output asset (e.g., comma- separated values or a bar chart) resulted from DAF.

Figure 2 presents an example and its explanations.

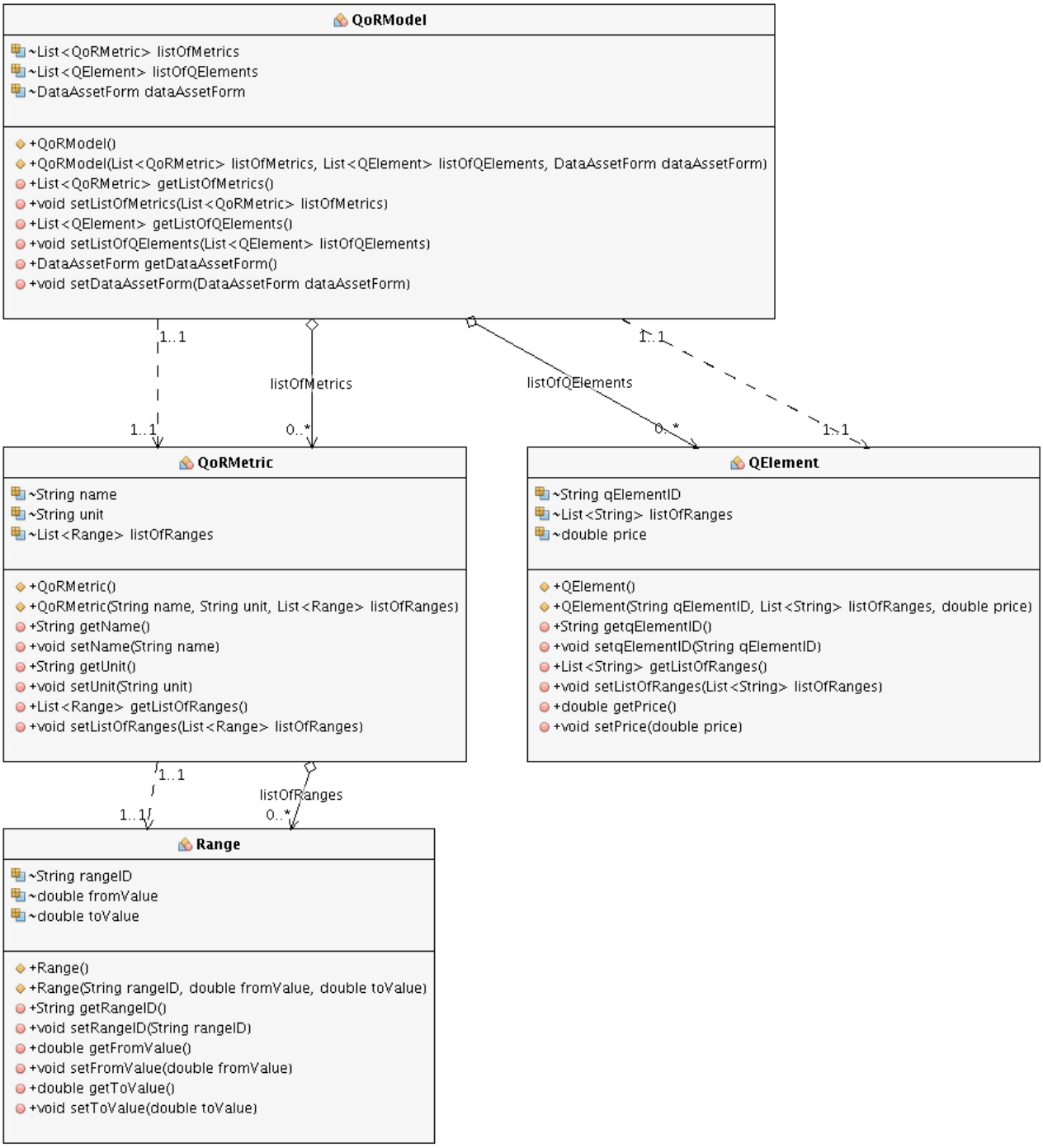

Figure 1: Class diagram of QoR Model

```
- !at.ac.tuwien.dsg.depic.common.entity.qor.QoRMetric
  name: vehicleArc
  listOfRanges:
  - !at.ac.tuwien.dsg.depic.common.entity.qor.Range
    rangeID: vehicleArc_co1
    toValue: 80.0
  - !at.ac.tuwien.dsg.depic.common.entity.qor.Range
    fromValue: 81.0
    rangeID: vehicleArc_co2
    toValue: 100.0
  unit: '%'
- !at.ac.tuwien.dsg.depic.common.entity.qor.QoRMetric
  name: deliveryTime
  listOfRanges:
  - !at.ac.tuwien.dsg.depic.common.entity.qor.Range
    fromValue: 54.0
    rangeID: deliveryTime_co1
    toValue: 1.7976931348623157E308
  - !at.ac.tuwien.dsg.depic.common.entity.qor.Range
    rangeID: deliveryTime_co2
    toValue: 53.0
  unit: s
listOfQElements:
- !at.ac.tuwien.dsg.depic.common.entity.qor.QElement
  listOfRanges:
  - speedArc_co2
  - vehicleArc_co2
  - deliveryTime_co2
  price: 0.009
  qElementID: qElement1
```

Expected ranges of values of QoR metric vehicleArc

Expected ranges of values of QoR metric deliveryTime

A qElement is a combination of ranges of QoR metrics and cost

Figure 2: An example of QoR

b. Data Analytics Function (DAF)

For provisioning data assets from the data sources, each data asset can be characterized by a data analytics function (DAF). The DaaS provider can provision data assets they want to sell by defining the function $f_{dataAsset}$.

Fig. 3 presents the class diagram of a data analytic function. The class `DataAnalyticsFunction` has:

- `name`: name of DAF
- `dataAssetForm`: form of data asset resulted from this DAF
- `dbType`: type of database is used to store data asset temporarily
- `daw`: this is the specification of DAF

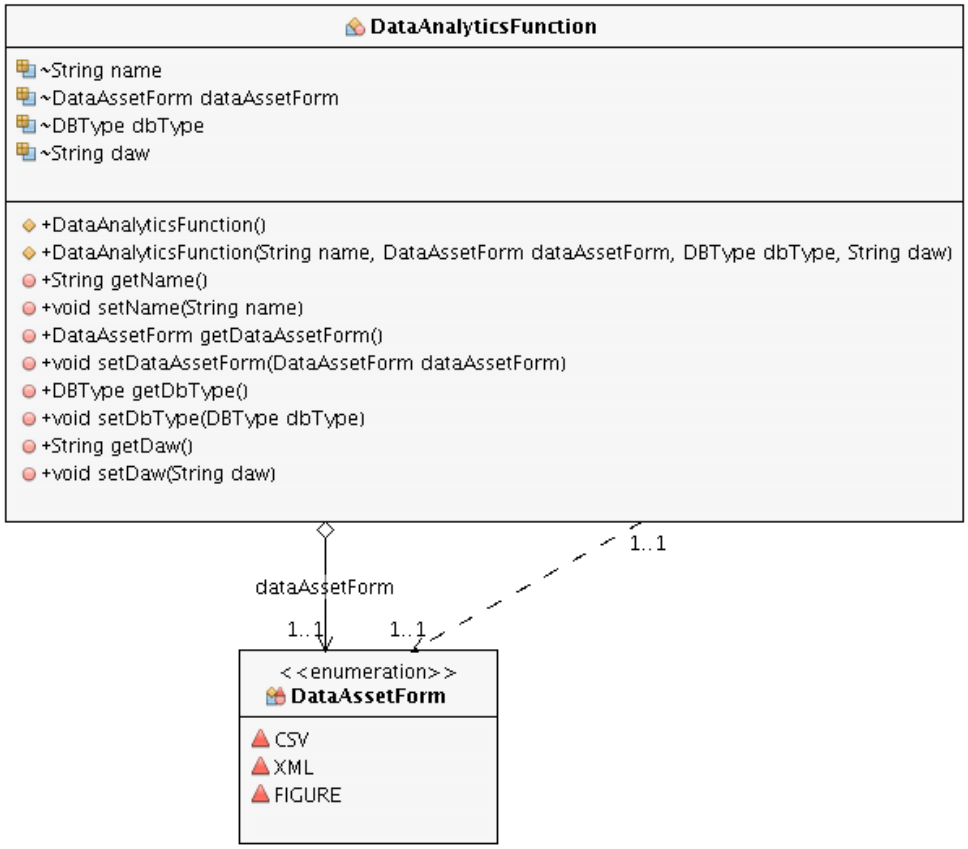

Figure 3: Class diagram of data analytics function

Fig. 4 presents an example of DAF. This DAF is a workflow of sequential activities, which are reading streaming GPS data, clustering vehicle location, estimating the average speed of vehicles in each cluster and outputting data in the form of comma-separated values (csv). Fig. 5 presents annotated information of analytic tasks in this DAF.

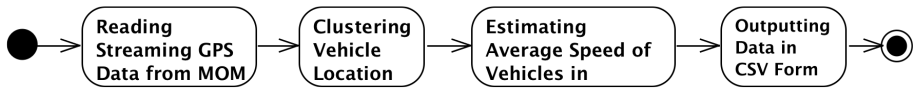

Figure 4: An example of data analytics function

```
                </net.sf.taverna.t2.annotation.AnnotationAssertionImpl>
            </annotationAssertions>
        </net.sf.taverna.t2.annotation.AnnotationChainImpl>
    </annotation_chain_2_2>
  </annotations>
</dataflow>
<depic>
  <AnalyticTask>
    <taskName>kmeans</taskName>
    <parameters>
      <parameterName>stopCondition</parameterName>
      <type>int</type>
      <value>5</value>
    </parameters>
  </AnalyticTask>
</depic>
```

annotated information of analytic tasks in daf

Figure 5: Annotated information of analytics task in DAF

c. **Primitive Action Metadata**

The document describing primitive action metadata is put here https://github.com/tuwiendsg/EPICS/blob/master/depic/docs/pamDocs.pdf