Project: Quality Assessment on Enterprise Linked Data
Authors: Rauf Agayev, Vugar Agayev, Carlos Montoya Sanchez
Date: 05.14.2014
Status: Ongoing

Introduction

During the development of Web Semantic technologies huge volume of data has being published on the web as a Linked Open Data. In order to use LOD for specific purpose, we should check its quality in advance. The quality term means in Semantic Web is fitness of use. Then we should have into account that not all the data is meaningful and if the quality is poor it leads to Enterprise problems such as data standardization, multiple data with duplicates in the data sets, meaningless information and so on.

To assure that the Enterprise is trustful, and then in the design we should take into account that the quality measurement depends on which domain the data will be used. On the other hand, we should assure that our LOD the data is measure for the correct quality metrics, which means that the given data in one LOD be useful for one of the use cases but not favorable for other ones. Therefore, for measurement of LOD we are going to use metrics and we identify the fitness of it use.

Who benefits?

Everyone who wants to use Linked Open Data should do quality assessment in advance. In our case metrics are domain independent. Thus, anyone who wants to use dataset which consist of RDF triples can use them for his/her datasets.

Functional Requirements.

1. Data Processing:

    1.1 Data Acquisition: System gets data from SPARQL endpoint and process it.

    Source/Initiator: Developers

    Priority: High

    1.2 Analyzing Data: Process the mentioned data and compute qualities

    Source/Initiator: Developers

    Priority: High

2. Data Visualization:

> Visualizing data: Show computed values to the user

> Source/Initiator: Developers and Users

> Priority: High

3. User interface

It consists of two parts: 1. Evaluating dataset which has been added to our system. User can select dimensions and metrics respect to those dimensions.

2. If user wants to evaluate new dataset he/she can select second window in main page and add data set there. After evaluation he/she will get mail about assessment.

Non Functional requirements

1. Usability
   The visualized information should be in such way that users can easily understand, and not in too much details
   Source/Initiator: Developers and Users
   Priority: Medium

2. Performance
   In case that there is no internet connection, the system is not up to date.
   Source/Initiator: Developers and Users
   Priority: high

3. Implementation Requirement
   For running the system some configurations must be done , JBoss server must be set up and should be started.
   Source/Initiator: Developers and Users
   Priority: high